Note di Algebra Lineare

Pino Vigna Suria

2 9 2003

Capitolo 1

Preliminari

1.1 Un po' d'insiemistica

La matematica moderna si fonda sui concetti di *elemento*, *insieme*, *appartiene* che non vengono definiti, confidando sul fatto che tutti ci intendiamo perfettamente sul loro significato. Tutti gli altri concetti devono essere definiti a partire da questi tre. La notazione per esprimere il fatto che un elemento x appartiene a un insieme X è $x \in X$. Un insieme può essere definito sostanzialmente in due modi.

1. Elencando tutti i suoi elementi, ad esempio $X = \{1, 5, \text{Tarzan}\}$.
2. Assegnando una proprietà soddisfatta da tutti e soli gli elementi dell'insieme, ad esempio

$$X = \{x \text{ tale che } x \text{ ha partecipato alla spedizione dei mille}\}.$$

Un'autentica pietra miliare della matematica, che si presenta fin da questo punto, è, una volta per tutte, la considerazione che segue

> Possiamo dire di conoscere un insieme A soltanto quando, comunque ci venga proposto un elemento x, siamo in grado di decidere se $x \in A$.

Un insieme importante è l' *insieme vuoto* che non ha elementi e viene indicato con il simbolo $\emptyset$, a lui riservato e che quindi non verrà usato in nessun altra circostanza. Lo conosciamo precisamente perché, messi di fronte ad un qualunque elemento x siamo in grado di decidere se $x \in \emptyset$, per la precisione la nostra risposta è sempre "no".

Si dice che un insieme A è *contenuto* in un insieme B o che B *contiene* A o che A è un *sottoinsieme* di B, se ogni elemento di A è elemento di B; si usa allora la notazione $A \subseteq B$ o $B \supseteq A$; in simboli

$$A \subseteq B \Longleftrightarrow \forall x, \quad x \in A \Longrightarrow x \in B$$

dove i simboli $\Longleftrightarrow, \forall, \Rightarrow$ si leggono rispettivamente *se e solo se, per ogni, implica*. Negare che A è contenuto in B vuol dire che esiste un elemento che appartiene a A ma non a B, in simboli $\exists x \in A$ tale che $x \notin B$, e si scrive $A \nsubseteq B$. A titolo di esempio: $\forall x \quad x \notin \emptyset$.

Vale la pena di precisare l'esatto significato che la matematica dà alla parola "implica "(o, naturalmente, al simbolo $\Rightarrow$): esso è sempre compreso tra due affermazioni, chiamiamole P, Q e fornisce una nuova affermazione $P \Rightarrow Q$;

$$P \Rightarrow Q \text{ è vera salvo quando } P \text{ è vera ma } Q \text{ è falsa.}$$

Alla luce di questa precisazione si ha che l'insieme vuoto è contenuto in ogni insieme e che due insiemi vuoti sono uguali. Questo giustifica l'uso dell'articolo determinativo che abbiamo finora usato abusivamente, imbrogliando la lettrice e costringendola ad un primo atto di umiltà.

Se A è un insieme e $\mathcal{A}$ è un insieme di sottoinsiemi di A (per non essere linguisticamente noiosi si dice che $\mathcal{A}$ è una *famiglia* di sottoinsiemi di A),

- l' *unione* di $\mathcal{A}$ è l'insieme

$$\bigcup_{B \in \mathcal{A}} B = \{a \in A \text{ tali che } \exists B \in \mathcal{A} \text{ tale che } a \in B\}.$$

- l' *intersezione* di $\mathcal{A}$ è l'insieme

$$\bigcap_{B \in \mathcal{A}} B = \{a \in A \text{ tali che } \forall B \in \mathcal{A},\, a \in B\}.$$

Definizione 1. *Siano a, b elementi. La* coppia ordinata *determinata da a e b è l'insieme*

$$(a, b) = \{\{a\}, \{a, b\}\}$$

Teorema 1. *Siano a, b, c, d elementi. sono equivalenti*

1. $a = c$ *e* $b = d$.

2. $(a, b) = (c, d)$.

Dimostrazione. È evidente che $1 \Rightarrow 2$. Per quanto riguarda l'altra implicazione chiamiamo rispettivamente S e D gli insiemi menzionati in 2.

$\{a\} \in S \subseteq D$ e quindi $\{a\} = \{c\}$ oppure $\{a\} = \{c, d\}$; nel primo caso $a = c$, nel secondo $\{c, d\} \subseteq \{a\}$ e quindi, di nuovo $c = a$.

Supponiamo $b = a$, allora, visto che $\{c, d\} \in D$ abbiamo che $\{c, d\} = \{a\}$ e quindi $d = a = b$.

Se invece $b \neq a$, visto che $\{a, b\} \in S$ avremo necessariamente che $\{a, b\} \subseteq \{c, d\}$. Siccome $b \neq c$ segue che $b = d$. □

Definizione 2. *Siano A, B insiemi; il loro* prodotto cartesiano *è l'insieme*

$$A \times B = \{(a, b) \text{ tali che } a \in A, b \in B\}$$

Definizione 3. *Siano A, B insiemi. Una* funzione *da A a B è coppia ordinata $((A, B), f)$ dove f è un sottoinsieme di $A \times B$ che soddisfa la seguente condizione:*

per ogni $a \in A$ esiste un unico $b \in B$ tale che $(a, b) \in f$. L'insieme A viene detto dominio *della funzione e B è il* codominio.

Se $A = \emptyset$ allora $A \times B = \emptyset$ e quindi l'unico suo sottoinsieme è l'insieme vuoto. Esso soddisfa la condizione imposta nelle definizione: abbiamo verificato che $((\emptyset, B), \emptyset)$ è l'unica funzione da $\emptyset$ a B. Viene chiamata la *vuota.*

Se invece $A \neq \emptyset$ e $B = \emptyset$ di nuovo avremo che $A \times B = \emptyset$ e l'unico suo sottoinsieme è l'insieme vuoto; tuttavia esso non soddisfa la condizione richiesta; cioè non esistono funzioni da un insieme non vuoto all'insieme vuoto.

Per indicare una funzione con dominio A e codominio B si usa la notazione $A \xrightarrow{f} B$ o semplicemente f se non c'è pericolo di confusione.

Se $a \in A$ l'unico elemento $b \in B$ tale che $(a, b) \in f$ viene chiamato *immagine* di a mediante f e indicato con il simbolo $f(a)$ o, alle volte, f_a; quindi $b = f(a)$ è sinonimo di $(a, b) \in f$.

$\mathrm{id}_A = \{(a, b) \in A \times A \text{ tali che } a = b\}$ è chiamata la *identica* o *identità* di A; con la notazione appena introdotta tale funzione è dunque caratterizzata da $\mathrm{id}_A(a) = a \; \forall a \in A$.

Se abbiamo due funzioni $A \xrightarrow{f} B$ e $B \xrightarrow{g} C$ definiamo la loro *composizione* come la funzione $A \xrightarrow{g \circ f} C$ definita da $g \circ f(x) = g(f(x))$. È facile vedere che se f, g, h sono funzioni, allora $h \circ (g \circ f) = (h \circ g) \circ f$ a condizione che tali

composizioni abbiano senso. Questa regola si chiama *proprietà associativa della composizione*. Se le composizioni hanno senso è facile controllare che $f \circ \mathrm{id}_A = f = \mathrm{id}_A \circ f$. Una funzione $A \xrightarrow{f} B$ viene detta *iniettiva* se $\forall a, a' \in A \quad f(a) = f(a') \Rightarrow a = a'$ e viene detta *suriettiva* se $\forall b \in B \quad \exists a \in A$ tale che $f(a) = b$. Viene detta *biettiva* o *corrispondenza biunivoca* se è iniettiva e suriettiva. Se è così esiste una funzione, indicata con il simbolo $B \xrightarrow{f^{-1}} A$, tale che $f^{-1} \circ f = \mathrm{id}_A$ e $f \circ f^{-1} = \mathrm{id}_B$.

Sia $n \in \mathbb{N}$ un numero naturale, indichiamo con il simbolo $\overline{n}$ l'insieme $\overline{n} = \{1, \ldots, n\}$ costituito da tutti i numeri compresi tra 1 e n, (osserviamo che $\overline{0} = \emptyset$).

Se A è un insieme una $n-$*upla ordinata* di elementi di A è una funzione $\mathcal{A} : \overline{n} \longrightarrow A$, cioè un elemento di $A^{\overline{n}}$. Con un piccolo ma universalmente adoperato risparmio notazionale indichiamo quest'ultimo insieme con A^n. Una notazione molto accattivante ed universalmente usata è quella di denotare una $n-$upla ordinata $\mathcal{A} \in A^n$ con il simbolo $(a_1, \ldots, a_n)$. Nei casi $n = 3, 4$ si usa, per motivi storici, la terminologia particolare di *terne ordinate* e *quaterne ordinate*. Naturalmente l'unica $0-$upla ordinata è la funzione vuota; le $1-$uple ordinate sono semplicemente gli elementi di A nel senso che adesso precisiamo: la funzione $A^1 \longrightarrow A$ che ad ogni $\mathcal{A} \in A^1$ associa l'elemento $a_1 \in A$ è una corrispondenza biunivoca che viene sistematicamente utilizzata per identificare A^1 con A.

Similmente la funzione che ad ogni $\mathcal{A} \in A^2$ associa la coppia ordinata $(a_1, a_2) \in A \times A$ è una corrispondenza biunivoca e la si adopera per identificare A^2 con $A \times A$; attraverso tale corrispondenza si identifica ogni $2-$upla ordinata di elementi di A con un'opportuna coppia ordinata di elementi di A.

1.2 Strutture algebriche

Definizione 4. *Se A, B, C sono insiemi un'* operazione *su A e B a valori in C è una funzione $A \times B \xrightarrow{\heartsuit} C$. Si parla semplicemente di operazione* interna *su A quando $A = B = C$.*

Se $(a, b) \in A \times B$ si preferisce denotare l'elemento $\heartsuit((a, b))$ con il molto più succinto $a \heartsuit b$.

Definizione 5. *Un* gruppo *è una coppia ordinata $(G, +)$ dove G è un insieme e $+$ è un'operazione interna su G che soddisfa i seguenti assiomi:*

G1 *Per ogni* $f, g, h \in G \quad f+(g+h) = (f+g)+h$. *(*proprietà associativa.*)*

G2 *esiste* $e \in G \quad$ *tale che* $e+f = f+e = f \quad \forall f \in G$. *Ogni elemento con questa proprietà viene detto* elemento neutro.

G3 *Per ogni* $g \in G$ *esiste* $h \in G \quad$ *tale che* $g + h = h + g = e$. *Ogni elemento con questa proprietà viene detto* opposto *di* g *e indicato con il simbolo* $-g$.

Nella definizione che precede le parentesi (e) sono state usate con il significato (di precedenza) che la lettrice ben conosce.

Esiste un solo elemento neutro, infatti se $e, e' \in G$ sono elementi neutri allora $e = e + e' = e'$. Ogni $g \in G$ ha un solo opposto, infatti se $h, h' \in G$ sono opposti di g allora $h = h+e = h+(g+h') = (h+g)+h' = e+h' = h'$.

Definizione 6. *Un gruppo* $(G, +)$ *è detto* commutativo *o* abeliano *se* $\forall g, h \in G \quad g + h = h + g$.

Se X è un insieme indichiamo con $\mathcal{S}(X)$ l'insieme di tutte le biezioni $f : X \longrightarrow X$. Allora $(\mathcal{S}(X), \circ)$ è un gruppo: l'elemento neutro é id_X e l'inverso di $g \in \mathcal{S}(X)$ è g^{-1}. Questo gruppo è commutativo se e sono se X non ha più di due elementi: la lettrice è incoraggiata a verificarlo.

Questo è l'unico esempio concreto che possiamo produrre in questo momento, ma chi legge si sarà subito accorto che l'insieme dei numeri relativi e l'operazione di somma su cui la maestra delle elementari ha così a lungo commentato costituiscono un altro esempio di gruppo abeliano.

Definizione 7. *Un* anello *è una coppia ordinata* $(A, (+, \cdot))$ *dove* A *è un insieme e* $(+, \cdot)$ *è una coppia ordinata di operazioni interne su* A *che soddisfano i seguenti assiomi:*

A1 *Per ogni* $f, g, h \in A \quad f+(g+h) = (f+g)+h$. *(*proprietà associativa.*)*

A2 *Esiste* $0 \in A \quad$ *tale che* $0 + f = f + 0 = f \quad \forall f \in A$.

A3 *Per ogni* $g \in A$ *esiste* $h \in G \quad$ *tale che* $g + h = h + g = 0$.

A4 $\forall g, h \in A \quad g + h = h + g$.

A5 *Per ogni* $f, g, h \in A \quad f \cdot (g \cdot h) = (f \cdot g) \cdot h$.

A6 *Per ogni* $f, g, h \in A \quad f \cdot (g + h) = (f \cdot g) + (f \cdot h)$ *e* $(f + g) \cdot h = (f \cdot h) + (g \cdot h)$.

I primi quattro assiomi ci dicono che $(A, +)$ è un gruppo abeliano. Le due operazioni su un anello vengono chiamate rispettivamente *somma* e *prodotto*. Se $f, g \in A$ l'elemento $f \cdot g$ viene indicato semplicemente con fg. Si usa, per alleggerire la notazione, l'usuale convenzione di dare la precedenza al prodotto sulla somma: a titolo di esempio $fg + h$ sta per $(fg) + h$.

Qui non abbiamo al momento degli esempi che dipendano solo da quanto abbiamo detto in queste note, ma la lettrice si è certamente accorta che, con le operazioni di cui ci hanno parlato prima la maestra e poi i professori delle medie e delle superiori, i numeri interi, quelli razionali e quelli reali sono tutti anelli.

In ogni anello 0 è *nullificatore del prodotto*, cioè, $\forall g \in A \quad g0 = 0$ e $0g = 0$; infatti, partendo da $0 + 0 = 0$ e usando l'assioma $A6$ si vede che $0g = (0 + 0)g = 0g + 0g$, sia h l'opposto di $0g$, allora $0 = h + 0g = h + (0g + 0g) = (h + 0g) + 0g = 0 + 0g = 0g$. Analogamente per l'altra uguaglianza.

Definizione 8. *Un* campo *è una coppia ordinata* $(\mathbb{K}, (+, \cdot))$ *dove* $\mathbb{K}$ *è un insieme e* $(+, \cdot)$ *è una coppia ordinata di operazioni interne su* $\mathbb{K}$ *che soddisfano i seguenti assiomi:*

C1 *Per ogni* $f, g, h \in \mathbb{K} \quad f + (g + h) = (f + g) + h$. *(*proprietà associativa.*)*

C2 *Esiste* $0 \in \mathbb{K} \quad$ *tale che* $0 + f = f + 0 = f \quad \forall f \in \mathbb{K}$.

C3 *Per ogni* $g \in \mathbb{K}$ *esiste* $h \in G \quad$ *tale che* $g + h = h + g = 0$.

C4 $\forall g, h \in \mathbb{K} \quad g + h = h + g$.

C5 *Per ogni* $f, g, h \in \mathbb{K} \quad f(gh) = (fg)h$.

C6 *Per ogni* $f, g, h \in \mathbb{K} \quad f(g + h) = fg + fh$.

C7 *Per ogni* $f, g \in \mathbb{K} \quad fg = gf$.

C8 *Esiste* $1 \in \mathbb{K} \quad$ *tale che* $1f = f \quad \forall f \in \mathbb{K}$.

C9 *Per ogni* $0 \neq g \in \mathbb{K}$ *esiste* $g^{-1} \in \mathbb{K}$ *tale che* $gg^{-1} = 1$.

C10 $1 \neq 0$.

I primi sei assiomi ci assicurano che ogni campo è un anello. L'insieme dei numeri interi con le operazioni di cui ci ha detto la maestra è un anello ma non un campo. Lo indicheremo sistematicamente con il simbolo, a lui riservato, $\mathbb{Z}$. Invece i numeri razionali e quelli reali, con le solite operazioni sono entrambi dei campi, indicati rispettivamente con $\mathbb{Q}$ e $\mathbb{R}$.

La definizione rigorosa di tali insiemi e delle rispettive operazioni, come anche dei numeri naturali (indicati con $\mathbb{N}$), non è affatto semplice ed in queste note li useremo, per la verità con estrema parsimonia, usando la confidenza che la lettrice ha acquisito con essi attraverso anni di convivenza. Chi non gradisce questo atteggiamento può consultare

`http://alpha.science.unitn.in/~vigna/numeri.dvi`

ma consigliamo di acquisire un pochino di confidenza con i metodi della matematica moderna prima di affrontarne la lettura.

Un esempio autonomo può essere ottenuto come segue: prendiamo due elementi qualunque che indicheremo, per motivi che saranno subito evidenti, con i simboli 0 e 1. Sull'insieme $\mathbb{K} = \{0, 1\}$ definiamo somma e prodotto mediante le tabelle che seguono, da interpretarsi come alle scuole elementari.

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Un po' di pazienza consente di verificare che i dieci assiomi sono tutti soddisfatti. Questo campo, ovviamente il più semplice possibile, ha una sua importanza specialmente in crittografia. Lo indicheremo con il simbolo $\mathbb{Z}_2$.

Ogni campo ha ovviamente le proprietà elementari dei gruppi e degli anelli, a cui aggiungiamo la seguente: se a, b sono elementi di un campo e $ab = 0$ allora $a = 0$ o $b = 0$; infatti se $a \neq 0$ allora $0 = a^{-1}0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$.

D'ora in poi denoteremo un campo $(\mathbb{K}, (+, \cdot))$ semplicemente con $\mathbb{K}$; questo per semplificare la notazione. Analogamente per gli anelli ed i gruppi.

1.3 Numeri complessi

Poniamo $\mathbb{C} = \mathbb{R}^2$ e definiamo su questo insieme un'operazione interna, detta *somma* mediante

$$(a, b) + (c, d) = (a + c, b + d)$$

e un'operazione interna, detta *prodotto* mediante

$$(a,b)(c,d) = (ac - bd, ad + bc)$$

Gli elementi di $\mathbb{C}$ sono chiamati *numeri complessi*; con un po' di pazienza si verifica che $(\mathbb{C}, (+,\cdot))$ è un anello. La coppia ordinata $(1,0)$ è elemento neutro rispetto al prodotto, inoltre, se $(a,b) \neq (0,0)$ si vede subito che $(a,b)(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}) = (1,0)$.

Quindi, con queste operazioni, $\mathbb{C}$ è un campo. La funzione $\phi : \mathbb{R} \longrightarrow \mathbb{C}$ data da $\phi(a) = (a,0)$ è inettiva, inoltre $\phi(a+b) = \phi(a) + \phi(b)$ e $\phi(ab) = \phi(a)\phi(b)$ e questo ci permette di pensare ogni numero reale come un numero complesso. La lettrice non si scandalizzi per questo piccolo ed utilissimo abuso: è lo stesso che le consente di interpretare, se lo desidera, un numero intero come se fosse un razionale.

Cerchiamo i numeri complessi (x,y) tali che $(x,y)^2 = -1$, cioè tali che $(x^2 - y^2, 2xy) = (-1,0)$; è facile convincersi che $x = 0$ e $y = \pm 1$. Il numero complesso $(0,1)$ viene indicato con il simbolo i; allora, se $a, b \in \mathbb{R}$

$$a + ib = (a,0) + (0,1)(b,0) = (a,0) + (0,b) = (a,b)$$

che ci fornisce una notazione eccellente per i numeri complessi che d'ora in poi verrà adottata sistematicamente.

Se $z = a + ib \in \mathbb{C}$ il suo *complesso coniugato* è il numero complesso $\overline{z} = a - ib$ ed il suo *modulo* è il numero reale non negativo $|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}$.

Si vede con irrisoria facilità che, se $z, w \in \mathbb{C}$ allora $\overline{z+w} = \overline{z} + \overline{w}$ e che $\overline{zw} = \overline{z}\,\overline{w}$.

Un numero complesso z è reale se e solo se $z = \overline{z}$.

Il motivo principale che consiglia l'introduzione dei numeri complessi è il seguente

Teorema 2 (fondamentale dell'algebra). *Ogni polinomio non costante a coefficienti in* $\mathbb{C}$ *ha almeno una radice in* $\mathbb{C}$.

La dimostrazione è fuori dalla nostra portata.

Capitolo 2

Spazi vettoriali

2.1 Spazi vettoriali

Definizione 9. *Siano* $\mathbb{K}$ *un campo,* $(V, (+, \cdot))$ *una coppia ordinata dove* V *è un insieme* $+$ *è un'operazione interna su* V*, detta* somma *e* $\cdot : \mathbb{K} \times V \longrightarrow V$ *è un'operazione su* $\mathbb{K}$ *e* V *a valori in* V*, detta* operazione esterna*; si dice che* $(V, (+, \cdot))$ *è uno* spazio vettoriale *su* $\mathbb{K}$ *se valgono i seguenti assiomi*

SV1 $\forall v, w, u \in V \quad (v + w) + u = v + (w + u)$.

SV2 $\exists 0 \in V$ *tale che* $0 + v = v + 0 = v, \quad \forall v \in V$.

SV3 $\forall v \in V \, \exists w \in V$ *tale che* $v + w = w + v = 0$. *Tale* w *viene chiamato l'* opposto v *e indicato con il simbolo* $-v$*; inoltre, se* $u \in V$*, l'elemento* $u + (-v)$ *verrà indicato con* $u - v$.

SV4 $\forall v, w \in V \quad v + w = w + v$.

SV5 $\forall \lambda, \mu \in \mathbb{K} \quad \forall v \in V \quad \lambda(\mu v) = (\lambda\mu)v$.

SV6 $\forall \lambda, \mu \in \mathbb{K} \quad \forall v \in V \quad (\lambda + \mu)v = \lambda v + \mu v$.

SV7 $\forall \lambda \in \mathbb{K} \quad \forall v, w \in V \quad \lambda(v + w) = \lambda v + \lambda w$.

SV8 $\forall v \in V \quad 1v = v$.

Se $(V, (+, \cdot))$ è uno spazio vettoriale gli elementi di V vengono chiamati *vettori* mentre gli elementi di $\mathbb{K}$ sono detti *scalari*. L'elemento 0 citato in

SV2 viene chiamato il *vettore nullo*. Se non c'è pericolo di confusione uno spazio vettoriale $(V, (+, \cdot))$ verrà indicato semplicemente con V.

Ecco qualche esempio: sia $V = \{a\}$ un insieme con un solo elemento a, e definiamo le due operazioni nel solo modo possibile. Si vede facilissimamente che si ottiene uno spazio vettoriale. In particolare l'elemento 0 menzionato nell'assioma *SV2* deve essere a, e quindi $V = \{0\}$; questo semplicissimo spazio vettoriale viene chiamato lo *nullo*.

Un altro spazio vettoriale, talmente generale che lo chiameremo *supere-sempio*, può essere ottenuto come segue: prendiamo un qualunque insieme X e denotiamo con il simbolo $\mathbb{K}^X$ l'insieme di tutte le funzioni $X \xrightarrow{f} \mathbb{K}$; su questo insieme definiamo le operazioni come segue

$$(f+g)(x) = f(x) + g(x), \forall f, g \in \mathbb{K}^X, \forall x \in X,$$

$$(\lambda f)(x) = \lambda f(x), \forall \lambda \in \mathbb{K}\ \forall g \in \mathbb{K}^X, \forall x \in X.$$

Di nuovo è facile vedere che $\mathbb{K}^X$ è uno spazio vettoriale.

Adesso diamo un altro esempio che si ottiene scegliendo opportunamente l'insieme X di cui si parla nel paragrafo precedente. Conviene illustrare l'argomento in modo leggermente più generale.

Se $\mathbb{K}$ è un campo, abbiamo già visto come assegnare le operazioni su $\mathbb{K}^n$, ma l'eccezionale importanza pratica di questo esempio ci suggerisce di tradurre nella nuova notazione le operazioni che abbiamo definito: avremo dunque, per ogni $(x_1, \ldots, x_n) \in \mathbb{K}^n$, $(y_1, \ldots, y_n) \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Ogni spazio vettoriale gode delle seguenti proprietà:

1. Il vettore nullo è unico. Siano indatti $0, 0'$ elementi di V che soddisfano entrambi la regola *SV2*, allora $0 = 0 + 0' = 0'$.

2. Se $v \in V$, l'opposto di v è unico. Siano infatti $w, w' \in V$ vettori che soddisfano entrambi la regola *SV3,* allora $w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'$.

3. $\forall v \in V \quad 0v = 0$. Infatti $0v = (0 + 0)v = 0v + 0v$ da cui segue che $0 = 0v - 0v = (0v + 0v) - 0v = 0v + (0v - 0v) = 0v$.

4. $\forall \lambda \in \mathbb{K} \quad \lambda 0 = 0$. Infatti $\lambda 0 = \lambda(0+0) = \lambda 0 + \lambda 0$ da cui segue che $0 = \lambda 0 - \lambda 0 = (\lambda 0 + \lambda 0) - \lambda 0 = \lambda 0 + (\lambda 0 - \lambda 0) = \lambda 0$.

5. Se $\lambda v = 0$ allora $\lambda = 0$ oppure $v = 0$. Sia infatti $\lambda \neq 0$ allora $0 = \lambda^{-1} 0 = \lambda^{-1}(\lambda v) = (\lambda^{-1}\lambda)v = 1v = v$.

6. $\forall v \in V \quad (-1)v = -v$. Infatti $(-1)v + v = (-1)v + 1v = (-1+1)v = 0v = 0$.

2.2 Sottospazi vettoriali

In tutta questa sezione supporremo di aver scelto una volta per tutte un campo $\mathbb{K}$.

Definizione 10. *Sia V uno spazio vettoriale e sia $W \subseteq V$. Si dice che W è un* sottospazio vettoriale di V *e si scrive $W \lhd V$ se valgono le condizioni*

S1 $0 \in W$.

S2 $\forall v, w \in W \quad v + w \in W$.

S3 $\forall \lambda \in \mathbb{K}$, $\forall v \in W \quad \lambda v \in W$.

In particolare se $W \lhd V$ allora $(W, +, \cdot)$ è uno spazio vettoriale.

Si può controllare facilmente che $W = \{0\}$ e $W = V$ sono sottospazi vettoriali, ma si possono fornire altri esempi non banali: sia $W = \{(x, y, z) \in \mathbb{R}^3$ tali che $2x - y + 5z = 0\}$; si vede facilmente che $W \lhd \mathbb{R}^3$.

Se U, W sono sottospazi vettoriali di V anche la loro *intersezione* e cioè l'insieme $U \cap W = \{v \in V$ tale che $v \in U$ e $v \in W\}$ lo è. Un altro esempio ci viene fornito dalla seguente

Definizione 11. *Siano V uno spazio vettoriale e U, W sottospazi vettoriali di V. La* somma *di U e W è l'insieme*

$$U + W = \{v \in V \text{ tali che esistono } u \in U \text{ e } w \in W \text{ tali che } v = u + w\}$$

Non è difficile verificare che $U + W$ gode delle seguenti proprietà:

1. $U + W$ è un sottospazio vettoriale di V.

2. $U \subseteq U + W$ e $W \subseteq U + W$.

3. Se $M \lhd V$, $U \subseteq M \quad W \subseteq M$ allora $U + W \subseteq M$.

In lingua $U+W$ è il più piccolo sottospazio vettoriale di V che contiene sia U che W. Analogamente si vede che $U \cap W$ è il più grande sottospazio vettoriale di V contenuto sia in U che in W.

2.3 Matrici

Anche in questa sezione $\mathbb{K}$ è un campo scelto una volta per tutte.

Definizione 12. *Siano m, n numeri naturali positivi. Una* matrice $m \times n$ *(a coefficienti in $\mathbb{K}$) è una funzione $A : \overline{m} \times \overline{n} \longrightarrow \mathbb{K}$.*

La notazione accattivante, anche se un tantino contradditoria con altre convenzioni in auge, per una matrice è

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

L'insieme di $\mathbb{K}^{\overline{m} \times \overline{n}}$ di tutte le matrici $m \times n$ viene indicato con $\mathcal{M}(m \times n; \mathbb{K})$.

Se $A \in \mathcal{M}(m \times n; \mathbb{K})$ la *trasposta* di A è la matrice $B \in \mathcal{M}(n \times m; \mathbb{K})$ definita da $b_{ji} = a_{ij}$. Essa viene indicata con il simbolo A^t, ovviamente $(A^t)^t = A$. Una matrice $m \times 1$ viene chiamata *vettore colonna* e una matrice $1 \times n$ viene chiamata *vettore riga*.

Se $A : \overline{1} \times \overline{n} \longrightarrow \mathbb{K}$ è un vettore riga, componendolo con la funzione, a cui non vale la pena di dare un nome, $\overline{n} \longrightarrow \overline{1} \times \overline{n}$ che manda j nella coppia ordinata $(1, j)$ si ottiene una $n-$upla ordinata, cioè un elemento di $\mathbb{K}^n$. Si stabilisce in questo modo una corrispondenza biunivoca tra $\mathcal{M}(1 \times n, \mathbb{K})$ e $\mathbb{K}^n$ che chiamiamo *isomorfismo canonico* ed useremo quando necessario per identificare questi due spazi vettoriali.

Per ogni $i = 1, \dots, m$ denotiamo con il simbolo $\alpha_i : \overline{1} \times \overline{n} \longrightarrow \overline{m} \times \overline{n}$ la funzione data da $\alpha_i(1, j) = (i, j)$. Se $A \in \mathcal{M}(m \times n, \mathbb{K})$ la sua composizione con α_i ci dà un vettore riga che si chiama la $i-$*esima riga* di A e si indica con A_i. In sintesi $A_i = (a_{i1}, \dots, a_{in})$.

Analogamente, per ogni $j = 1, \dots, n$ denotiamo con $\beta_j : \overline{m} \times \overline{1} \longrightarrow \overline{m} \times \overline{n}$ la funzione data da $\beta_j(i, 1) = (i, j)$. Se $A \in \mathcal{M}(m \times n, \mathbb{K})$ la sua composizione

con β_j ci dà un vettore colonna che si chiama la $j-$*esima colonna* di A e si indica con A^j. In sintesi

$$A^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = (a_{1j}, \ldots, a_{mj})^t.$$

Ovviamente, equipaggiando $\mathcal{M}(n \times m; \mathbb{K})$ con le operazioni descritte nel superesempio, si ottiene uno spazio vettoriale. Con la notazione appena introdotta $A + B$ è la matrice C definita da $c_{ij} = a_{ij} + b_{ij}$, più brevemente $(a + b)_{ij} = a_{ij} + b_{ij}$.

Se $\lambda \in \mathbb{K}$ e $A \in \mathcal{M}(n \times m; \mathbb{K})$, $\lambda A \in \mathcal{M}(n \times m; \mathbb{K})$ è definita da $(\lambda a)_{ij} = \lambda a_{ij}$.

Inoltre

$$\forall A \in \mathcal{M}(m \times n; \mathbb{K}),\ B \in \mathcal{M}(m \times n; \mathbb{K}) \quad (A + B)^t = A^t + B^t.$$

Se $A \in \mathcal{M}(m \times n; \mathbb{K})$ e $B \in \mathcal{M}(n \times p; \mathbb{K})$ definiamo il loro *prodotto* AB come la matrice $C \in \mathcal{M}(m \times p; \mathbb{K})$ definita da

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

dove $i = 1, \ldots, m$ e $k = 1, \ldots, p$. Il prodotto di matrici gode delle seguenti proprietà, tutte facili da dimostrare meno la $MP1$; ad ogni buon conto sono molto semplici da ricordare:

MP1 $\forall A \in \mathcal{M}(m \times n; \mathbb{K}),\ B \in \mathcal{M}(n \times p; \mathbb{K}),\ C \in \mathcal{M}(p \times q; \mathbb{K})$

$$A(BC) = (AB)C.$$

MP2 $\forall A \in \mathcal{M}(m \times n; \mathbb{K}),\ B \in \mathcal{M}(n \times p; \mathbb{K}),\ \lambda \in \mathbb{K}$

$$A(\lambda C) = (\lambda A)C = \lambda(AC).$$

MP3 $\forall A \in \mathcal{M}(m \times n; \mathbb{K}),\ B \in \mathcal{M}(n \times p; \mathbb{K}),\ C \in \mathcal{M}(n \times p; \mathbb{K})$

$$A(B + C) = AB + AC.$$

MP4 $\forall A \in \mathcal{M}(m \times n; \mathbb{K})$, $B \in \mathcal{M}(m \times n; \mathbb{K})$, $C \in \mathcal{M}(n \times p; \mathbb{K})$

$$(A + B)C = AC + BC.$$

MP5 $\forall A \in \mathcal{M}(m \times n; \mathbb{K})$, $B \in \mathcal{M}(n \times p; \mathbb{K}) \quad (AB)^t = B^t A^t$.

Notiamo che, ponendo in $MP3$ $B = C = 0$ si ottiene che $A0 = 0$.

In particolare il prodotto è un'operazione sull'insieme $\mathcal{M}(n \times n; \mathbb{K})$ che è quindi un anello; se $n \geq 2$ il prodotto non è commutativo, qualunque sia il campo $\mathbb{K}$.: siano infatti A la matrice data da $a_{11} = a_{12} = 1$ e $a_{ij} = 0$ in tutti gli altri posti e B la matrice data da $b_{11} = b_{21} = 1$ e $b_{ij} = 0$ in tutti gli altri posti. Allora $(ab)_{11} = 1 + 1$ mentre $(ba)_{11} = 1$ e quindi $AB \neq BA$.

Per ogni naturale positivo n la *identica* ${}^nI \in \mathcal{M}(n \times n; \mathbb{K})$ (o solo I se non c'è pericolo di confusione) è definita mediante

$$i_{jk} = \begin{cases} 1 & \text{se } j = k \\ 0 & \text{se } j \neq k. \end{cases}$$

Valgono allora le seguenti ulteriori proprietà:

MP6 $\forall A \in \mathcal{M}(m \times n; \mathbb{K}) \quad AI = A$.

MP7 $\forall A \in \mathcal{M}(m \times n; \mathbb{K}) \quad IA = A$.

Una matrice $A \in \mathcal{M}(n \times n; \mathbb{K})$ si dice *invertibile* se esiste una matrice $B \in \mathcal{M}(n \times n; \mathbb{K})$ tale che $AB = BA = I$. Tale B, che, se esiste, è unica viene chiamata l'*inversa di* A e denotata con il simbolo A^{-1}.

2.4 Generatori e basi

Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia A un sottoinsieme di V. Consideriamo la famiglia $\mathcal{S}(A) = \{W \lhd V \text{ tali che } A \subseteq W\}$, costituita da tutti i sottospazi vettoriali di V che contengono A; tale famiglia non è vuota perché $V \in \mathcal{S}(A)$.

Indichiamo con $\mathcal{L}(A)$ l'intersezione di tale famiglia. È facile controllare che si tratta ancora di un sottospazio di V, anzi è il più piccolo sottospazio di V che contiene A. Tanto per familiarizzarci notiamo che $\mathcal{L}(\emptyset) = \{0\}$.

Se $A \subseteq V$ allora, dato $v \in V$, avremo che $v \in \mathcal{L}(A) \Leftrightarrow \forall W \lhd V$ tale che $A \subseteq W, \quad v \in W$: tremendo, ma adesso impariamo una legge molto più semplice da verificare in pratica.

Definizione 13. *Siano V uno spazio vettoriale, n un intero positivo e $(v_1, \ldots, v_n)$ una n-upla ordinata di elementi di V. Si dice che un vettore $v \in V$ è* combinazione lineare *di $(v_1, \ldots, v_n)$ (o, per alleggerire la notazione, dei vettori $v_1, \ldots, v_n$) se esiste una n-upla ordinata di scalari $(\lambda_1, \ldots, \lambda_n)$ tali che*

$$v = \sum_{i=1}^{n} \lambda_i v_i.$$

L'insieme di tutte le combinazioni lineari di $(v_1, \ldots, v_n)$ viene indicato con $\mathcal{L}(v_1, \ldots, v_n)$.

Proposizione 1. *$\mathcal{L}(v_1, \ldots, v_n)$. è il più piccolo sottospazio vettoriale di V che contiene tutti i vettori $v_1, v_2, \ldots, v_n$, cioè*

$$\mathcal{L}(v_1, v_2, \ldots, v_n) = \mathcal{L}(\{v_1, v_2, \ldots, v_n\}).$$

Dimostrazione. 1. $\mathcal{L}(v_1, v_2, \ldots, v_n)$ è un sottospazio vettoriale di V infatti

S1 Ponendo $\lambda_i = 0 \quad \forall i = 1, \ldots, n$ si ottiene che $0 \in \mathcal{L}(v_1, v_2, \ldots, v_n)$.

S2 Se $v = \sum_{i=1}^{n} \lambda_i v_i$ e $w = \sum_{i=1}^{n} \mu_i v_i$ allora

$$v + w = \sum_{i=1}^{n} (\lambda_i + \mu_1) v_i \in \mathcal{L}(v_1, v_2, \ldots, v_n).$$

S3 Se $v = \sum_{i=1}^{n} \lambda_i v_i$ e $\lambda \in \mathbb{K}$ allora

$$\lambda v = \sum_{i=1}^{n} (\lambda \lambda_i) v_i \in \mathcal{L}(v_1, v_2, \ldots, v_n).$$

2. per ogni $i = 1, \ldots, n$ otteniamo che $v_i \in \mathcal{L}(v_1, v_2, \ldots, v_n)$ prendendo, per ogni $k = 1, \ldots, n$, $\lambda_k = 1$ se $k = i$ e $\lambda_k = 0$ se $k \neq i$.

3. Se $W \lhd V, \quad v_i \in W \quad \forall i = 1, \ldots, n$ e $v = \sum_{i=1}^{n} \lambda_i v_i$ allora $\lambda_i v_i \in W$ per S3 e $v \in W$ per S2.

□

Si dice che l'insieme A *genera* un sottospazio $W \lhd V$ (o che A è un *insieme di generatori* di W) se $\mathcal{L}(A) = W$.

Un qualunque insieme A si dice *finito* se esiste un numero naturale n ed una corrispondenza biunivoca tra $\overline{n}$ ed A. Tale numero n, che dipende solo

da A.[1], è detto il *numero degli elementi* A e indicato con $\sharp A$. L'insieme vuoto è finito e $\sharp\emptyset = 0$.

Si dice che uno spazio vettoriale V è *finitamente generato* se esiste un suo sottoinsieme finito che lo genera.

In queste note ci occuperemo esclusivamente di spazi vettoriali finitamente generati.

Ovviamente se $(v_1, \ldots, v_n)$ è un'n-upla ordinata di vettori di V, $m \leq n$ e $v_1, \ldots, v_m$ generano V allora anche $v_1, \ldots, v_n$ lo fanno.

A titolo di esempio, prendiamo in $\mathbb{K}^n$ l'n-upla ordinata di vettori $(e_1, \ldots, e_n)$ data da

$$\begin{array}{rcl} e_1 & = & (1, 0, \ldots, 0) \\ e_2 & = & (0, 1, \ldots, 0) \\ \vdots & \vdots & \vdots \quad \vdots \\ e_n & = & (0, 0, \ldots, 1) \end{array} \tag{2.1}$$

abbiamo che, per ogni $v = (x_1, \ldots, x_n) \in \mathbb{K}^n$, $v = \sum_{i=1}^n x_i e_i$ e quindi $e_1, \ldots, e_n$ generano $\mathbb{K}^n$.

Definizione 14. *1. Si dice un'n-upla ordinata di vettori $(v_1, \ldots, v_n)$ è* linearmente dipendente *(o, brevemente, che i vettori $v_1, \ldots, v_n$ sono* linearmente dipendenti*) se esiste un' $n-$upla ordinata di scalari non tutti nulli $(\lambda_1, \ldots, \lambda_n)$ tali che $\sum_{i=1}^n \lambda_i v_i = 0$.*

2. Si dice che i vettori $v_1, \ldots, v_n$ sono linearmente indipendenti *se non sono linearmente dipendenti.*

Detto in maniera più simbolica una n-upla ordinata di vettori $(v_1, \ldots, v_n)$ è linearmente dipendente se e solo se

$$\text{Esiste } \lambda : \overline{n} \longrightarrow \mathbb{K} \text{ tale che } \exists i \in \overline{n} \text{ tale che } \lambda_i \neq 0 \text{ e } \sum_{i=1}^n \lambda_i v_i = 0$$

il che, visto il significato di "implica", ci assicura che la funzione vuota $\overline{0} \longrightarrow V$ è linearmente indipendente.

Se esiste $i = 1, \ldots, n$ tale che $v_i = 0$ allora scegliendo

$$\lambda_j = \begin{cases} 1 & \text{se} \quad j = i \\ 0 & \text{se} \quad j \neq i \end{cases}$$

[1] stiamo solo dicendo che, contando in modi diversi un insieme finito si ottiene sempre lo stesso numero. Chi legge si sentirà frustrato nel non riuscire a dimostrare questa banalità senza la quale la vita sarebbe veramente difficile

si ottiene che $v_1, \ldots, v_n$ sono linearmente dipendenti. Analogamente se l'$n-$upla ordinata $(v_1, \ldots, v_n)$ non è iniettiva, cioè esistono $i \neq j$ tali che $v_i = v_j$ allora basta prendere

$$\lambda_k = \begin{cases} 1 & \text{se} \quad k = i \\ -1 & \text{se} \quad k = j \\ 0 & \text{se} \quad k \notin \{i, j\} \end{cases}$$

si vede di nuovo che si tratta di vettori linearmente dipendenti.

In sintesi $v_1, \ldots, v_n$ sono linearmente indipendenti se dall'equazione

$$\sum_{i=1}^{n} \lambda_i v_i = 0$$

riusciamo a dedurre che $\lambda_i = 0 \quad \forall i = 1, \ldots, n$.

Questa è di gran lunga la tecnica che si applica più frequentemente per riconoscere dei vettori sono linearmente indipendenti, ma saranno utili anche quelle esposte nella seguente

Proposizione 2. *Siano $v_1, \ldots, v_n$ elementi dello spazio vettoriale V.*

1. *$v_1, \ldots, v_n$ sono linearmente dipendenti se e solo se uno di essi è combinazione lineare degli altri.*

2. *Se $v_1, \ldots, v_{n-1}$ sono linearmente indipendenti allora $v_1, \ldots, v_n$ sono linearmente dipendenti se e solo se $v_n \in \mathcal{L}(v_1, \ldots, v_{n-1})$.*

Dimostrazione. 1. $\Rightarrow$ Siano $\lambda_1, \ldots, \lambda_n$ scalari non tutti nulli tali che

$$\sum_{i=1}^{n} \lambda_i v_i = 0,$$

esiste un indice i tale che $\lambda_i \neq 0$ e con facili passaggi algebrici si ottiene

$$v_i = \sum_{s=1}^{i-1} -\lambda_i^{-1} \lambda_s v_s + \sum_{s=i+1}^{n} -\lambda_i^{-1} \lambda_s v_s.$$

$\Leftarrow$ Se $v_i = \sum_{s=1}^{i-1} \mu_s v_s + \sum_{s=i+1}^{n} \mu_s v_s$ con semplici passaggi algebrici si ottiene

$$\sum_{s=1}^{i-1} \mu_s v_s + (-1) v_i + \sum_{s=i+1}^{n} \mu_s v_s = 0.$$

2. $\Rightarrow$ Siano $\lambda_1, \ldots, \lambda_n$ scalari non tutti nulli tali che

$$\sum_{i=1}^{n} \lambda_i v_i = 0,$$

siccome $v_1, \ldots, v_{n-1}$ sono linearmente indipendenti si ottiene che $\lambda_n \neq 0$, e poi si procede come nella dimostrazione appena esposta prendendo $i = n$.

$\Leftarrow$ Già fatta.

□

Definizione 15. *Si dice che $\mathcal{B} = (b_1, \ldots, b_n)$ è una* base *per V se*

- *$b_1, \ldots, b_n$ sono linearmente indipendenti.*
- *$b_1, \ldots, b_n$ generano V.*

Per esempio è facile vedere che la $n-$ upla ordinata di vettori definita in 2.1 costituiscono una base per $\mathbb{K}^n$, chiamata la *standard* e indicata con il simbolo $\mathcal{E}$ (o $\mathcal{E}^n$ se esiste pericolo di confusione).

Anche lo spazio vettoriale $\mathcal{M}(m \times n; \mathbb{K})$ ha una base standard: cominciamo con il definire, per ogni $(i, j) \in \overline{m} \times \overline{n}$, la matrice E^{ij} mediante

$$e^{ij}_{rs} = \begin{cases} 1 & \text{se } (r, s) = (i, j) \\ 0 & \text{se } (r, s) \neq (i, j) \end{cases} \tag{2.2}$$

È facile vedete che l'insieme di tutte queste matrici genera $\mathcal{M}(m \times n, \mathbb{K})$.

Definiamo poi una funzione $f : \overline{m} \times \overline{n} \longrightarrow \overline{mn}$ mediante $f(i, j) = (i - 1)n + j$; si controlla facilmente che si tratta di una corrispondenza biunivoca e chiamiamo g la sua inversa.

Infine definiamo $\mathcal{E} : \overline{mn} \longrightarrow \mathcal{M}(m \times n, \mathbb{K})$ ponendo $e_k = E_{g(k)}$. Si verifica senza patemi che la $mn-$upla ordinata $\mathcal{E}$ è linearmente indipendente e quindi una base per $\mathcal{M}(m \times n, \mathbb{K})$, si veda anche la propoisizione 4.

Se $n \in \mathbb{N}$, A è un insieme, $\mathcal{A} = (a_1, \ldots, a_n)$ e $\mathcal{B} = (b_1, \ldots, b_n)$ sono n-uple ordinate di elementi di A si dice che $\mathcal{B}$ è *ottenuta da $\mathcal{A}$ cambiando l'ordine* se esiste una biezione $\sigma : \overline{n} \longrightarrow \overline{n}$ tale che $\mathcal{B} = \mathcal{A} \circ \sigma$, cioè, per ogni $i = 1, \ldots, n \quad b_i = a_{\sigma(i)}$. Se $(v_1, \ldots, v_n)$ genera uno spazio vettoriale V (rispettivamente è linearmente indipendente) e $(w_1, \ldots, w_n)$ è ottenuta da $(v_1, \ldots, v_n)$ cambiando l'ordine allora anche $(w_1, \ldots, w_n)$ genera V (rispettivamente è linearmente indipendente). In parole povere l'essere generatori o linearmente indipendenti non dipende dall'ordine.

Il risultato più decisivo di tutta l'algebra lineare è il seguente

Teorema 3. *Sia V uno spazio vettoriale, $(v_1, \ldots, v_n)$ linearmente indipendente e $(w_1, \ldots, w_m)$ generatori di V. Allora $n \leq m$.*

Dimostrazione. Supponiamo per assurdo che sia $n > m$; siccome $v_1 \in V$ esistono degli scalari tali che

$$v_1 = \sum_{i=1}^{m} \lambda_i w_i$$

La proposizione 2 ci assicura che $v_1 \neq 0$ e quindi almeno uno degli scalari è non nullo; salvo cambiare l'ordine in $w_1, \ldots, w_m$, possiamo supporre sia $\lambda_1 \neq 0$. Con facili passaggi algebrici si ottiene

$$w_1 = \lambda_1^{-1} v_1 - \lambda_1^{-1}\lambda_2 w_2 - \ldots - \lambda_1^{-1}\lambda_m w_m$$

da cui si deduce immediatamente che anche $v_1, w_2, \ldots, w_m$ generano V. Ma allora esistono degli scalari tali che

$$v_2 = \mu_1 v_1 + \mu_2 w_2 + \ldots + \mu_m w_m.$$

Visto che $v_2 \notin \mathcal{L}(v_1)$ si ha che $\mu_2, \ldots, \mu_m$ non sono tutti nulli e, cambiando se necessario l'ordine in $w_2, \ldots, w_m$, possiamo supporre che $\mu_2 \neq 0$ e allora

$$w_2 = -\mu_2^{-1}\mu_1 v_1 + \mu_2^{-1} v_2 - \mu_2^{-1}\mu_3 w_3 - \ldots - \mu_2^{-1}\mu_m w_m$$

e quindi anche $v_1, v_2, w_3, \ldots, w_m$ generano V. Ripetendo m volte questo ragionamento si arriva a dire che $v_1, v_2, \ldots, v_m$ generano V e quindi $v_{m+1} \in \mathcal{L}(v_1, v_2, \ldots, v_m)$, e questo contraddice la proposizione 2. □

La lettrice che sia infastidita dalle volgarità contenute nella dimostrazione precedente (il particolare quel "ripetendo m volte questo ragionamento") avrà maggior soddisfazione dalla

Dimostrazione corretta. Sia $U = \{p \in \overline{m}$ tali che $\exists u : \overline{m} \longrightarrow V$ tale che $\forall i \in \overline{p}$, $u_i = v_i$, $\forall i > p$, $u_i \in \{w_1, \ldots, w_m\}$, u genera $V\}$. $0 \in U$, che quindi non è vuoto. Sia $r = \max(U)$; affermo che $r \geq n$ da cui in particolare segue la tesi.

Supponiamo infatti che $r < n$. Per definizione di U esiste $u : \overline{m} \longrightarrow V$ tale che $\forall i \in \overline{r}$, $u_i = v_i$, $\forall i > r$, $u_i \in \{w_1, \ldots, w_m\}$, u genera V; allora esiste $\lambda \in \mathbb{K}^n$ tale che $v_{r+1} = \sum_{i=1}^{m} \lambda_i u_i = \sum_{i=1}^{r} \lambda_i v_i + \sum_{i=r+1}^{m} \lambda_i u_i$; siccome $(v_1, \ldots, v_n)$ è linearmente indipendente, esiste $j \in \{r+1, \ldots, m\}$ tale che $\lambda_i \neq 0$. Ne segue che $(v_1, \ldots, v_{r+1}, u_{r+1}, \ldots, u_{j-1}, u_{j+1}, \ldots, u_m)$ genera V e quindi $r + 1 \in U$, contraddizione. □

Questo teorema ha un numero impressionante di conseguenze:

Proposizione 3. *Sia V uno spazio vettoriale*

1. *Se $\mathcal{B} = (b_1, \ldots, b_n)$ e $\mathcal{C} = (c_1, \ldots, c_m)$ sono basi per V allora $m = n$. Questo numero è detto* dimensione *di V e indicato con* $\dim V$.

2. *Se $\dim V = n$ e $W \lhd V$ allora W è finitamente generato,* $\dim W \leq n$ *e se $\dim W = n$ allora $W = V$.*

3. *Se $\dim V = n$ e $v_1, \ldots, v_p$ sono linearmente indipendenti allora $p \leq n$; se $p = n$ allora $(v_1, \ldots, v_p)$ è una base di V, se invece $p < n$ allora esistono $v_{p+1}, \ldots, v_n$ tali che $(v_1, \ldots, v_p, \ldots, v_n)$ è una base di V. Cioè ogni $p-$upla ordinata di vettori linearmente indipendenti può essere* completata *a base.*

4. *Se $\dim V = n$ e $v_1, \ldots, v_m$ generano V allora $m \geq n$; se $m = n$ allora $(v_1, \ldots, v_m)$ è una base di V, se invece $m > n$ possiamo scartare $m - n$ vettori in modo che i rimanenti formino una base, cioè ogni $n-$upla ordinata di generatori può essere* sfoltita *a base.*

Dimostrazione. 1. Usando il fatto che $b_1, \ldots, b_n$ sono linearmente indipendenti e che $c_1, \ldots, c_n$ generano V si ha che $n \leq m$, e scambiando le basi tra di loro si trova che $m \leq n$.

2. La prima affermazione è la più delicata da dimostrare, anche se questo sembra paradossale: in fondo stiamo solo dicendo che se V è "piccolo"anche W lo è: in realtà tutto quello che sappiamo è che V ha un insieme di generatori con n elementi e vorremmo concludere che anche W ha un insieme finito di generatori; possiamo procedere così: sia $M = \{m \in \mathbb{N}$ tali che esiste un'm-upla ordinata linearmente indipendente di vettori di $W\}$. Questo insieme non è vuoto perché $0 \in M$, ed il teorema 3 ci assicura che, per ogni $m \in M$, $m \leq n$, quindi M ha un massimo, chiamiamolo m. Naturalmente questo significa che possiamo trovare m vettori linearmente indipendenti in W ma non possiamo trovarne $m + 1$. Sia $(v_1, \ldots, v_m)$ un' m-upla con tale proprietà; se $w \in W$ la $(m + 1)$-upla $(v_1, \ldots, v_m, w)$ è linearmente dipendente e quindi, per il punto 2 della proposizione 2, $w \in \mathcal{L}(v_1, \ldots, v_m)$: dunque $v_1, \ldots, v_m$ generano W e siamo a posto.

Una base di W è costituita da vettori linearmente indipendenti; inoltre se $\dim W = n$ e, per assurdo, esiste $v \in V$ tale che $v \notin W$ allora, aggiungendo v a una base di W di otterrebbero $n+1$ vettori linearmente indipendenti in V, per il punto 2 della proposizione 2, e questo contraddice il teorema.

3. La prima affermazione segue ovviamente dal teorema. Per quanto riguarda la seconda basta prendere $W = \mathcal{L}(v_1, \ldots, v_p)$ e applicare il punto precedente. Se invece $p < n$ allora $\mathcal{L}(v_1, \ldots, v_p) \neq V$ e quindi esiste $v_{p+1} \notin \mathcal{L}(v_1, \ldots, v_p)$. Per la proposizione 2 abbiamo che $v_1, \ldots, v_p, v_{p+1}$ sono ancora linearmente indipendenti. Ripetendo $n - p$ volte questo ragionamento si ha la tesi.

4. Anche qui basta dimostrare la terza affermazione. Visto che $m > n$ i vettori $v_1, \ldots, v_m$ sono linearmente dipendenti. Per la proposizione 2 uno almeno di essi è combinazione lineare degli altri. Escludendolo si continua ad avere un insieme di generatori, e ripetendo $m - n$ volte si ha la tesi.

□

Se $\mathcal{B} = (b_1, \ldots, b_n)$ è una base per uno spazio vettoriale V allora esiste (perché $b_1, \ldots, b_n$ generano V) ed è unica (perché $b_1, \ldots, b_n$ sono linearmente indipendenti) una $n-$upla ordinata di scalari $(x_1, \ldots, x_n)$ tale che

$$v = \sum_{i=1}^{n} x_i b_i.$$

Tali scalari sono chiamati le *coordinate* di v rispetto a $\mathcal{B}$ (se V ha una base standard e $\mathcal{B}$ è tale base, si parla di *coordinate* tout court).

Lo spazio vettoriale nullo ha per base $\emptyset$ e dunque dimensione zero, $\mathbb{K}^n$ ha dimensione n e $\dim \mathcal{M}(m \times n, \mathbb{K}) = mn$.

Abbiamo giurato che in queste note avremmo parlato soltanto di spazi vettoriali finitamente generati, ma vogliamo soddisfare la legittima curiosità della lettrice che a questo punto si sia domandato se ne esistano di *non* finitamente generati.

La risposta, pienamente esauriente, è fornita dal superesempio. Questa discussione è leggermente più impegnativa e non è il caso di farsi amareggiare da una comprensione non immediata;

Proposizione 4. *Siano $\mathbb{K}$ un campo e X un insieme:*

- *Se X è finito allora $\mathbb{K}^X$ è finitamente generato e* $\dim \mathbb{K}^X = \sharp X$.
- *Se X non è finito $\mathbb{K}^X$ non è finitamente generato.*

Dimostrazione. Per ogni $x \in X$ indichiamo con x^* l'elemento di $\mathbb{K}^X$ definito da

$$x^*(y) = \begin{cases} 1 & \text{se} \quad y = x. \\ 0 & \text{se} \quad y \neq x \end{cases}$$

Supponiamo che X sia finito ed abbia n elementi, cioè esiste una corrispondenza biunivoca $\mathcal{X} : \overline{n} \longrightarrow X$. Affermo che l' n-upla ordinata $(x_1^*, \ldots, x_n^*)$ è una base di $\mathbb{K}^X$.

Supponiamo infatti di avere un' n-upla ordinata $(\lambda_1, \ldots, \lambda_n)$ di scalari tali che $\sum_{i=1}^n \lambda_i x_i^* = 0$; questo 0 è naturalmente in vettore nullo di $\mathbb{K}^X$, cioè la funzione che associa lo scalare 0 a tutti gli elementi di X. Ma allora, per ogni $j = 1, \ldots, n$ abbiamo

$$0 = 0(x_j) = (\sum_{i=1}^n \lambda_i x_i^*)(x_j) = \sum_{i=1}^n \lambda_i (x_i^*(x_j)) = \lambda_j.$$

E quindi $x_1^*, \ldots, x_n^*$ sono linearmente indipendenti; il penultimo passaggio segue dalla definizione delle operazioni in $\mathbb{K}^X$ e l'ultimo da quella di x_i^*.

Sia poi $f \in \mathbb{K}^X$ e, per $i = 1, \ldots, n$, poniamo $\lambda_i = f(x_i) \in \mathbb{K}$; vogliamo dimostrare che $f = \sum_{i=1}^n \lambda_i x_i^*$, cioè che queste due funzioni coincidono su ogni elemento di X; visto che $X = \{x_1, \ldots, x_n\}$ abbiamo che, per ogni j

$$(\sum_{i=1}^n \lambda_i x_i^*)(x_j) = \sum_{i=1}^n \lambda_i (x_i^*(x_j)) = \lambda_j = f(x_j),$$

e si vince.

Supponiamo ora che X non sia finito, e dimostriamo che, per ogni $n \in \mathbb{N}$, possiamo trovare n vettori linearmente indipendenti in $\mathbb{K}^X$. Questo, usando il teorema 3, ci permette di concludere che $\mathbb{K}^X$ non è finitamente generato.

Per ogni $n \in \mathbb{N}$ possiamo trovare una funzione iniettiva $(x_1, \ldots, x_n) : \overline{n} \longrightarrow X$ e, come nella prima parte della dimostrazione, è facile vedere che $x_1^*, \ldots, x_n^*$ sono linearmente indipendenti. □

La prima parte di questa proposizione, scegliendo opportunamente X, ci permette di ricalcolare le dimensioni degli spazi vettoriali che già conosciamo.

Un altro strumento utilissimo per calcolare le dimensioni è dato dalla seguente

Proposizione 5 (Formula di Grassmann). *Sia V uno spazio vettoriale, $U \lhd V$ e $W \lhd V$ allora vale la formula*

$$\dim(U+W) = \dim U + \dim W - \dim U \cap W.$$

Dimostrazione. Poniamo $\dim U = p, \quad \dim W = q, \quad \dim U \cap W = s$. Alla luce della proposizione appena dimostrata avremo che $s \leq p$, $s \leq q$ e dobbiamo dimostrare che $\dim(U+W) = p+q-s$. Sia $\mathcal{B} = (b_1, \ldots, b_s)$ una base per $U \cap W$ (ovviamente prendiamo $\mathcal{B} = \emptyset$ se $s = 0$) e completiamola a base per U mediante l'aggiunta di altri $p-s$ vettori $c_{s+1}, \ldots, c_p$ (nessun vettore se $p = s$). Completiamo $\mathcal{B}$ a base di W mediante altri $q-s$ vettori $d_{s+1}, \ldots, d_q$ (nessun vettore se $q = s$). La proposizione sarà dimostrata se riusciremo a verificare che i $p+q-s$ vettori

$$b_1, \ldots, b_s, c_{s+1}, \ldots, c_p, d_{s+1}, \ldots, d_q$$

sono una base per $U+W$.

- Sono linearmente indipendenti: siano infatti

  $$\beta_1, \ldots, \beta_s, \gamma_{s+1}, \ldots, \gamma_p, \delta_{s+1}, \ldots, \delta_q$$

  degli scalari tali che

  $$\beta_1 b_1 + \ldots + \beta_s b_s + \gamma_{s+1} c_{s+1} + \ldots + \gamma_p c_p + \delta_{s+1} d_{s+1} + \ldots + \delta_q d_q = 0$$

  e poniamo $v = \beta_1 b_1 + \ldots + \beta_s b_s + \gamma_{s+1} c_{s+1} + \ldots + \gamma_p c_p \in U$, cosicché $-v = \delta_{s+1} d_{s+1} + \ldots + \delta_q d_q \in W$, ma, siccome W è un sottospazio vettoriale di V anche $v \in W$.

  Quindi $v \in U \cap W$ ed allora esistono s scalari $\lambda_1, \ldots, \lambda_s$ tali che

  $$v = \lambda_1 b_1 + \ldots + \lambda_s b_s$$

  da cui segue che

  $$\lambda_1 b_1 + \ldots + \lambda_s b_s + \delta_{s+1} d_{s+1} + \ldots + \delta_q d_q = v - v = 0$$

  e, essendo $b_1, \ldots, b_s, d_{s+1}, \ldots, d_q$ linearmente indipendenti, da qui si può dedurre che $\lambda_1, \ldots, \lambda_s, \delta_{s+1}, \ldots, \delta_q$ sono tutti nulli, e in particolare che $v = 0$; dalla definizione di v e dal fatto che $b_1, \ldots, b_s, c_{s+1}, \ldots, c_p$ sono linearmente indipendenti si conclude che $\beta_1, \ldots, \beta_s, \gamma_{s+1}, \ldots, \gamma_p$ sono tutti nulli.

- Generano $U+W$; sia infatti $v \in U+W$, allora esistono $u \in U$ e $w \in W$ tali che $v = u + w$. Ma esistono degli scalari tali che

$$u = \beta_1 b_1 + \ldots + \beta_s b_s + \gamma_{s+1} c_{s+1} + \ldots + \gamma_p c_p$$

$$w = \alpha_1 b_1 + \ldots + \alpha_s b_s + \delta_{s+1} d_{s+1} + \ldots + \delta_q d_q$$

ma allora

$$v = (\beta_1+\alpha_1)b_1+\ldots+(\beta_s+\alpha_s)b_s+\gamma_{s+1}c_{s+1}+\ldots+\gamma_p c_p+\delta_{s+1}c_{s+1}+\ldots+\delta_q d_q$$

□

La somma di sottospazi $U \lhd V$ e $W \lhd V$ si dice *diretta* e si scrive $U \oplus W$ se ogni vettore in $U+W$ si scrive in modo unico come somma di un elemento di U e uno di W. Si può riconoscere facilmente se una somma è diretta applicando la seguente

Proposizione 6. *Siano V uno spazio vettoriale, $U \lhd V$ e $W \lhd V$. Le seguenti condizioni sono equivalenti:*

1. *La somma di U e W è diretta.*
2. *Se $u \in U$ e $w \in W$ sono tali che $u + w = 0$ allora $u = 0$ e $w = 0$.*
3. *$U \cap W = \{0\}$.*

Dimostrazione. $2 \Rightarrow 1$ Siano $u, u' \in U$ e $w, w' \in W$ tali che $u + w = u' + w'$. allora abbiamo che $(u - u') + (w - w') = 0$, e, usando 2 si ha subito che $u = u'$, $w = w'$.

$1 \Rightarrow 3$ Sia $v \in U \cap W$ allora $v \in U$ e $v \in W$ e allora $v + 0 = 0 + v \in U + W$. Da 1 si deduce subito che $v = 0$.

$3 \Rightarrow 2$ Se $u \in U$ e $w \in W$ sono tali che $u + w = 0$ allora $u = -w$ da cui si deduce che $u \in U \cap W$. Usando 3 si vede che $u = 0$ e subito dopo che $w = 0$.

□

Se la somma di U e W è diretta la formula di Grassmann ci assicura che $\dim U + W = \dim U + \dim W$ e la sua dimostrazione ci dice che, se $(v_1, \ldots, v_p)$ e $(w_1, \ldots, w_q)$ sono basi per Ue W rispettivamente, allora $(v_1, \ldots, v_p, w_1, \ldots, w_q)$ è base per $U + W$.

2.5 Sistemi lineari

Definizione 16. *Siano n, m interi positivi. Un* sistema lineare *di m equazioni in n incognite è una coppia ordinata (A, b) dove $A \in \mathcal{M}(m \times n; \mathbb{K})$ e $b \in \mathcal{M}(m \times 1; \mathbb{K})$.*

Il sistema lineare (A, b) viene tradizionalmente indicato con la notazione $AX = b$, e così faremo anche in queste note. A viene chiamata la *matrice dei coefficienti* del sistema $AX = b$ e la matrice $(A|b) \in \mathcal{M}(m \times (n+1); \mathbb{K})$ le cui prime n colonne sono quelle di A e l'ultima è b viene detta la *matrice completa del sistema.*

L'*insieme delle soluzioni* di tale sistema lineare è $\mathrm{Sol}(A, b) = \{X \in \mathcal{M}(n \times 1; \mathbb{K})$ tali che $AX = b\}$. Si dice che il sistema è *risolubile* se $\mathrm{Sol}(A, b) \neq \emptyset$. Se $b = 0$ il sistema si dice *omogeneo.*

Proposizione 7. *Siano $A \in \mathcal{M}(m \times n; \mathbb{K})$.*

1. $\mathrm{Sol}(A, 0)$ *è un sottospazio vettoriale di $\mathcal{M}(n \times 1; \mathbb{K})$.*
2. $\{v \in \mathbb{K}^n$ *tali che* $Av^t = 0\}$ *è un sottospazio vettoriale di $\mathbb{K}^n$.*

Dimostrazione. 1. Si deve verificare

- S1 $A0 = 0$ e questa è conseguenza della proprietà MP3 del prodotto di matrici citata sopra.
- S2 Se $AX = 0$, $AY = 0$ allora $A(X + Y) = AX + AY = 0 + 0 = 0$.
- S3 Se $AX = 0$ e $\lambda \in \mathbb{K}$ allora $A(\lambda X) = \lambda(AX) = \lambda 0 = 0$.

2. È ovvio.

□

Proposizione 8. *Sia $AX = b$ un sistema lineare risolubile e sia $X_0 \in \mathrm{Sol}(A, b)$. Allora*

$$\mathrm{Sol}(A, b) = X_0 + \mathrm{Sol}(A, 0) =$$

$$\{X \in \mathcal{M}(n \times 1; \mathbb{K}) \text{ tali che } \exists Y \in \mathrm{Sol}(A, 0) \text{ tale che } X = X_0 + Y\}$$

Dimostrazione. Se $X \in \mathrm{Sol}(A, b)$, ponendo $Y = X - X_0$ si ha che $X = X_0 + Y$ e che $AY = A(X - X_0) = AX - AX_0 = b - b = 0$. Viceversa, se $AY = 0$ allora $A(X_0 + Y) = AX_0 + AY = b + 0 = b$. □

Sia $A \in \mathcal{M}(m \times n; \mathbb{K})$. Abbiamo già detto che cosa sono le sue righe $A_1, \ldots, A_m \in \mathcal{M}(1 \times n; \mathbb{K})$, e definiamo il *rango per righe* di A come il numero

$$\mathrm{rg}_R(A) = \dim \mathcal{L}(A_1, \ldots, A_m).$$

Naturalmente $\mathrm{rg}_R(A) \leq m$, ma anche $\mathrm{rg}_R(A) \leq n$, in quanto dimensione di un sottospazio di $\mathcal{M}(1 \times n; \mathbb{K})$.

Analogamente possiamo definire il *rango per colonne* di A come

$$\mathrm{rg}_C(A) = \dim \mathcal{L}(A^1, \ldots, A^n).$$

Di nuovo avremo che $\mathrm{rg}_C(A) \leq \min(m, n)$ in quanto è la dimensione di un sottospazio di $\mathcal{M}(m \times 1; \mathbb{K})$.

Dimostreremo più avanti, nel corollario 3, il seguente importante risultato

Teorema 4. $\forall n, m, A \in \mathcal{M}(m \times n; \mathbb{K}) \quad \mathrm{rg}_R(A) = \mathrm{rg}_C(A)$.

Questo numero verrà chiamato semplicemente *rango* di A e indicato con $\mathrm{rg}(A)$. Siamo ora in grado di enunciare e provare il teorema che ci dice quando un sistema lineare è risolubile:

Teorema 5. (Rouché-Capelli) *Il sistema lineare $AX = b$ è risolubile se e solo se* $\mathrm{rg}(A) = \mathrm{rg}(A|b)$.

Dimostrazione. Conviene pensare al rango per colonne. Evidentemente

$$\mathcal{L}(A^1, \ldots, A^n) \lhd \mathcal{L}(A^1, \ldots, A^n, b)$$

e l'eguaglianza vale se e solo se $b \in \mathcal{L}(A^1, \ldots, A^n)$. Perciò abbiamo provato che

$$\mathrm{rg}(A) = \mathrm{rg}(A|b) \Leftrightarrow b \in \mathcal{L}(A^1, \ldots, A^n).$$

Ma quest'ultima condizione vuol dire che esistono degli scalari $x_1, \ldots, x_n$ tali che $\sum_{i=1}^n x_i A^i = b$, e questo significa che il vettore $(x_1, \ldots, x_n)^t \in \mathrm{Sol}(A, b)$. □

Perché questo risultato diventi praticamente efficace abbiamo bisogno di un metodo che ci consenta di calcolare il rango di una matrice. Quello che proponiamo calcola il rango per righe, ma per illustrarlo dovremo introdurre qualche concetto nuovo. L'idea è di modificare una matrice, ma senza cambiarne il rango, finché si trovi in una forma in cui il rango sia evidente. Le modifiche avverranno attraverso le cosiddette *operazioni elementari sulle righe* che sono

1. Scambiare tra loro due righe: se $A \in \mathcal{M}(m \times n; \mathbb{K})$ e $1 \leq i < j \leq m$ la matrice modificata ha per righe $A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n$ dove A_j si trova all'$i-$esimo posto e A_i al $j-$esimo. Evidentemente questa operazione non modifica il rango per righe della matrice.

2. Moltiplicare una riga per uno scalare non nullo: se $A \in \mathcal{M}(m \times n; \mathbb{K})$ e $1 \leq i \leq m$ e $0 \neq \lambda \in \mathbb{K}$ la matrice modificata ha per righe $A_1, \ldots, \lambda A_i, \ldots, A_n$ dove λA_i si trova all'$i-$esimo posto. Anche in questo caso è facile vedere che questa operazione non modifica il rango per righe.

3. Sostituire una riga con la somma tra lei e un multiplo (anche nullo) di una riga che la precede: se $A \in \mathcal{M}(m \times n; \mathbb{K})$ e $1 \leq i < j \leq m$ e $\lambda \in \mathbb{K}$, la matrice modificata ha per righe $A_1, \ldots, A_i, \ldots, A_j + \lambda A_i, \ldots, A_n$ dove A_i si trova all'$i-$esimo posto e $A_j + \lambda A_i$ al $j-$ esimo. Anche qui il rango non cambia.

Diremo che un vettore non nullo $X = (x_1, \ldots, x_n) \in \mathbb{K}^n$ ha s *zeri iniziali* se $x_i = 0$ quando $i \leq s$ ma $x_{s+1} \neq 0$. Questo numero s viene indicato con il simbolo $\mathrm{zi}(X)$ e ovviamente si ha $0 \leq \mathrm{zi}(X) < n$.

Definizione 17. *Si dice che una matrice $A \in \mathcal{M}(m \times n; \mathbb{K})$ è* in forma ridotta *se si verificano le seguenti circostanze:*

- *Per ogni $i = 1, \ldots, m-1$, se $A_i = 0$ anche $A_{i+1} = 0$, cioè le (eventuali) righe nulle sono confinate agli ultimi posti disponibili.*

- *Per ogni $i = 1, \ldots, m-1$, se $A_{i+1} \neq 0$ allora $\mathrm{zi}(A_i) < \mathrm{zi}(A_{i+1})$; cioè ogni riga non nulla ha più zeri iniziali di quella che la precede.*

Proposizione 9. *Attraverso operazioni elementari sulle righe, e quindi senza modificarne il rango, possiamo trasformare una matrice qualunque in una in forma ridotta.*

La dimostrazione, che non è altro che un rifacimento teorico del processo che si attua negli esempi, è pesante dal punto di vista notazionale e queste note non l'ospiteranno.

Il problema del calcolo del rango, per righe, di una matrice sarà dunque risolto se possiamo calcolare il rango delle matrici ridotte.

Proposizione 10. *Il rango per righe di una matrice ridotta è uguale al numero delle sue righe non nulle.*

Dimostrazione. Sia $A \in \mathcal{M}(m \times n; \mathbb{K})$ una matrice in forma ridotta, e sia A_t la sua ultima riga non nulla. Evidentemente

$$\mathcal{L}(A_1, \ldots, A_m) = \mathcal{L}(A_1, \ldots, A_t)$$

e ci basta dimostrare che $A_1, \ldots, A_t$ sono linearmente indipendenti. Siano $\lambda_1, \ldots, \lambda_t$ degli scalari tali che

$$\lambda_1 A_1 + \ldots + \lambda_t A_t = 0.$$

Abbiamo subito che $\lambda_1 a_{1zi(A_1)} = 0$, da cui si deduce che $\lambda_1 = 0$ visto che $a_{1zi(A_1)} \neq 0$. Usando questo risultato si vede analogamente che $\lambda_2 = 0$ e, uno dopo l'altro, che tutti i λ_i sono nulli. □

Le operazioni elementari sulle righe ci permettono dunque di scoprire quando un sistema lineare è risolubile, ma c'è di meglio:

Proposizione 11. *Sia $AX = b$ un sistema lineare. Se la matrice completa $(A'|b')$ è ottenuta da $(A|b)$ mediante operazioni elementari sulle righe allora* $\mathrm{Sol}(A, b) = \mathrm{Sol}(A', b')$.

Dimostrazione. Si tratta evidentemente di controllare che ogni operazione elementare sulle righe non modifica l'insieme delle soluzioni di un sistema, e questo è lasciato per esercizio. □

Capitolo 3

Applicazioni lineari

3.1 Applicazioni lineari

Definizione 18. *Siano V, W spazi vettoriali e $V \xrightarrow{T} W$ una funzione. Si dice che T è un'* applicazione lineare *se soddisfa le seguenti regole:*

L1 $\forall v, w \in V \quad T(v+w) = T(v) + T(w)$.

L2 $\forall v \in V, \forall \lambda \in \mathbb{K} \quad T(\lambda v) = \lambda T(v)$.

Ad esempio la funzione nulla $V \xrightarrow{0} W$ definita da $0(v) = 0 \, \forall v \in V$ è un'applicazione lineare. Anche id_V è lineare. L'insieme di tutte le applicazioni lineari $V \xrightarrow{T} W$ viene indicato con il simbolo $\mathrm{Hom}(V, W)$. Una funzione biettiva di $\mathrm{Hom}(V, W)$ viene detta *isomorfismo*; è facile vedere che se T è un isomorfismo, anche la sua inversa T^{-1} lo è. Si dice che gli spazi vettoriali V e W sono *isomorfi* se esiste un isomorfismo in $\mathrm{Hom}(V, W)$. Un importante esempio è il seguente: Sia $\mathcal{B}$ una base per lo spazio vettoriale V di dimensione n. La funzione

$$\mathbb{K}^n \xrightarrow{C_{\mathcal{B}}} V$$

definita da $C_{\mathcal{B}}(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i b_i$ è un isomorfismo.

Ogni elemento di $\mathrm{Hom}(V, W)$ gode automaticamente delle seguenti proprietà:

- $T(0) = 0$, infatti $T(0) = T(00) = 0T(0) = 0$.
- $\forall v \in V \quad T(-v) = -T(v)$, infatti $T(v) + T(-v) = T(v - v) = T(0) = 0$.

Definizione 19. *Sia* $T \in \mathrm{Hom}(V,W)$ *il* nucleo *di* T *è l'insieme*

$$\mathrm{Ker}T = \{v \in V \text{ tale che } T(v) = 0\}$$

e l' immagine *di* T *è l'insieme*

$$\mathrm{Im}T = \{w \in W \text{ tali che } \exists v \in V \text{ tale che } T(v) = w\}.$$

È facile vedere che $\mathrm{Ker}T \lhd V$ e che $\mathrm{Im}T \lhd W$. La dimensione di $\mathrm{Ker}T$ viene detta la *nullità* di T e quella di $\mathrm{Im}T$ si chiama *rango* di T e si indica con $\mathrm{rg}(T)$. Ovviamente T è suriettiva se e solo se $\mathrm{Im}T = W$, se e solo se $\mathrm{rg}(T) = \dim W$.

Quindi il rango ci fornisce un'informazione sull'applicazione lineare; altrettanto fa il nucleo:

Proposizione 12. *Sia* $T \in \mathrm{Hom}(V,W)$. T *è iniettiva se e solo se* $\mathrm{Ker}T = \{0\}$.

Dimostrazione. Supponiamo che T sia iniettiva e che $v \in \mathrm{Ker}T$, allora $T(v) = 0 = T(0)$ e quindi $v = 0$. Viceversa supponiamo che $\mathrm{Ker}T = \{0\}$ e che $v, w \in V$ soddisfino $T(v) = T(w)$; ma allora $T(v - w) = T(v) - T(w) = 0$ e quindi $v - w = 0$. □

Proposizione 13. *Sia* $T \in \mathrm{Hom}(V,W)$ *e sia* U *un sottospazio di* V. *Allora*

1. $T(U) = \{T(u) \text{ tali che } u \in U\}$ *è un sottospazio di* W.
2. *Se* $u_1, \ldots, u_p$ *generano* U *allora* $T(u_1), \ldots, T(u_p)$ *generano* $T(U)$.
3. *Se* $u_1, \ldots, u_p \in U$ *sono linearmente indipendenti e* T *è iniettiva allora* $T(u_1), \ldots, T(u_p)$ *sono linearmente indipendenti.*
4. *Se* T *è iniettiva allora* $\dim U = \dim T(U)$.

Dimostrazione. 1. $0 = T(0) \in T(U)$ perché $0 \in U$. Se $w, w' \in T(U)$ allora esistono $u, u' \in U$ tali che $T(u) = w$, $T(u') = w'$; ma allora $w + w' = T(u) + T(u') = T(u + u') \in T(U)$ perché $u + u' \in U$. Se $w \in T(U)$ e $\lambda \in \mathbb{K}$ esiste $u \in U$ tale che $T(u) = w$; ma allora $\lambda w = \lambda T(u) = T(\lambda u) \in T(U)$ perché $\lambda u \in U$.

2. Se $w \in T(U)$ esiste $u \in U$ tale che $T(u) = w$, ma esiste una p-upla ordinata di scalari $(\lambda_1, \ldots, \lambda_p)$ tale che $u = \sum_{i=1}^{p} \lambda_i u_i$ e allora $w = T(u) = T(\sum_{i=1}^{p} \lambda_i u_i) = \sum_{i=1}^{p} \lambda_i T(u_i)$.

3. Se $0 = \sum_{i=1}^{p} \lambda_i T(u_i) = T(\sum_{i=1}^{p} \lambda_i u_i)$ allora $\sum_{i=1}^{p} \lambda_i u_i \in \mathrm{Ker}T = \{0\}$; siccome $u_1, \ldots, u_p$ sono linearmente indipendenti segue che $\lambda_i = 0 \ \forall i = 1, \ldots p$.

4. Se $(u_1, \ldots, u_p)$ è una base di U allora $(T(u_1), \ldots, T(u_p))$ genera $T(U)$ per 2 ed è linearmente indipendente per 3.

□

Uno strumento indispensabile dell'algebra lineare è dato dal seguente

Teorema 6 (nullità + rango). *Se V è uno spazio vettoriale finitamente generato, W è uno spazio vettoriale e $T \in \mathrm{Hom}(V, W)$ allora*

$$\dim \mathrm{Ker}T + \dim \mathrm{Im}T = \dim V.$$

Dimostrazione. Cominciamo con l'osservare che, grazie al punto 2 della proposizione precedente, ImT è finitamente generato; anche Ker$T$ lo è in quanto sottospazio di V. Quindi le dimensioni menzionate nell'enunciato del teorema hanno senso.

Sia $\mathcal{B} = (b_1, \ldots, b_p)$ una base di Ker$T$ ($\mathcal{B} = \emptyset$ se Ker$T = \{0\}$) e completiamo $\mathcal{B}$ a base di V mediante i vettori $b_{p+1}, \ldots, b_n$ (nessun vettore se Ker$T = V$, cioè se T è l'applicazione nulla). Il teorema sarà dimostrato se riusciamo a provare che $(T(b_{p+1}), \ldots, T(b_n))$ è una base per ImT.

- $T(b_{p+1}), \ldots, T(b_n)$ sono linearmente indipendenti, infatti se $\lambda_{p+1}, \ldots, \lambda_n$ sono scalari tali che

  $$\lambda_{p+1} T(b_{p+1}) + \ldots + \lambda_n T(b_n) = 0$$

  allora, grazie al fatto che T è lineare, possiamo scrivere

  $$T(\lambda_{p+1} b_{p+1} + \ldots + \lambda_n b_n) = 0$$

  e cioè abbiamo che $\lambda_{p+1} b_{p+1} + \ldots + \lambda_n b_n \in \mathrm{Ker}T$, esistono allora degli scalari $\lambda_1, \ldots, \lambda_p$ tali che

  $$\lambda_1 b_1 + \ldots + \lambda_p b_p - \lambda_{p+1} b_{p+1} - \ldots - \lambda_n b_n = 0.$$

  Siccome $b_1, \ldots, b_n$ sono linearmente indipendenti si deduce che $\lambda_{p+1}, \ldots, \lambda_n$ sono tutti nulli.

- $T(b_{p+1}), \ldots, T(b_n)$ generano $\mathrm{Im}T$. Sia infatti $w \in \mathrm{Im}T$, allora esiste $v \in V$ tale che $T(v) = w$. Dato che $b_1, \ldots, b_n$ generano V esistono degli scalari $\lambda_1, \ldots, \lambda_n$ tali che

$$v = \lambda_1 b_1 + \ldots + \lambda_n b_n,$$

da cui deduciamo subito che

$$w = T(v) = T(\lambda_1 b_1 + \ldots + \lambda_n b_n) = \lambda_1 T(b_1) + \ldots + \lambda_n T(b_n)$$

$$= \lambda_{p+1} T(b_{p+1}) + \ldots + \lambda_n T(b_n),$$

visto che $T(b_i) = 0$ per ogni $i = 1, \ldots, p$.

Questo conclude la dimostrazione. □

Anche il teorema nullità più rango ha una serie di illuminanti conseguenze: poniamo $\dim V = n$ e $\dim W = m$.

1. Se $T \in \mathrm{Hom}(V, W)$ è iniettiva allora $\dim \mathrm{Ker}T = 0$ e quindi il teorema ci dice che $\dim \mathrm{Im}T = n$; dato che $\mathrm{Im}T \lhd W$ si evince che $n \leq m$; in altri termini, se $n > m$ non esistono funzioni iniettive (e quindi nemmeno biettive) in $\mathrm{Hom}(V, W)$.

2. Se $T \in \mathrm{Hom}(V, W)$ è suriettiva allora $\dim \mathrm{Im}T = m$ e quindi il teorema ci dice che $n \geq m$; in altri termini, se $n < m$ non esistono funzioni suriettive (e quindi nemmeno biettive) in $\mathrm{Hom}(V, W)$.

3. D'altra parte se $(b_1, \ldots, b_n), \quad (c_1, \ldots, c_m)$ sono basi per V e W rispettivamente, definendo

$$T(\sum_{i=1}^{n} x_i b_i) = \begin{cases} \sum_{i=1}^{n} x_i c_i & \text{se } n \leq m, \\ \sum_{i=1}^{m} x_i c_i & \text{se } n \geq m, \end{cases}$$

si trova, nel primo caso, un'applicazione lineare iniettiva, e nel secondo una suriettiva. Se poi $n = m$ si ha un isomorfismo.

Tra le righe abbiamo anche provato che due spazi vettoriali sono isomorfi se e solo se hanno la stessa dimensione.

4. Se $\dim V = \dim W$ e $L \in \mathrm{Hom}(V, W)$ è iniettiva (rispettivamente suriettiva) allora è anche suriettiva (iniettiva) e quindi è un isomorfismo.

Proposizione 14. *Siano V, W spazi vettoriali, $\mathcal{B} = (b_1, \ldots, b_n)$ una base di V e $\tau : \{b_1, \ldots, b_n\} \longrightarrow W$ una funzione. Esiste un'unica*

$$T \in \mathrm{Hom}(V, W) \text{ tale che } T(b_j) = \tau(b_j)\ \forall j = 1, \ldots, n.$$

Tale T è detta l' estensione lineare *di τ.*

Dimostrazione. Cominciamo col provare l'unicità: siano dunque $T, T' \in \mathrm{Hom}(V, W)$ funzioni che soddisfano quanto richiesto. Se $v \in V$ esiste un'unica $n-$upla ordinata di scalari $(x_1, \ldots, x_n)$ tale che $v = \sum_{j=1}^{n} x_j b_j$ ma allora

$$T(v) = T(\sum_{j=1}^{n} x_j b_j) = \sum_{j=1}^{n} x_j T(b_j) = \sum_{j=1}^{n} x_j \tau(b_j) =$$

$$\sum_{j=1}^{n} x_j T'(b_j) = T'(\sum_{j=1}^{n} x_j b_j) = T'(v).$$

Per quanto riguarda l'esistenza basta definire, per ogni $v = \sum_{j=1}^{n} x_j b_j$, $T(v) = \sum_{j=1}^{n} x_j \tau(b_j)$ e poi verificare, molto facilmente, che tale funzione soddisfa quanto richiesto. □

Questa proposizione ci dice, rapidamente: ogni applicazione lineare è completamente determinata dal suo comportamento su una base.

La prossima definizione è un'altro dei cardini su cui ruota l'algebra lineare:

Definizione 20. *Siano V uno spazio vettoriale di dimensione n, W uno spazio vettoriale di dimensione m, $\mathcal{B} = (b_1, \ldots, b_n)$, $\mathcal{C} = (c_1, \ldots, c_m)$ basi per V e W rispettivamente; sia infine $T \in \mathrm{Hom}(V, W)$. La* matrice di T rispetto alle basi $\mathcal{B}$ *e* $\mathcal{C}$ *è la matrice $\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T) = A \in \mathcal{M}(m \times n; \mathbb{K})$ data dalla seguente regola:*

$$\forall j = 1, \ldots, n \quad T(b_j) = \sum_{i=1}^{m} a_{ij} c_i$$

In parole, per ogni j, la colonna $j-$esima della matrice è data dalle coordinate di $T(b_j)$ rispetto alla base $\mathcal{C}$. Se $\mathcal{B}$ e $\mathcal{C}$ sono basi standard si parla semplicemente di matrice associata a T.

In modo del tutto analogo a quanto abbiamo fatto nel superesempio, possiamo dare delle operazioni all'insieme W^V in modo che diventi uno spazio vettoriale; è facile vedere che $(T+T')\circ S = T\circ S + T'\circ S$, che $S\circ(T+T') = S\circ T + S\circ T'$ e che $(\lambda T)\circ S = \lambda(T\circ S) = T\circ(\lambda S)$ tutte le volte che le operazioni hanno senso. Ma soprattutto si vede facilmente che $\mathrm{Hom}(V,W)$ è un sottospazio vettoriale di W^V.

Proposizione 15. $\dim \mathrm{Hom}(V,W) = (\dim V)(\dim W)$.

Dimostrazione. Posto $n = \dim V, \quad m = \dim W$ e scelte le basi $\mathcal{B}$ per V e $\mathcal{C}$ per Wè facile verificare che la funzione

$$\mathrm{Hom}(V,W) \xrightarrow{\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}} \mathcal{M}(m\times n;\mathbb{K})$$

è un isomorfismo. □

L'utilità principale dalla matrice associata risiede nella seguente

Proposizione 16. *Se* $v = \sum_{i=1}^n x_i b_i \quad A = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T) \quad T(v) = \sum_{j=1}^m y_j c_j$ *allora*

$$A(x_1,\ldots,x_n)^t = (y_1,\ldots,y_m)^t.$$

In particolare $v \in \mathrm{Ker}T \Leftrightarrow (x_1,\ldots,x_n)^t \in \mathrm{Sol}(A,0)$.

Dimostrazione.

$$\sum_{j=1}^m y_j c_j = T(v) = T(\sum_{i=1}^n x_i b_i) = \sum_{i=1}^n x_i T(b_i) =$$
$$\sum_{i=1}^n x_i(\sum_{j=1}^m a_{ji}c_j) = \sum_{j=1}^m(\sum_{i=1}^n a_{ji}x_i)c_j$$

e, per l'unicità delle coordinate, abbiamo che $\forall j \quad y_j = \sum_{i=1}^n a_{ji}x_i$. □

Proposizione 17. *Se* $A = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T)$ *allora* $\mathrm{rg}(A) = \mathrm{rg}(T)$.

Dimostrazione. Siccome la trasposizione $\mathcal{M}(m\times 1;\mathbb{K}) \longrightarrow \mathbb{K}^m$ e $C_{\mathcal{C}} : \mathbb{K}^m \longrightarrow W$ sono isomorfismi e

$$\forall i = 1,\ldots,m \quad C_{\mathcal{C}}(A^i)^t = T(b_i)$$

per il punto 4 della proposizione 13 è immediato che

$$\mathrm{rg}(A) = \dim \mathcal{L}(A^1,\ldots,A^n) = \dim \mathcal{L}(T(b_1),\ldots,T(b_n)).$$

La tesi segue ora dal fatto che $\mathcal{L}(T(b_1),\ldots,T(b_n)) = \mathrm{Im}T$, per il punto 2 della stessa proposizione. □

Abbiamo già visto un buon motivo per definire in quel modo apparentemente strampalato il prodotto di matrici, precisamente che consente di tradurre la terminologia dei sistemi lineari; ma ecco un altra ragione:

Teorema 7. *Siano $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ basi per gli spazi vettoriali V, W, U rispettivamente. Siano poi $T \in \mathrm{Hom}(V, W)$, $S \in \mathrm{Hom}(W, U)$. Allora*

1. *$S \circ T \in \mathrm{Hom}(V, U)$.*
2. *$\mathcal{M}^{\mathcal{B}}_{\mathcal{D}}(S \circ T) = \mathcal{M}^{\mathcal{C}}_{\mathcal{D}}(S)\mathcal{M}^{\mathcal{B}}_{\mathcal{C}}(T)$.*

Dimostrazione. La prima affermazione è una semplicissima verifica. La seconda un calcolo: chiamiamo n, m, p le dimensioni dei tre spazi vettoriali, $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{C}}(T), \quad B = \mathcal{M}^{\mathcal{C}}_{\mathcal{D}}(S), \quad C = \mathcal{M}^{\mathcal{B}}_{\mathcal{D}}(S \circ T)$.

$$\begin{aligned}
&\forall i = 1, \dots m \quad \sum_{s=1}^{p} c_{si} d_s = S \circ T(b_i) = S(T(b_i)) \\
&= S(\sum_{j=1}^{m} a_{ji} c_j) = \sum_{j=1}^{m} a_{ji} S(c_j) = \sum_{j=1}^{m} a_{ji} (\sum_{s=1}^{p} b_{sj} d_s) \\
&= \sum_{s=1}^{p} (\sum_{j=1}^{m} b_{sj} a_{ji}) d_s
\end{aligned}$$

e, per l'unicità delle coordinale abbiamo che, $\forall i, s \quad c_{si} = \sum_{j=1}^{m} b_{sj} a_{ji}$, che è la tesi. □

Corollario 1. *Siamo $\mathcal{B}$, $\mathcal{C}$ basi per gli spazi vettoriali V, W.*

1. *Se $T \in \mathrm{Hom}(V, W)$ è un isomorfismo allora $\mathcal{M}^{\mathcal{C}}_{\mathcal{B}}(T^{-1}) = \mathcal{M}^{\mathcal{B}}_{\mathcal{C}}(T)^{-1}$.*
2. *Se $\mathcal{B}', \mathcal{C}'$ sono altre basi per V, W e $T \in \mathrm{Hom}(V, W)$ allora*

$$\mathcal{M}^{\mathcal{B}'}_{\mathcal{C}'}(T) = \mathcal{M}^{\mathcal{C}'}_{\mathcal{C}}(\mathrm{id}_W)^{-1} \mathcal{M}^{\mathcal{B}}_{\mathcal{C}}(T) \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V)$$

Dimostrazione.

1. Visto che T è un isomorfismo, deve essere $n = m$. La tesi segue dal teorema osservando che $T^{-1} \circ T = \mathrm{id}_V, \quad T \circ T^{-1} = \mathrm{id}_W$ e $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(\mathrm{id}_V) = \mathcal{M}^{\mathcal{C}}_{\mathcal{C}}(\mathrm{id}_W) = I$.
2. Basta applicare due volte il teorema osservando che $\mathrm{id}_W \circ T = T \circ \mathrm{id}_V$.

□

Il prossimo risultato dimostra come l'algebra delle matrici possa trarre vantaggio dalla conoscenza delle applicazioni lineari.

Corollario 2. *Sia* $A \in \mathcal{M}(n \times n; \mathbb{K})$:

1. A *è invertibile se e solo se* $\mathrm{rg}(A) = n$.
2. *Se esiste* $B \in \mathcal{M}(n \times n; \mathbb{K})$ *tale che* $AB = I$ *allora* A *è invertibile e* $A^{-1} = B$.
3. *Se esiste* $B \in \mathcal{M}(n \times n; \mathbb{K})$ *tale che* $BA = I$ *allora* A *è invertibile e* $A^{-1} = B$.

Dimostrazione. Scegliamo uno spazio vettoriale V di dimensione n e una sua base $\mathcal{B}$ (va benissimo, anche se non è obbligatorio, scegliere $\mathbb{K}^n$ con la base standard), e sia $T \in \mathrm{Hom}(V, V)$ l'unica applicazione lineare tale che $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T) = A$.

1. Se A è invertibile e A^{-1} è la sua inversa sia $S \in \mathrm{Hom}(V, V)$ l'unica applicazione lineare tale che $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(S) = A^{-1}$. Allora $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T \circ S) = AA^{-1} = I$, da cui segue che $T \circ S = \mathrm{id}_V$; allora, per ogni $w \in V$ abbiamo che $w = T(S(w))$ e dunque T è suriettiva, cioè $\mathrm{rg}(T) = n$; dato che $\mathrm{rg}(A) = \mathrm{rg}(T)$ abbiamo la tesi.

   Viceversa se $\mathrm{rg}(A) = \mathrm{rg}(T) = n$ allora T è suriettiva e quindi è un isomorfismo. $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T^{-1})$ è l'inversa di A.

2. Sia $S \in \mathrm{Hom}(V, V)$ tale che $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(S) = B$. Allora $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T \circ S) = AB = I = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V)$, e dunque $T \circ S = \mathrm{id}_V$; come prima segue che T è un isomorfismo e quindi A è invertibile. Inoltre $A^{-1} = A^{-1}I = A^{-1}(AB) = (A^{-1}A)B = IB = B$.

3. Sia $S \in \mathrm{Hom}(V, V)$ tale che $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(S) = B$. Allora $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(S \circ T) = BA = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V)$, e dunque $S \circ T = \mathrm{id}_V$, ma allora, se $v \in \mathrm{Ker}T$ abbiamo che $v = S(T(v)) = S(0) = 0$, cioè T è iniettiva e quindi un isomorfismo; allora A è invertibile. Inoltre $A^{-1} = IA^{-1} = (BA)A^{-1} = B(AA^{-1}) = BI = B$.

□

3.2 Spazio duale

Definizione 21. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Lo spazio vettoriale* $\mathrm{Hom}(V, \mathbb{K})$ *viene indicato con V^* e chiamato le* spazio duale *di V.*

Ovviamente $\dim V^* = \dim V$; se $\mathcal{B} = (b_1, \ldots, b_n)$ è una base di V, per ogni $i = 1, \ldots, n$, definiamo un vettore $b_i^* \in V^*$ mediante

$$b_i^*(b_j) = \begin{cases} 1 & \text{se} \quad j = i \\ 0 & \text{se} \quad j \neq i \end{cases}$$

La $n-$upla ordinata $\mathcal{B}^* = (b_1^*, \ldots, b_n^*)$ è linearmente indipendente, supponiamo infatti di avere una n-upla ordinata $(\lambda_1, \ldots, \lambda_n)$ di scalari tali che $\sum_{i=1}^n \lambda_i b_i^* = 0$; questo 0 è naturalmente in vettore nullo di V^*, cioè la funzione che associa lo scalare 0 a tutti gli elementi di V. Ma allora, per ogni $j = 1, \ldots, n$ abbiamo

$$0 = 0(b_j) = (\sum_{i=1}^n \lambda_i b_i^*)(b_j) = \sum_{i=1}^n \lambda_i (b_i^*(b_j)) = \lambda_j.$$

Sarebbe altrettanto semplice dimostrare che tale n-upla è linearmente indipendente, ma, per motivi di dimensioni, è inutile. $\mathcal{B}^*$ è detta la *base duale* di $\mathcal{B}$.

Se $U \lhd V$ l'*ortogonale* di U è $U^\perp = \{\alpha \in V^* \text{ tali che } \alpha(u) = 0 \quad \forall u \in U\}$. È facile verificare che si tratta di un sottospazio di V^* ma possiamo fare di meglio:

Proposizione 18. *Sia U un sottospazio di V. Allora*

$$\dim U + \dim U^\perp = \dim V.$$

Dimostrazione. Poniamo $p = \dim U$ e $n = \dim V$.

Sia $(b_1, \ldots, b_p)$ una base di U e completiamola a una base $\mathcal{B}$ di V. Affermo che la $(n-p)$-upla ordinata $(b_{p+1}^*, \ldots, b_n^*)$ è una base per $U^\perp$, che avrà quindi dimensione $n - p$.

Non occorre dimostrare che questi vettori sono linearmente indipendenti, in quanto estratti da una base, ci resta solo da verificare che appartengono a $U^\perp$ e che lo generano. Se $u \in U$ allora $u = \sum_{i=1}^p x_i b_i$ e quindi, per ogni $j = p+1, \ldots, n$ abbiamo che $b_j^*(u) = b_j^*(\sum_{i=1}^p x_i b_i) = \sum_{i=1}^p x_i b_j^*(b_i) = 0$ per la definizione della base duale.

Inoltre, se $\alpha \in U^\perp \subseteq V^*$ esiste una $n-$upla ordinata di scalari $(x_1, \ldots, x_n)$ tale che $\alpha = \sum_{i=1}^n x_i b_i^*$ ma, per ogni $j = 1, \ldots, p, \quad b_j \in U$ e quindi $0 = \alpha(b_j) = (\sum_{i=1}^n x_i b_i^*)(b_j) = x_j$ e quindi $\alpha = \sum_{i=p+1}^n x_i b_i^*$. □

Se V, W sono spazi vettoriali sul campo $\mathbb{K}$ e $T \in \mathrm{Hom}(V, W)$ la *trasposta* di T è la funzione

$$T^t : W^* \longrightarrow V^*$$

definita da $T^t(\beta) = \beta \circ T$. È facile controllare che $T^t \in \mathrm{Hom}(W^*, V^*)$ ed il motivo del nome è dato dalla seguente

Proposizione 19. *Siano $\mathcal{B}$ una base di V, $\mathcal{C}$ una base di W e $T \in \mathrm{Hom}(V, W)$ allora*

$$\mathcal{M}_{\mathcal{B}^*}^{\mathcal{C}^*}(T^t) = (\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T))^t$$

Dimostrazione. Sia $n = \dim V$, $m = \dim W$, $A = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T)$, $B = \mathcal{M}_{\mathcal{B}^*}^{\mathcal{C}^*}(T^t)$. Allora, per ogni $i = 1, \ldots, m$ e per ogni $j = 1, \ldots, n$ abbiamo

$$T^t(c_i^*)(b_j) = (\sum_{k=1}^n b_{ki} b_k^*)(b_j) = \sum_{k=1}^n b_{ki}(b_k^*(b_j)) = b_{ji}$$

ma anche

$$T^t(c_i^*)(b_j) = (c_i^* \circ T)(b_j) = c_i^*(\sum_{k=1}^m a_{kj} c_k) = \sum_{k=1}^m a_{kj} c_i^*(c_k) = a_{ij}$$

e quindi $B = A^t$. □

Se $T \in \mathrm{Hom}(V, W)$ abbiamo i sottospazi vettoriali $\mathrm{Im}T^t$ e $(\mathrm{Ker}T)^\perp$ di V^* e $(\mathrm{Im}T)^\perp$ e $\mathrm{Ker}T^t$ di W^*.

Teorema 8. *Nelle condizioni appena descritte valgono le seguenti uguaglianze:*

1. $\mathrm{Im}T^t = (\mathrm{Ker}T)^\perp$.
2. $\mathrm{Ker}T^t = (\mathrm{Im}T)^\perp$.
3. $\mathrm{rg}T = \mathrm{rg}T^t$.

Dimostrazione. Consiste nel provare direttamente due delle inclusioni dei primi due punti, e poi nell'ottenere le altre inclusioni e il terzo punto per motivi di dimensione. Poniamo $n = \dim V = \dim V^*$, $m = \dim W = \dim W^*$ e cominciamo

Se $\alpha \in \mathrm{Im}T^t$ esiste $\beta \in W^*$ tale che $\alpha = T^t(\beta) = \beta \circ T$, ed allora, per ogni $v \in \mathrm{Ker}T$ avremo

$$\alpha(v) = \beta(T(v)) = \beta(0) = 0.$$

Quindi $\mathrm{Im}T^t \subseteq (\mathrm{Ker}T)^\perp$, e, in particolare $\dim \mathrm{Im}T^t \leq \dim(\mathrm{Ker}T)^\perp$.

Se $\beta \in \mathrm{Ker}T^t$ allora $0 = T^t(\beta) = \beta \circ T$; per ogni $w \in \mathrm{Im}(T)$ esiste $v \in V$ tale che $T(v) = w$ e quindi

$$\beta(w) = \beta(T(v)) = (\beta \circ T)(v) = 0(v) = 0.$$

Quindi $\mathrm{Ker}T^t \subseteq (\mathrm{Im}T)^\perp$, e, in particolare $\dim \mathrm{Ker}T^t \leq \dim(\mathrm{Im}T)^\perp$. Infine

$$\dim \mathrm{Im}T^t \leq \dim(\mathrm{Ker}T)^\perp \stackrel{*}{=} n - \dim \mathrm{Ker}T \stackrel{**}{=} \dim \mathrm{Im}T \stackrel{*}{=}$$

$$m - \dim(\mathrm{Im}T)^\perp \leq m - \dim \mathrm{Ker}T^t \stackrel{**}{=} \dim \mathrm{Im}T^t,$$

dove le uguaglianze indicate con $**$ sono dovute al teorema nullità+rango e quelle segnate con $*$ sono motivate nella proposizione 18.

Quindi tutte le disuguaglianze che compaiono nella catena sono in realtà uguaglianze ed il teorema è dimostrato. □

Corollario 3. *Siano m, n numeri naturali positivi e $A \in \mathcal{M}(m \times n, \mathbb{K})$ allora*

$$\mathrm{rg}_R(A) = \mathrm{rg}_C(A)$$

Dimostrazione. Come nella dimostrazione del corollario 2, scegliamo uno spazio vettoriale V di dimensione n e una sua base $\mathcal{B}$,uno spazio vettoriale W di dimensione m e una sua base $\mathcal{C}$ e sia $T \in \mathrm{Hom}(V, W)$ l'unica applicazione lineare tale che $\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(T) = A$. Allora $\mathcal{M}_{\mathcal{B}^*}^{\mathcal{C}^*}(T^t) = A^t$ e dunque

$$\mathrm{rg}_C(A) = \mathrm{rg}T = \mathrm{rg}T^t = \mathrm{rg}_C(A^t) = \mathrm{rg}_R(A).$$

□

Capitolo 4

Diagonalizzazione

4.1 Determinanti

In tutta questa sezione parleremo di matrici *quadrate* cioè appartenenti a $\mathcal{M}(n \times n; \mathbb{K})$; una matrice del genere è detta *singolare* se il suo rango è minore di n. Useremo la seguente notazione: se $A \in \mathcal{M}(n \times n; \mathbb{K})$ scriveremo $A = (A^1, \ldots, A^n)$.

Esiste uno strumento che permette, attraverso un calcolo facile fatto soltanto di somme e prodotti di elementi di $\mathbb{K}$, di rispondere a una sola domanda, ma decisiva: quand'è che una matrice quadrata ha le colonne (e quindi le righe) linearmente indipendenti? (cioè è non singolare?)

Teorema 9. *Per ogni naturale positivo n esiste un'unica funzione*

$$\det : \mathcal{M}(n \times n; \mathbb{K}) \longrightarrow \mathbb{K}$$

detta determinante *che gode delle seguenti proprietà:*

d1 tenendo fisse tutte le colonne meno una det *è un'applicazione lineare: cioè, per ogni $i = 1, \ldots, n$*

a)

$$\begin{aligned}&\det(A^1, \ldots, A^{i-1}, A^i + B^i, A^{i+1}, \ldots, A^n) = \\ &\det(A^1, \ldots, A^{i-1}, A^i, A^{i+1}, \ldots, A^n) + \\ &\det(A^1, \ldots, A^{i-1}, B^i, A^{i+1}, \ldots, A^n)\end{aligned}$$

b)

$$\begin{aligned}&\det(A^1, \ldots, A^{i-1}, \lambda A^i, A^{i+1}, \ldots, A^n) = \\ &\lambda \det(A^1, \ldots, A^{i-1}, A^i, A^{i+1}, \ldots, A^n)\end{aligned}$$

d2 *Se la matrice ha due colonne uguali allora il suo determinante vale* 0 *cioè, se esistono* $i < j \in \{1, \dots, n\}$ *tali che* $A^i = A^j$ *allora*

$$\det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) = 0$$

d3 $\det(I) = 1$.

La dimostrazione esula dai propositi di queste note; cerchiamo piuttosto di convincerci che le cose stanno proprio così: esaminiamo il caso in cui $n = 1$: una matrice $A \in \mathcal{M}(1 \times 1; \mathbb{K})$ è del tipo $A = (a)$ dove $a \in \mathbb{K}$; allora, usando *d1* e *d3* abbamo che $\det(A) = a \det((1)) = a$.

Notiamo soltanto che se in una matrice si scambiano due colonne allora il determinante cambia di segno, infatti

$$0 = \det(A^1, \dots, A^i + A^j, \dots, A^j + A^i, \dots, A^n) =$$
$$\det(A^1, \dots, A^i, \dots, A^j + A^j, \dots, A^n) + \det(A^1, \dots, A^j, \dots, A^j + A^j, \dots, A^n) =$$
$$\det(A^1, \dots, A^i, \dots, A^i, \dots, A^n) + \det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) +$$
$$\det(A^1, \dots, A^j, \dots, A^i, \dots, A^n) + \det(A^1, \dots, A^j, \dots, A^j, \dots, A^n) =$$
$$0 + \det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) + \det(A^1, \dots, A^j, \dots, A^i, \dots, A^n) + 0.$$

Prendiamo $n = 2$ e sia $A \in \mathcal{M}(2 \times 2; \mathbb{K})$ cioè

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

La prima colonna può essere scritta come $a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, e la seconda come $b \begin{pmatrix} 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, per cui, usando le proprietà caratterizzanti il determinante avremo

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \det \begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix} + c \det \begin{pmatrix} 0 & b \\ 1 & d \end{pmatrix} =$$
$$ab \det \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + ad \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + cb \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + cd \det \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} =$$
$$ad \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + cb \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = ad - bc.$$

Il determinante delle matrici $n \times n$ si calcola supponendo di conoscere quello delle matrici $(n-1) \times (n-1)$ secondo la regola che segue: si sceglie a

piacere una colonna, diciamo la $j-$esima (se ha dentro tanti zeri, tutto di guadagnato) e poi, per ogni $i = 1, \dots, n$ si chiama A_{ij} la matrice $(n-1) \times (n-1)$ ottenuta da A cancellando la $j-$ esima riga e la $i-$ esima colonna: poi si calcola

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

Qualcuno ha verificato per noi che la funzione così definita soddisfa le regole *d1,d2,d3*. Le proprietà rilevanti del determinante sono

1. $\det(AB) = \det(A)\det(B)$; (*formula di Binet).*
2. $\det(A^t) = \det(A)$.

In particolare la seconda proprietà ci assicura che tutto quanto è stato detto in questa sezione può essere ripetuto scambiando tra loro le parole *righe* e *colonne*. Ma il vero motivo per cui vale la pena di calcolare un determinante è il seguente

Teorema 10. $\forall A \in \mathcal{M}(n \times n; \mathbb{K}) \quad \det(A) = 0 \Leftrightarrow \mathrm{rg}(A) < n.$

Dimostrazione. $\Rightarrow$ Se $\mathrm{rg}(A) = n$ allora A è invertibile e

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

e quindi $\det(A) \neq 0$.

$\Leftarrow$ Supponiamo che $\mathrm{rg}(A) < n$, cioè le colonne di A sono linearmente dipendenti, e allora una di loro è combinazione lineare delle altre. Siccome il determinante cambia solo segno quando si scambiano le colonne, possiamo supporre che sia

$$A^n = \sum_{s=1}^{n-1} \lambda_s A^s.$$

Usando la proprietà *d1* si ha che

$$\det(A) = \det(A^1, \dots, A^{n-1}, \sum_{s=1}^{n-1} \lambda_s A^s) = \sum_{s=1}^{n-1} \lambda_s \det(A^1, \dots, A^{n-1}, A^s)$$

e quest'ultimo numero è la somma di $n-1$ addendi tutti uguali a zero in quanto determinanti di matrici con due colonne uguali. $\square$

4.2 Diagonalizzazione

In questa sezione ci occuperemo di *operatori* su uno spazio vettoriale V, cioè degli elementi di $\mathrm{Hom}(V,V)$; in particolare , dato un operatore T su V cercheremo una base $\mathcal{B}$ per V in modo che la matrice $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$ si presenti in una forma particolarmente maneggevole.

Definizione 22. *Si dice che uno scalare λ è un* autovalore *di T se esiste un vettore $v \neq 0$ tale che $T(v) = \lambda v$.*

Osserviamo che 0 è un autovalore se e solo se $\mathrm{Ker} T \neq \{0\}$, cioè T non è iniettivo. Più in generale λ è un autovalore se e solo se l'operatore $\lambda \,\mathrm{id}_V - T$ non è iniettivo.

Definizione 23. *Si dice che un vettore $v \neq 0$ è un* autovettore *di T se esiste uno scalare λ tale che $T(v) = \lambda v$, cioè se $v \in \mathrm{Ker}(\lambda \,\mathrm{id}_V - T)$. Tale λ è detto l'*autovalore *di v.*

Definizione 24. *Un operatore $T \in \mathrm{Hom}(V,V)$ è detto* diagonalizzabile *se esiste una base di V costituita interamente da autovettori.*

Il motivo della definizione è che, se $\mathcal{B}$ è una base formata da autovettori di T, $\forall b_i \in \mathcal{B}$, λ_i è l'autovalore di b_i e $\dim V = n$ allora

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Cioè $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$ è una matrice *diagonale*. Viceversa se $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$ è diagonale allora tutti gli elementi di $\mathcal{B}$ sono autovettori.

Il nostro problema sarà, dato un operatore, scoprire se è diagonalizzabile e, in caso affermativo, di trovare una base formata da autovettori.

Prima di tutto si devono cercare tutti gli autovalori del nostro operatore. Sia $T \in \mathrm{Hom}(V,V)$, $\mathcal{B}$ una base di V e sia $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$. Consideriamo la funzione $P_A : \mathbb{K} \longrightarrow \mathbb{K}$ definita da

$$P_A(t) = \det(tI - A) = \det \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix}$$

La funzione $P_A(t)$ è un polinomio di grado n, il coefficiente di t^n è 1 e il termine noto è $P_A(0) = \det(-A) = (-1)^n \det(A)$.

Naturalmente $P_A(t)$ dipende dall'operatore T e, ad occhio, ci si potrebbe aspettare che dipenda anche dalla scelta della base che si usa per trasformare T in una matrice; ma non è così:

Proposizione 20. *Siano $T \in \mathrm{Hom}(V, V)$, $\mathcal{B}$, $\mathcal{C}$ basi di V, $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$ e $B = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(T)$. Allora le funzioni $P_A(t)$ e $P_B(t)$ coincidono.*

Dimostrazione. Poniamo $P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V)$, per il punto 2 del teorema 7 avremo che $B = P^{-1}AP$ e le proprietà del prodotto di matrici ci permettono di affermare che, per ogni $t \in \mathbb{K}$:

$$\begin{aligned}
&P_B(t) = \det(t\,I - B) = \det(t\,P^{-1}IP - P^{-1}AP) = \\
&\det(P^{-1}(tIP - AP)) = \det(P^{-1}(tI - A)P) = \\
&\det(P^{-1})\det(tI - A)\det(P) = \det(P^{-1})\det(P)\det(tI - A) = \\
&\det(P^{-1}P)\det(tI - A) = \det I \det(tI - A) = P_A(t).
\end{aligned}$$

□

Definizione 25. *Il polinomio $P_A(t)$, che, come abbiamo visto dipende solo dall'operatore T è detto* polinomio caratteristico *dell'operatore T e indicato con il simbolo $P_T(t)$.*

Proposizione 21. *Gli autovalori di T sono esattamente tutte le radici del polinomio $P_T(t)$.*

Dimostrazione. Sia $\mathcal{B}$ una base di V e poniamo $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$; avremo che λ è un autovalore di T se e solo se l'operatore $\lambda I - T$ non è iniettivo. Ma questo avviene se e solo se $\mathrm{rg}(\lambda I - T) < n$, e, visto che $\lambda I - A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\lambda I - T)$, questa condizione è equivalente a $\mathrm{rg}(\lambda I - A) < n$, cioè $0 = \det(\lambda I - A) = P_A(\lambda) = P_T(\lambda)$. □

Ricordiamo qualche proprietà dei polinomi con cui la lettrice sarà certamente in buona confidenza: questo richiamo è tuttavia opportuno almeno per fissare la notazione.

L'insieme di tutti i polinomi viene indicato con il simbolo $\mathbb{K}[t]$; due polinomi possono essere sommati e moltiplicati; si può anche dividere per un polinomio non nullo, per la precisione: siano $P(t)$, $S(t) \neq 0 \in \mathbb{K}[t]$ allora esistono e sono unici due polinomi $Q(t)$, $R(t) \in \mathbb{K}[t]$ tali che $P(t) = S(t)Q(t) + R(t)$

e, se $R(t) \neq 0$, allora $\partial R(t) < \partial S(t)$, dove con ∂ si indica il grado di un polinomio, che è definito come la lettrice sa bene per tutti e soli i polinomi non nulli. $R(t)$ che è chiamato il *resto* della divisione di $P(t)$ per $S(t)$. Se è il polinomio nullo si dice che $P(t)$ è *divisibile* per $S(t)$. Il caso più interessante si verifica quando $S(t)$ è del tipo $t-\alpha$, in particolare di grado 1. In tal caso si ha che $P(t)$ è divisibile per $t-\alpha$ se e solo se $P(\alpha)=0$, cioè α è una radice di $P(t)$. Si dice che la radice α del polinomio non nullo $P(t)$ ha *molteplicità algebrica* m se $P(t)$ è divisibile per $(t-\alpha)^m$ ma non per $(t-\alpha)^{m+1}$, cioè se esiste un polinomio $Q(t)$ tale che $P(t)=(t-\alpha)^m Q(t)$ e $Q(\alpha)\neq 0$. Questo numero m viene indicato con il simbolo m.a.(α).

Naturalmente la somma di tutte le molteplicità algebriche di tutte le radici di un polinomio non supera il suo grado, e ci sono dei casi in cui è minore, per esempio il polinomio $P(t)=t^3+t$ che ha solo la radice 0 e m.a.$(0)=1<3=\partial P(t)$.

In particolare, se λ è un autovalore dell'operatore $T\in \mathrm{Hom}(V,V)$ ha senso parlare della sua molteplicità algebrica pensando λ come radice del polinomio caratteristico.

Inoltre se λ è un autovalore di T il sottospazio $\mathrm{Ker}(\lambda\,\mathrm{id}_V - T) = \{v \in V \text{ tali che } T(v)=\lambda v\}$ è chiamato l' *autospazio* di λ, e indicato con il simbolo V_λ o $V_{\lambda,T}$ se sussiste pericolo di confusione; la sua dimensione si chiama *molteplicità geometrica* di λ e si indica con m.g.(λ). Gli elementi di V_λ sono il vettore nullo e tutti gli autovettori che fanno capo a λ.

Proposizione 22. *Se λ è un autovalore dell'operatore T allora*

$$\text{m.g.}(\lambda) \leq \text{m.a.}(\lambda).$$

Dimostrazione. Possiamo calcolare il polinomio caratteristico usando una qualsiasi base per trasformare T in una matrice; procediamo in questo modo: siano $p=\text{m.g.}(\lambda)$, $n=\dim V$ e sia $(b_1,\dots,b_p)$ una base di V_λ ed estendiamola a una base $\mathcal{B}$ di V. La matrice $A=\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$ sarà dunque del tipo

$$A=\begin{pmatrix} \lambda & 0 & \dots & 0 & a_{1(p+1)} & \dots & a_{1n} \\ 0 & \lambda & \dots & 0 & a_{2(p+1)} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & a_{p(p+1)} & \dots & a_{pn} \\ 0 & 0 & \dots & 0 & a_{(p+1)(p+1)} & \dots & a_{(p+1)n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{n(p+1)} & \dots & a_{nn} \end{pmatrix}$$

per cui il polinomio caratteristico sarà

$$\det(tI-A) = \det \begin{pmatrix} t-\lambda & 0 & \dots & 0 & -a_{1(p+1)} & \dots & -a_{1n} \\ 0 & t-\lambda & \dots & 0 & -a_{2(p+1)} & \dots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & t-\lambda & -a_{p(p+1)} & \dots & -a_{pn} \\ 0 & 0 & \dots & 0 & t-a_{(p+1)(p+1)} & \dots & -a_{(p+1)n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & -a_{n(p+1)} & \dots & t-a_{nn} \end{pmatrix}$$

Chiamando B la matrice quadrata $(n-p)\times(n-p)$ ottenuta usando le ultime righe e colonne di A e sviluppando il determinante di ottiene

$$P_T(t) = (t-\lambda)^p \det(tI-B)$$

da cui segue che $p \leq \text{m.a.}(\lambda)$. □

Abbiamo parlato di somma di due sottospazi, ma niente ci impedisce di ripetere la definizione quando ne abbiamo di più.

Definizione 26. *Sia V uno spazio vettoriale, $U_1, \dots, U_k$ siano sottospazi vettoriali di V; la* somma *di $U_1, \dots, U_k$ è il sottospazio di V dato da*

$$\sum_{i=1}^{k} U_i = \{v \text{ tali che } \forall i = 1, \dots, k \; \exists u_i \in U_i \text{ tale che } v = \sum_{i=1}^{k} u_i\}$$

La somma dei sottospazi $U_1, \dots, U_k$ è detta *diretta* se l'unico modo di ottenere il vettore nullo nella forma $v = \sum_{i=1}^{k} u_i$ con $u_i \in U_i$, $\forall i = 1; \dots, k$ è di prendere $u_i = 0$, $\forall i$. Se la somma è diretta si usa la notazione $\bigoplus_{i=1}^{k} U_i$. Non esiste una formula di Grassmann per la somma di tre o più sottospazi, ma, se la somma è diretta allora $\dim \bigoplus_{i=1}^{k} U_i = \sum_{i=1}^{k} \dim U_i$. La dimostrazione, concettualmente non difficile, è complicata da una notazione necessariamente pesante: eccola. Per ogni $i = 1, \dots, k$ poniamo $n_i = \dim U_i$ e sia

$$\mathcal{B}_i = (b_{i1}, \dots, b_{in_i})$$

una base di U_i. La nostra tesi seguirà dal fatto, che proveremo, che l'insieme

$$\mathcal{B} = (b_{11}, \dots, b_{1n_1} \dots, b_{k1}, \dots, b_{kn_k})$$

è una base di $\bigoplus_{i=1}^{k} U_i$. Che $\mathcal{B}$ generi la somma si vede facilmente, proviamo dunque che sono linearmente indipendenti. Siano

$$x_{11}, \ldots, x_{1,n_1}, \ldots, x_{k1}, \ldots, x_{k,n_k}$$

degli scalari tali che

$$\begin{array}{rccccccc} 0 = & x_{11}b_{11} & + & \ldots & + & x_{1n_1}b_{1n_1} & + \\ & \ldots & + & \ldots & + & \ldots & + \\ & x_{k1}b_{k1} & + & \ldots & + & x_{kn_k}b_{kn_k} & \end{array}$$

e poniamo, $\forall i = 1, \ldots, k\,, v_i = \sum_{j=1}^{n_i} x_{ij}b_{ij} \in U_i$. Avremo allora che $\sum_{i=1}^{k} v_i = 0$ e, visto che la somma è diretta, troviamo che $v_i = 0\,\forall i = 1, \ldots, k$. Dato che i vettori in $\mathcal{B}_i$ sono linearmente indipendenti si avrà che $\forall i = 1, \ldots, k\,, \forall j = 1, \ldots, n_i, \quad x_{ij} = 0$.

Proposizione 23. *Siano $\lambda_1, \ldots, \lambda_k$ autovalori distinti di un operatore $T \in \mathrm{Hom}(V, V)$, per ogni $i = 1, \ldots, k$ sia v_i un autovettore in V_{λ_i}. Allora $v_1, \ldots, v_k$ sono linearmente indipendenti.*

Dimostrazione. Si effettua per induzione sul numero k. Se $k = 1$ abbiamo un solo vettore, non nullo in quanto autovettore, e quindi linearmente indipendente.

Supponiamo ora che la proposizione sia vera per $k-1$ autovettori (ipotesi induttiva) e dimostriamola quando ne abbiamo k. Sia

$$\mu_1 v_1 + \ldots + \mu_{k-1} v_{k-1} + \mu_k v_k = 0 \tag{4.1}$$

Moltiplicando questa equazione per λ_k otteniamo

$$\lambda_k \mu_1 v_1 + \ldots + \lambda_k \mu_{k-1} v_{k-1} + \lambda_k \mu_k v_k = \lambda_k 0 = 0. \tag{4.2}$$

Applicando T all'equazione (4.1) si ottiene

$$\lambda_1 \mu_1 v_1 + \ldots + \lambda_{k-1} \mu_{k-1} v_{k-1} + \lambda_k \mu_k v_k = T(0) = 0. \tag{4.3}$$

Sottraendo (4.3) da (4.2) si ha

$$(\lambda_k - \lambda_1)\mu_1 v_1 + \ldots (\lambda_k - \lambda_{k-1})\mu_{k-1} v_{k-1} = 0 - 0 = 0$$

da cui segue, visto che $v_1, \ldots, v_{k-1}$ sono linearmente indipendenti per l'ipotesi induttiva, che $\forall i = 1, \ldots, k-1, \quad (\lambda_k - \lambda_i)\mu_i = 0$ e, dato che $\lambda_k - \lambda_i \neq 0$, abbiamo che $\mu_i = 0\,\forall i = 1, \ldots, k-1$. Infine, sostituendo quest'ultimo risultato in (4.1) in cui $v_k \neq 0$, si ha che anche $\mu_k = 0$. □

Corollario 4. *La somma di autospazi relativi ad autovalori distinti è diretta.*

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori in questione e, per ogni $i = 1, \ldots, k$ siano $v_i \in V_{\lambda_i}$ tali che $\sum_{i=1}^{k} v_i = 0$. Supponiamo per assurdo che $s \geq 1$ di essi siano non nulli. Cambiando se necessario l'ordine degli addendi possiamo supporre che sia $\sum_{i=1}^{s} v_i = 0$, con $v_i \neq 0 \quad \forall i = 1, \ldots, s$. Ma questa è una combinazione lineare di autovettori che fornisce il vettore nullo, pur avendo tutti i coefficienti uguali a 1, e contraddice la proposizione appena dimostrata. □

Teorema 11. *Sia T un operatore sullo spazio vettoriale V di dimensione n e sia $\mathcal{A}(T)$ l'insieme di tutti i suoi autovalori. Le seguenti condizioni sono equivalenti:*

1. *T è diagonalizzabile.*
2. *$\sum_{\lambda \in \mathcal{A}(T)} \text{m.a.}(\lambda) = n$ e $\forall \lambda \in \mathcal{A}(T)\, \text{m.a.}(\lambda) = \text{m.g.}(\lambda)$.*
3. *$\bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda = V$.*

Dimostrazione. Cominciamo con il premettere una successione di diseguaglianze che abbiamo provato in questa sezione:

$$n = \partial P_T(t) \geq \sum_{\lambda \in \mathcal{A}(T)} \text{m.a.}(\lambda) \geq \sum_{\lambda \in \mathcal{A}(T)} \text{m.g.}(\lambda) = \dim \bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda \leq \dim V = n.$$

1) $\Rightarrow$ 3). Sia $\mathcal{B}$ una base di V costituita da autovettori e, per ogni $\lambda \in \mathcal{A}(T)$ poniamo $\mathcal{B}_\lambda = \mathcal{B} \cap V_\lambda$, cioè raggruppiamo insieme gli elementi di $\mathcal{B}$ che hanno lo stesso autovalore. Le seguenti osservazioni sono ovvie

- Se $\lambda \neq \mu$ allora $\mathcal{B}_\lambda \cap \mathcal{B}_\mu = \emptyset$.
- $\bigcup_{\lambda \in \mathcal{A}(T)} \mathcal{B}_\lambda = \mathcal{B}$.
- Se indichiamo con $\sharp(\mathcal{B}_\lambda)$ il numero degli elementi di $\mathcal{B}_\lambda$ abbiamo che $\sharp(\mathcal{B}_\lambda) \leq \text{m.g.}(\lambda)$ in quanto sono vettori estratti da una base e quindi linearmente indipendenti.

Da qui si deduce subito che

$$n = \sharp(\mathcal{B}) = \sum_{\lambda \in \mathcal{A}(T)} \sharp(\mathcal{B}_\lambda) \leq \sum_{\lambda \in \mathcal{A}(T)} \text{m.g.}(\lambda) = \dim \bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda$$

da cui si ha subito la tesi.

3) $\Rightarrow$ 1). Per ogni $\lambda \in \mathcal{A}(T)$ prendiamo una base $\mathcal{B}_\lambda$ di V_λ; abbiamo già visto che $\mathcal{B} = \bigcup_{\lambda \in \mathcal{A}(T)} \mathcal{B}_\lambda$ è una base per $\bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda = V$, ed è chiaramente costituita da autovettori.

3) $\Rightarrow$ 2) Se vale la condizione *3)* allora l'ultima disuguaglianza che abbiamo premesso a questa dimostrazione è un'uguaglianza, e quindi anche le altre due lo sono. La seconda, in particolare, può essere verificata solo quando $\forall \lambda \in \mathcal{A}(T)$, m.a.$(\lambda)$ = m.g.(λ) e quindi vale la *2).*

2) $\Rightarrow$ 3) Se vale la *2)* allora le prime due disuguaglianze sono delle uguaglianze e quindi anche la terza lo è; cioè $\dim \bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda = n = \dim V$. Ne segue che $\bigoplus_{\lambda \in \mathcal{A}(T)} V_\lambda = V$. □

Capitolo 5

Spazi euclidei

5.1 Forme bilineari

Definizione 27. *Sia V uno spazio vettoriale. Una* forma bilineare *su V è una funzione*

$$f : V \times V \longrightarrow \mathbb{K}$$

che soddisfa le seguenti condizioni:

B1 $\forall v, v', w \in V \quad f(v + v', w) = f(v, w) + f(v', w)$

B2 $\forall v, w \in V, \forall \lambda \in \mathbb{K} \quad f(\lambda v, w) = \lambda f(v, w)$

B3 $\forall v, w, w' \in V \quad f(v, w + w') = f(v, w) + f(v, w')$

B4 $\forall v, w \in V, \forall \lambda \in \mathbb{K} \quad f(v, \lambda w) = \lambda f(v, w)$

L'insieme di tutte le forma bilineari su V viene indicato con il simbolo $\mathrm{Bil}(V)$; è facile verificare che si tratta di un sottospazio vettoriale di $\mathbb{K}^{V \times V}$. Se $f \in \mathrm{Bil}(V)$ allora $f(0, w) = f(00, w) = 0f(0, w) = 0$, analogamente $f(v, 0) = 0$.

Un esempio importantissimo di forma bilineare può essere ottenuto come segue: Sia $A \in \mathcal{M}(n \times n; \mathbb{K})$, prendiamo $V = \mathbb{K}^n$ e definiamo la funzione $V \times V \xrightarrow{f} \mathbb{K}$ mediante $f(v, w) = vAw^t$. Le proprietà delle operazioni sulle matrici ci dicono che f è effettivamente una forma bilineare.

Se $\mathcal{B} = (b_1, \ldots, b_n)$ è una base per V e $f \in \mathrm{Bil}(V)$ la *matrice di f* rispetto a $\mathcal{B}$ è la matrice $A \in \mathcal{M}(n \times n; \mathbb{K})$ data da $a_{ij} = f(b_i, b_j)$; questa matrice viene indicata con il simbolo $\mathcal{M}_{\mathcal{B}}(f)$.

Se $v = \sum_{i=1}^n x_i b_i$, $w = \sum_{j=1}^n y_j b_j$ avremo

$$f(v,w) = f(\sum_{i=1}^n x_i b_i, w) = \sum_{i=1}^n x_i f(b_i, w) = \sum_{i=1}^n x_i f(b_i, \sum_{j=1}^n y_j b_j) =$$
$$\sum_{i,j=1}^n x_i f(b_i, b_j) y_j = \sum_{i,j=1}^n x_i a_{ij} y_j = (x_1, \ldots, x_n) A (y_1, \ldots, y_n)^t.$$

È facile, vedere che la funzione

$$\mathrm{Bil}(V) \xrightarrow{\mathcal{M}_{\mathcal{B}}} \mathcal{M}(n \times n; \mathbb{K})$$

è un isomorfismo, e quindi $\dim \mathrm{Bil}(V) = (\dim V)^2$. Se $\mathcal{C}$ è un'altra base di V, $B = \mathcal{M}_{\mathcal{C}}(f)$, $P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V)$, allora

$$b_{ij} = f(c_i, c_j) = f(\sum_{s=1}^n p_{si} b_s, c_j) = \sum_{s=1}^n p_{si} f(b_s, c_j) =$$
$$\sum_{s=1}^n p_{si} f(b_s, \sum_{t=1}^n p_{tj} b_t) = \sum_{s,t=1}^n p_{si} f(b_s, b_t) p_{tj} = \sum_{s,t=1}^n p_{si} a_{st} p_{tj} = (p^t a p)_{ij}$$

e quindi le matrici A, B, P sono legate dalla formula

$$B = P^t A P.$$

Definizione 28. *Sia $f \in \mathrm{Bil}(V)$. Il* nucleo sinistro *di f il sottoinsieme*

$$\mathrm{Ker}^S f = \{v \in V \text{ tali che } f(v,w) = 0 \quad \forall w \in V\}$$

mentre il nucleo destro *di f è*

$$\mathrm{Ker}^D f = \{v \in V \text{ tali che } f(w,v) = 0 \quad \forall w \in V\}.$$

Si vede molto facilmente che sono entrambi sottospazi di V.
Se $\mathcal{B}$ è una base di V possiamo controllare quanto segue:

$$v \in \mathrm{Ker}^S f \Leftrightarrow f(v, b_i) = 0 \quad \forall i = 1, \ldots, n$$

L'implicazione $\Rightarrow$ è ovvia. Viceversa se $v \in V$ soddisfa quanto richiesto allora, per ogni $w = \sum_{i=1}^n x_i b_i$, avremo

$$f(v,w) = \sum_{i=1}^n x_i f(v, b_i) = 0.$$

Analogamente

$$v \in \mathrm{Ker}^D f \Leftrightarrow f(b_i, v) = 0 \quad \forall i = 1, \ldots, n.$$

Da questo segue facilmente che, posto $A = \mathcal{M}_{\mathcal{B}}(f)$, $\quad v = \sum_{i=1}^{n} x_i b_i$ allora $v \in \mathrm{Ker}^S f \Leftrightarrow (x_1, \ldots, x_n)A = 0$ e $v \in \mathrm{Ker}^D f \Leftrightarrow A(x_1, \ldots, x_n)^t = 0$. Se prendiamo $\mathbb{K} = \mathbb{R}, n = 2$ e

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

allora $\mathrm{Ker}^S f = \mathcal{L}(3b_1 - b_2)$ mentre $\mathrm{Ker}^D f = \mathcal{L}(2b_1 - b_2)$, e dunque i due nuclei sono diversi. Eppure

Teorema 12. *Siano V uno spazio vettoriale di dimensione n, $\mathcal{B}$ una base di V, $f \in \mathrm{Bil}(V)$ e $A = \mathcal{M}_{\mathcal{B}}(f)$ allora*

$$\dim \mathrm{Ker}^S f = n - \mathrm{rg}(A) = \dim \mathrm{Ker}^D f$$

Il numero $n - \dim \mathrm{Ker}^S f = n - \dim \mathrm{Ker}^D f$ viene detto *rango* della forma bilineare f e indicato con $\mathrm{rg}(f)$; esso coincide con il rango di una qualunque matrice di f.

Dimostrazione. Conviene procedere per una via indiretta che produce il risultato in maniera altamente spettacolare.

Cominciamo con il definire una funzione

$$\phi : \mathrm{Hom}(V, V^*) \longrightarrow \mathrm{Bil}(V)$$

mediante $\phi(T)(v, w) = T(v)(w)$, le seguenti verifiche sono molto facili e lasciate alla lettrice:

- Per ogni $T \in \mathrm{Hom}(V, V^*), \phi(T)$ appartiene effettivamente a $\mathrm{Bil}(V)$.
- ϕ è lineare.
- ϕ è iniettiva, e dunque, visto che il suo dominio e codominio hanno la stessa dimensione n^2, è un isomorfismo, in particolare suriettiva.

Inoltre

$$v \in \mathrm{Ker}T \Leftrightarrow T(v) = 0 \Leftrightarrow T(v)(w) = 0 \,\forall w \in V \Leftrightarrow$$
$$\phi(T)(v,w) = 0 \,\forall w \in V \Leftrightarrow v \in \mathrm{Ker}^S\phi(T)$$

Definiamo poi

$$\psi : \mathrm{Hom}(V, V^*) \longrightarrow \mathrm{Bil}(V)$$

mediante $\phi(S)(v,w) = T(w)(v)$, anche qui si vede subito che

- Per ogni $S \in \mathrm{Hom}(V, V^*), \psi(T)$ appartiene effettivamente a $\mathrm{Bil}(V)$.
- ψ è lineare.
- ψ è iniettiva, e dunque, visto che il suo dominio e codominio hanno la stessa dimensione n^2, è un isomorfismo, in particolare suriettiva.

Anche qui si vede che $\mathrm{Ker}S = \mathrm{Ker}^D\psi(S)$.

Sia $\mathcal{B}$ una base per V, $T \in \mathrm{Hom}(V, V^*)$, e poniamo $B = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}^*}(T)$, $A = \mathcal{M}_{\mathcal{B}}(\phi(T))$ allora, per ogni $i, j = 1, \ldots, n$

$$a_{ij} = \phi(T)(b_i, b_j) = T(b_i)(b_j) = \sum_{k=1}^{n} b_{ki} b_k^*(b_j) = \sum_{k=1}^{n} b_{ki}(b_k^*(b_j)) = b_{ji}$$

Cioè $B = A^t$. Sia poi $S \in \mathrm{Hom}(V, V^*)$, $C = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}^*}(S)$, $A = \mathcal{M}_{\mathcal{B}}(\psi(S))$ allora, per ogni $i, j = 1, \ldots, n$

$$a_{ij} = \psi(S)(b_i, b_j) = S(b_j)(b_i) = \sum_{k=1}^{n} ckj b_k^*(b_i) = \sum_{k=1}^{n} c_{kj}(b_k^*(b_i)) = c_{ij}$$

Cioè $A = C$, adesso possiamo facilmente concludere la dimostrazione del teorema: siano $f, \mathcal{B}, A$ come nell'enunciato, siccome ϕ è suriettiva esiste $T \in \mathrm{Hom}(V, V^*)$ tale che $\phi(T) = f$ e, visto che anche ψ è suriettiva esiste $S \in \mathrm{Hom}(V, V^*)$ tale che $\psi(S) = f$; ma allora

$$\dim \mathrm{Ker}^S f = \dim \mathrm{Ker}^S \phi(T) = \dim \mathrm{Ker}T = n - \dim \mathrm{Im}T = n - \mathrm{rg}A^t = n - \mathrm{rg}(A),$$

$$\dim \mathrm{Ker}^D f = \dim \mathrm{Ker}^D \psi(S) = \dim \mathrm{Ker}S = n - \dim \mathrm{Im}S = n - \mathrm{rg}A$$

e la dimostrazione è conclusa. □

Una forma bilineare $f \in \mathrm{Bil}(V)$ si dice *non degenere* se $\mathrm{Ker}^S f = \{0\}$ o, equivalentemente, $\mathrm{Ker}^D f = \{0\}$; naturalmente questa condizione può essere facilmente verificata prendendo una matrice A che rappresenti f e controllando se ha rango massimo, per esempio vedendo che $\det(A) \neq 0$. Ovviamente f è non degenere se per ogni $0 \neq v \in V$ $\exists w \in V$ tale che $f(v,w) \neq 0$ (oppure, ma è equivalente, per ogni $0 \neq v \in V$ $\exists w \in V$ tale che $f(w,v) \neq 0$).

Una forma bilineare $f \in \mathrm{Bil}(V)$ si dice *simmetrica* se $\forall v, w \in V$ $f(v,w) = f(w,v)$; una matrice quadrata A si dice *simmetrica* se $A = A^t$. La simmetria di una forma bilineare può essere facilmente controllata guardando una sua matrice rispetto a qualunque base:

Proposizione 24. *Sia $f \in \mathrm{Bil}(V)$; le seguenti condizioni sono equivalenti:*

1. *f è simmetrica.*
2. *Per ogni base $\mathcal{C}$ di V la matrice $\mathcal{M}_{\mathcal{C}}(f)$ è simmetrica.*
3. *Esiste una base $\mathcal{B}$ di V tale che la matrice $\mathcal{M}_{\mathcal{B}}(f)$ è simmetrica.*

Dimostrazione. 1) $\Rightarrow$ 2) Sia $B = \mathcal{M}_{\mathcal{C}}(f)$, allora, per ogni $i, j = 1, \ldots, n$ abbiamo $b_{ij} = f(c_i, c_j) = f(c_j, c_i) = b_{ji}$.

2) $\Rightarrow$ 3) È ovvio.

3) $\Rightarrow$ 1) Siano $A = \mathcal{M}_{\mathcal{B}}(f)$, $\quad v = \sum_{i=1}^n x_i b_i$, $\quad w = \sum_{j=1}^n y_j b_j$ e osserviamo che ogni matrice 1×1 è simmetrica. Avremo

$$f(v,w) = (x_1, \ldots, x_n)A(y_1, \ldots, y_n)^t = ((x_1, \ldots, x_n)A(y_1, \ldots, y_n)^t)^t = (y_1, \ldots, y_n)A^t(x_1, \ldots, x_n)^t = (y_1, \ldots, y_n)A(x_1, \ldots, x_n)^t = f(w,v).$$

□

D'ora in poi imponiamo, fino alla fine di queste note, una limitazione sul campo $\mathbb{K}$: per la precisione richiediamo che $\forall 0 \neq \lambda \in \mathbb{K}, 2\lambda = \lambda + \lambda \neq 0$, cioè $\mathbb{K}$ non ha *caratteristica* 2.

Questa richiesta non è soddisfatta sul campo $\mathbb{Z}_2$ ma lo è su $\mathbb{Q}$ o $\mathbb{R}$.

Teorema 13. *Sia V uno spazio vettoriale di dimensione n, $f \in \mathrm{Bil}(V)$; le seguenti condizioni sono equivalenti*

1. *f è simmetrica.*
2. *Esiste una base $\mathcal{B}$ di V tale che $f(b_i, b_j) = 0$ se $i \neq j$; cioè $\mathcal{M}_{\mathcal{B}}(f)$ è diagonale.*

In parole, una forma bilineare è simmetrica se e solo se è diagonalizzabile.

Dimostrazione. $2 \Rightarrow 1$ Sia $\mathcal{B}$ una base tale che $\mathcal{M}_{\mathcal{B}}(f)$ è diagonale; siccome ogni matrice diagonale è simmetrica la conclusione segue dalla proposizione precedente.

$1 \Rightarrow 2$ La dimostrazione si effettua per induzione su $n = \dim V$. Se $n = 1$ la conclusione è banalmente vera. Supponiamo dunque $n > 1$ e che il teorema sia vero su tutti gli spazi di dimensione $n - 1$, cioè ogni forma bilineare simmetrica su uno spazio di dimensione $n - 1$ è diagonalizzabile (ipotesi induttiva).

Sia $\dim V = n$ e $f \in \mathrm{Bil}(V)$ una forma bilineare simmetrica. Se f è la forma bilineare nulla, certamente è diagonalizzabile (ogni base va bene). Supponiamo allora che non lo sia, cioè esistono $v, w \in V$ tali che $f(v, w) \neq 0$. Dalla simmetria di f deduciamo facilmente la seguente *formula di polarizzazione*

$$f(v + w, v + w) = f(v, v) + 2f(v, w) + f(w, w).$$

Se $f(v, v) \neq 0$ poniamo $b_1 = v$ se invece $f(v, v) = 0$ ma $f(w, w) \neq 0$ poniamo $b_1 = w$ se, infine, $f(v, v) = 0 = f(w, w)$ la formula di polarizzaione ci garantisce che $f(v + w, v + w) = 2f(v, w) \neq 0$ e allora poniamo $b_1 = v + w$.

In tutti i casi siamo riusciti a trovare un vettore $b_1 \in V$ tale che $f(b_1, b_1) \neq 0$. Poniamo $W = \{v \in V \text{ tali che } f(v, b_1) = 0\}$. Si vede immediatamente che W è un sottospazio di V, ma affermo che $V = \mathcal{L}(b_1) \oplus W$. Alla luce dalla proposizione 6 dobbiamo provare che

- $\mathcal{L}(b_1) \cap W = \{0\}$ sia $v = \lambda b_1 \in \mathcal{L}(b_1)$ tale che $v \in W$, allora $0 = f(v, b_1) = f(\lambda b_1, b_1) = \lambda f(b_1, b_1)$ e, siccome $f(b_1, b_1) \neq 0$ delle proprietà elementari dei campi segue che $\lambda = 0$ e quindi $v = 0$.

- $V = \mathcal{L}(b_1) + W$, cioè, per ogni $v \in V$ esistono $\lambda \in \mathbb{K}$, $w \in W$ tali che $v = \lambda b_1 + w$; è facile verificare che scegliendo $\lambda = f(v, b_1) f(b_1, b_1)^{-1}$ e $w = v - \lambda b_1$ si vince.

Quindi $\dim W = n - 1$; sia g la restrizione di f a $W \times W$ (per un leggero abuso la si chiama semplicemente *restrizione* di f a W); evidentemente g è una forma bilineare simmetrica su W, che, per l'ipotesi induttiva ha una base $(b_2, \ldots, b_n)$ tale che $g(b_i, b_j) = f(b_i, b_j) = 0 \, \forall i \neq j = 2, \ldots, n$; la base $(b_1, b_2, \ldots, b_n)$ di V è quanto andavamo cercando in quanto, per ogni $j = 2, \ldots, n$, $f(b_1, b_j) = f(b_j, b_1) = 0$ dato che $b_j \in W$.

La dimostrazione è conclusa. □

Corollario 5. *Sia $A \in \mathcal{M}(n \times n; \mathbb{K})$, esiste $P \in \mathcal{M}(n \times n; \mathbb{K})$ non singolare tale che $P^t AP$ è diagonale se e solo se A è simmetrica.*

D'ora in poi, fino ad avviso contrario, il campo degli scalari sarà $\mathbb{R}$. Il vantaggio rispetto ad un campo qualunque è che ha senso parlare di scalari positivi (anche in $\mathbb{Q}$ c'è questa possibilità) e che si può estrarre la radice quadrata dei numeri reali positivi (e questo non si può fare sui razionali).

Supponiamo dunque che V sia uno spazio vettoriale di dimensione n su $\mathbb{R}$ (brevemente uno *spazio vettoriale reale*) e che f sia una forma bilineare simmetrica su V, sia $\mathcal{B}$ una base tale che $f(b_i, b_j) = 0$ quando $i \neq j$. Cambiando, se necessario, l'ordine in $\mathcal{B}$ possiamo trovare un numero naturale $p = 0, \ldots, n$ tale che $f(b_i, b_i) > 0$ se $i = 1, \ldots p$ e un numero naturale $q = 0, \ldots, n - p$ tale che $f(b_i, b_i) < 0$ se $i = p + 1, \ldots p + q$ e $f(b_i, b_i) = 0$ se $i = p + q + 1, \ldots n$.

A questo punto, ponendo

$$c_i = \begin{cases} \dfrac{b_i}{\sqrt{f(b_i, b_i}} & \text{se} \quad i = 1, \ldots, p \\ \dfrac{b_i}{\sqrt{-f(b_i, b_i}} & \text{se} \quad i = p + 1, \ldots, p + q \\ b_i & \text{se} \quad i = p + q + 1, \ldots, n \end{cases}$$

Si ha che $\mathcal{C}$ è una base $f-$*adattata*, cioè $\mathcal{M}_{\mathcal{B}}(f)$ è diagonale, sulla diagonale principale ha 1 sui primi p posti, -1 sui seguenti q posti e 0 nei rimanenti.

La coppia di numeri naturali (p, q) viene detta *segnatura di* f (rispetto alla base f-adattata $\mathcal{C}$).

In realtà la segnatura dipende soltanto dalla forma bilineare e non dalla base che si sceglie.

Teorema 14 (legge di inerzia). *Siano V uno spazio vettoriale reale di dimensione n, f una forma bilineare simmetrica su V, $\mathcal{B}$, $\mathcal{C}$ basi $f-$ adattate di V, (p, q) la segnatura di f rispetto a $\mathcal{B}$, (p', q') la segnatura di f rispetto a $\mathcal{C}$. Allora $p = p'$, $q = q'$.*

Dimostrazione. Certamente $p+q = p'+q'$ in quanto coincidono entrambi con il rango di f. Affermo che la $p+n-p'$-upla ordinata $(b_1, \ldots, b_p, c_{p'+1}, \ldots, c_n)$ è linearmente indipendente, da cui segue, per il teorema 3 che $p \leq p'$; scambiando tra loro le basi si ha l'altra disuguaglianza e quindi $p = p'$ che prova la tesi.

Supponiamo dunque di avere il vettore nullo come combinazione lineare dei vettori citati: cioè

$$\sum_{i=1}^{p} x_i b_i + \sum_{j=p'+1}^{n} y_j c_j = 0 \tag{5.1}$$

e poniamo $v = \sum_{i=1}^{p} x_i b_i$ ottenendo di conseguenza $-v = \sum_{j=p'+1}^{n} y_j c_j$

$$f(v,v) = \sum_{i,k=1}^{p} x_i x_k f(b_i, b_k) = \sum_{i=1}^{p} x_i^2 f(b_i, b_i) = \sum_{i=1}^{p} x_i^2 \geq 0 \tag{5.2}$$

e l'uguaglianza si verifica solamente se $x_i = 0,\ \forall i = 1, \ldots, p$; inoltre

$$f(-v,-v) = \sum_{j,k=p'+1}^{n} y_j y_k f(c_j, c_k) = \sum_{i=p'+1}^{n} y_j^2 f(c_j, c_j) = \sum_{i=p'+1}^{p'+q'} -y_j^2 \leq 0 \tag{5.3}$$

Siccome $f(v,v) = (-1)^2 f(-v,-v) = f(-v,-v)$ confrontando le equazioni 5.2 e 5.3 si ottiene che entrambe le disuguaglianze sono in realtà degli uguale. In particolare $x_i = 0,\ \forall i = 1, \ldots, p$ e dunque $v = 0$. Sostituendo nell'equazione 5.1 si trova che anche $y_j = 0,\ \forall i = p' + 1, \ldots, n$ in quanto i vettori $c_{p'+1}, \ldots, c_n$ sono linearmente indipendenti. □

Una forma bilineare simmetrica su uno spazio vettoriale reale di dimensione n è detta

- *definita positiva* o *prodotto scalare* se la sua segnatura è $(n, 0)$, cioè se la sua matrice rispetto ad una qualunque base $f-$ adattata è la matrice identica I.
- *definita negativa* se la sua segnatura è $(0, n)$, cioè se la sua matrice rispetto ad una qualunque base $f-$ adattata è la matrice $-I$.
- *semidefinita positiva* se la sua segnatura è $(p, 0)$, con $p \leq n$.
- *semidefinita negativa* se la sua segnatura è $(0, q)$, con $q \leq n$.

5.2 Prodotti scalari

I fisici trovano particolarmente interessante lo *spazio di Minkowski*, cioè ogni coppia ordinata (M, f) dove M è uno spazio vettoriale reale di dimensione 4 e f è una forma bilineare simmetrica su M di segnatura $(3, 1)$. Chiamano *eventi* gli elementi di M, *riferimento inerziale* ogni sua base $f-$ adattata e *trasformazioni di Lorentz* gli operatori su M che soddisfano $f(T(v), T(w)) = f(v, w)\ \forall v, w \in M$. A malincuore dobbiamo abbandonare questo filone perché altre scelte appaiono prioritarie.

Ci occuperemo invece di prodotti scalari.

Proposizione 25. *Una forma bilineare simmetrica $f \in \mathrm{Bil}(V)$ è un prodotto scalare se e solo se $\forall\, 0 \neq v \in V, \quad f(v, v) > 0$.*

Dimostrazione. Sia $\mathcal{B}$ una base f-adattata di V. Se $0 \neq v \in V$ esistono degli scalari non tutti nulli tali che $v = \sum_{i=1}^{n} x_i b_i$ ma allora $f(v, v) = \sum_{i,k=1}^{n} x_i x_k f(b_i, b_k) = \sum_{i=1}^{n} x_i^2 f(b_i, b_i) = \sum_{i=1}^{n} x_i^2 > 0$.

Viceversa, se vale questa condizione, allora , $\forall i = 1, \ldots, n,\ f(b_i, b_i) = 1$ e quindi f ha segnatura $(n, 0)$.

□

Siano f una forma bilineare simmetrica su V, $\mathcal{B}, \mathcal{C}$ basi di V, $A = \mathcal{M}_{\mathcal{B}}(f)$, $B = \mathcal{M}_{\mathcal{C}}(f)$ e $P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V)$. Allora $B = P^t A P$ e $\det(B) = \det(P^t)\det(A)\det(P) = \det(A)\det(P)^2$ da cui segue che $\det(A)$ e $\det(B)$ sono concordi (quando non sono nulli); in particolare se f è un prodotto scalare e $\mathcal{C}$ è f-adattata allora $B = I$ e quindi $\det(A) > 0$;ma si può fare di meglio: se una forma bilineare simmetrica f sia un prodotto scalare può essere visto eseguendo facili calcoli sulla matrice (simmetrica) A di f rispetto a una qualunque base: basta fare così: per ogni $s = 1, \ldots, n$ definiamo la matrice ${}^sA \in \mathcal{M}(s \times s; \mathbb{R})$ prendendo solo le prime s righe e s colonne (è detta l'$s-$ esimo *minore di nord-ovest* di A):

Proposizione 26. *f è un prodotto scalare se e solo se*

$$\forall s = 1, \ldots, n,\ \det({}^sA) > 0.$$

Dimostrazione. Sia $\mathcal{B}$ una base qualunque e $A = \mathcal{M}_{\mathcal{B}}(f)$. Per ogni $s = 1, \ldots, n$ sia $V_s = \mathcal{L}(b_1, \ldots, b_s)$; evidentemente V_s è un sottospazio di V di dimensione s e la matrice della restrizione a V_s della forma bilineare f rispetto alla base $(b_1, \ldots, b_s)$ è proprio ${}^sA$.

Supponiamo allora che f sia un prodotto scalare; allora la sua restrizione a V_s lo è ancora, e quindi, per l'osservazione appena fatta, $\det({}^sA) > 0$ e una delle due implicazioni (disgraziatamente la meno utile) è dimostrata.

L'altra deduzione procederà per induzione sulla dimensione di V. Se $\dim V = 1$ allora $A = (a)$ dove $a = f(b_1, b_1)$; $\det A > 0$ significa dunque che $f(b_1, b_1) > 0$, ma allora, per ogni $0 \neq v \in v,\ v = \lambda b_1$ con $\lambda \neq 0$ e dunque $f(v, v) = \lambda^2 f(b_1, b_1) > 0$ e f è effettivamente un prodotto scalare.

Supponiamo ora $n > 1$, che la proposizione sia vera su tutti gli spazi di dimensione $n-1$ e dimostriamola per quelli di dimensione n. Siccome l'ipotesi induttiva è piuttosto involuta vale la pena di ricordarla esplicitamente: essa recita

Se W è uno spazio vettoriale di dimensione $n-1$, g è una forma bilineare simmetrica su W, $\mathcal{C}$ è una base di W e $B = \mathcal{M}_{\mathcal{C}}(g)$ ha tutti i suoi $n-1$ minori di nord-ovest con determinante positivo, allora g è un prodotto scalare su W.

Supponiamo $\dim V = n$, con la notazione introdotta nell'enunciato, osserviamo che, per $s = 1, \ldots, n-1$, ${}^sA$ è anche l's-esimo minore di nord-ovest di ${}^{n-1}A$, la matrice della restrizione di f a V_{n-1}; il fatto che tutti queste matrici abbiano determinante positivo e che $\dim V_{n-1} = n-1$, ci garantisce, tramite l'ipotesi induttiva, che tale restrizione è un prodotto scalare.

Sia $(c_1, \ldots, c_{n-1})$ una base di V_{n-1} adattata rispetto a tale forma, cioè

$$\forall i, j = 1, \ldots, n-1 \quad f(c_i, c_j) = \begin{cases} 0 & \text{se} \quad i \neq j \\ 1 & \text{se} \quad i = j \end{cases}$$

Poniamo

$$c_n = b_n - \sum_{i=1}^{n-1} f(b_n, c_i) c_i$$

Siccome $\mathcal{L}(b_1, \ldots, b_{n-1}) = V_{n-1} = \mathcal{L}(c_1, \ldots, c_{n-1})$ è facile vedere che $\mathcal{L}(b_1, \ldots, b_n) = \mathcal{L}(c_1, \ldots, c_n)$, cioè che $(c_1, \ldots, c_n)$ è una base $\mathcal{C}$ di V.

Inoltre, per ogni $j = 1, \ldots, n-1$,

$$f(c_n, c_j) = f(b_n - \sum_{i=1}^{n-1} f(b_n, c_i) c_i, c_j) =$$

$$f(b_n, c_j) - \sum_{i=1}^{n-1} f(b_n, c_i) f(c_i, c_j) = f(b_n, c_j) - f(b_n, c_j) = 0$$

e quindi la matrice di f rispetto a $\mathcal{C}$ è

$$B = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(c_n, c_n) \end{pmatrix}$$

e quindi f è un prodotto scalare se e solo se $f(c_n, c_n) > 0$; d'altra parte $f(c_n, c_n) = \det(B)$ che è concorde con $\det(A)$ e quindi effettivamente è positivo.

□

Se f è un prodotto scalare si preferisce scrivere $\langle v, w \rangle$ invece di $f(v, w)$. Per esempio se $V = \mathbb{R}^n$ la forma bilineare data da

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^{n} x_i y_i = (x_1, \dots, x_n)(y_1, \dots, y_n)^t$$

è un prodotto scalare, detto il *prodotto scalare standard*. La sua matrice rispetto alla base standard è la matrice identica I.

Una base di V è detta *ortogonale* per $\langle\ ,\ \rangle$ se la matrice di $\langle\ ,\ \rangle$ rispetto ad essa è diagonale, ed è detta *ortonormale* se è $\langle\ ,\ \rangle$-adattata.

Uno *spazio euclideo* è una coppia ordinata $(V, \langle\ ,\ \rangle)$ dove V è uno spazio vettoriale e $\langle\ ,\ \rangle$ è un prodotto scalare su V.

Se $(V, \langle\ ,\ \rangle)$ è uno spazio euclideo e $v \in V$ possiamo definire la *norma* di v come $\|v\| = \sqrt{\langle v, v \rangle} \geq 0$. Ovviamente avremo che $\|v\| = 0 \Leftrightarrow v = 0$. Si dice che v è *unitario* se $\|v\| = 1$.

Proposizione 27 (disuguaglianza di Cauchy-Schwartz). *Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo. $\forall v, w \in V$ si ha che*

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

e vale l'eguaglianza se e solo se v, w sono linearmente dipendenti.

Dimostrazione. La dimostrazione è ovvia se $w = 0$, supponiamo dunque che $w \neq 0$ e osserviamo che v, w sono linearmente dipendenti se e solo se esiste $\lambda \in \mathbb{R}$ tale che $v - \lambda w = 0$ o, equivalentemente $\|v - \lambda w\|^2 = 0$. Per gestirci questa osservazione definiamo una funzione $g : \mathbb{R} \longrightarrow \mathbb{R}$ mediante

$$g(t) = \|v - tw\|^2 = \langle v - tw, v - tw \rangle = \|w\|^2 t^2 - 2\langle v, w \rangle t + \|v\|^2.$$

Si tratta di un polinomio di secondo grado, per ogni $t \in \mathbb{R}$, $g(t) \geq 0$ e esiste λ tale che $g(\lambda) = 0$ se e solo se v, w sono linearmente dipendenti. Il discriminante $\triangle$ di questa funzione sarà dunque sempre ≤ 0 e vale l'uguaglianza se e solo se v, w sono linearmente dipendenti. Adesso basta calcolare

$$\triangle = \langle v, w \rangle^2 - \|v\|^2 \, w\|^2.$$

La tesi si deduce con un ovvio passaggio algebrico seguito da un'estrazione di radice quadrata. □

La disuguaglianza appena provata ha una serie di importanti conseguenze:

Proposizione 28. *Sia* $(V, \langle\ ,\ \rangle)$ *uno spazio euclideo,* $v, w \in V$, $\lambda \in \mathbb{R}$;

- $\|\lambda v\| = |\lambda| \, \|v\|$.
- $\|v + w\| \leq \|v\| + \|w\|$.

Dimostrazione. La prima affermazione è facilissima. Per quanto riguarda le seconda, che si chiama *disuguaglianza triangolare della norma*, osserviamo che, trattandosi di una disuguaglianza tra numeri reali non negativi, la sua validità non viene alterata se eleviamo entrambi i membri al quadrato. Procediamo al calcolo avvisando la lettrice che, ad un certo punto, useremo la disuguaglianza di Cauchy-Schwartz:

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \leq$$
$$\|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

□

Su uno spazio euclideo $(V, \langle\ ,\ \rangle)$, ponendo $d(v, w) = \|v - w\|$ definiamo una *distanza*, cioè una funzione $d : V \times V \longrightarrow \mathbb{R}$ che gode delle seguenti proprietà

- $\forall v, w \in V, \quad d(v, w) \geq 0$.
- $\forall v, w \in V, \quad d(v, w) = d(w, v)$.
- $\forall v, w \in V, \quad d(v, w) = 0 \Leftrightarrow v = w$.
- $\forall v, w, u \in V, \quad d(v, w) \leq d(v, u) + d(u, w)$.

L'ultima proprietà si chiama *disuguaglianza triangolare della distanza.*

Inoltre, se $v \neq 0$, $w \neq 0$ la disuguaglianza di Cauchy-Schwartz si può scrivere come

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \, \|w\|} \leq 1$$

e così possiamo definire l' *angolo* tra v e w come quell'unico numero $\theta \in [0, \pi]$ tale che

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

Due vettori $v, w \in V$ si dicono *ortogonali* se $\langle v, w \rangle = 0$. il vettore nullo è ortogonale a ogni vettore, ma nessun altro vettore è ortogonale a se stesso.

Sappiamo già che esistono basi ortonormali di uno spazio euclideo, anzi, si può fare di meglio:

Teorema 15 (ortonormalizzazione di Gram-Schmidt). *Sia $\mathcal{B} = (b_1, \ldots, b_n)$ una base per lo spazio euclideo $(V, \langle \, , \, \rangle)$; allora*

- *Esiste una base ortogonale $\mathcal{C}$ di V tale che $\forall i = 1, \ldots, n$,*

$$\mathcal{L}(c_1, \ldots, c_i) = \mathcal{L}(b_1, \ldots, b_i).$$

- *Esiste una base ortonormale $\mathcal{D}$ di V tale che $\forall i = 1, \ldots, n$,*

$$\mathcal{L}(d_1, \ldots, d_i) = \mathcal{L}(b_1, \ldots, b_i).$$

Dimostrazione. Ci basta provare la prima affermazione, infatti, una volta trovata la base $\mathcal{C}$ con le proprietà conclamate, ci basterà definire $d_i = \dfrac{c_i}{\|c_i\|}$ per ottenere quanto enunciato al secondo punto.

Cominciamo col definire $c_1 = b_1$, e poi procediamo per induzione: in altri termini, supponendo di avere già trovato $c_1, \ldots, c_{i-1}$ con le proprietà richieste che, ricordiamo, sono:

- $\mathcal{L}(c_1, \ldots, c_{i-1}) = \mathcal{L}(b_1, \ldots, b_{i-1})$.
- $\forall j \neq k \in \{1, \ldots, i-1\}, \quad \langle c_j, c_k \rangle = 0$

definiamo

$$c_i = b_i - \sum_{j=0}^{i-1} \frac{\langle b_i, c_j \rangle}{\|c_j\|^2} c_j.$$

È facilissimo vedere che $\mathcal{L}(c_1,\dots,c_i)=\mathcal{L}(b_1,\dots,b_i)$ e ci resta solo da provare che c_i è ortogonale a $c_k\ \forall k=1,\dots,i-1$. Calcoliamo dunque

$$\langle c_i,c_k\rangle=\langle b_i-\sum_{j=0}^{i-1}\frac{\langle b_i,c_j\rangle}{\|c_j\|^2}c_j,c_k\rangle=\langle b_i,c_k\rangle-\sum_{j=0}^{i-1}\frac{\langle b_i,c_j\rangle}{\|c_j\|^2}\langle c_j,c_k\rangle=$$
$$\langle b_i,c_k\rangle-\frac{\langle b_i,c_k\rangle}{\|c_k\|^2}\langle c_k,c_k\rangle=0.$$

E notiamo che la dimostrazione ci spiega come trovare le basi con le proprietà desiderate. □

Se $U\lhd V$ poniamo $U^\perp=\{v\in V \text{ tali che } \langle v,u\rangle=0\ \forall u\in U\}$, cioé l'insieme di tutti i vettori ortogonali a tutti i vettori di U; è facile controllare che si tratta di un sottospazio di V ed è chiamato il *sottospazio ortogonale* a U. Notiamo che, se $\mathcal{B}=(b_1,\dots,b_p)$ è una base di U, per verificare che $v\in U^\perp$ ci basterà controllare che $\langle v,b_i\rangle=0\ \forall i=1,\dots,p$.

Proposizione 29. *Per ogni sottospazio vettoriale U dello spazio euclideo $(V,\langle\ ,\ \rangle)$ si ha che*

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

in particolare $\dim U^\perp=\dim V-\dim U$.

Dimostrazione. L'enunciato è banalmente vero se $U=\{0\}$, supponiamo dunque che non sia così.

$U\cap U^\perp=\{0\}$ perché solo il vettore nullo è ortogonale a se stesso.

Ci resta dunque da provare che, per ogni $v\in V$ riusciamo a trovare $u\in U$ e $w\in U^\perp$ tali che $v=u+w$.

Prendiamo una base ortonormale $\mathcal{B}=(b_1,\dots,b_p)$ di U e poi definiamo $u=\sum_{i=1}^{p}\langle v,b_i\rangle b_i\in U$, a questo punto ci basta verificare che $w=v-u\in U^\perp$ e cioè che $\langle w,b_j\rangle=0\ \forall j=1,\dots,p$. Si tratta di un calcolo molto simile al procedimento di Gram-Schmidt:

$$\langle v-u,b_j\rangle=\langle v,b_j\rangle-\langle\sum_{i=1}^{p}\langle v,b_i\rangle b_i,b_j\rangle=$$
$$\langle v,b_j\rangle-\sum_{i=1}^{p}\langle v,b_i\rangle\langle b_i,b_j\rangle=\langle v,b_j\rangle-\langle v,b_j\rangle=0.$$

□

5.3 Spazi euclidei complessi

In questa sezione il campo degli scalari è $\mathbb{C}$, gli spazi vettoriali su $\mathbb{C}$ vengono chiamati *spazi vettoriali complessi.*.

Definizione 29. *Sia V uno spazio vettoriale complesso. Una* forma sesquilineare *su V è una funzione*

$$f : V \times V \longrightarrow \mathbb{C}$$

che soddisfa le seguenti condizioni:

S1 $\forall v, v', w \in V \quad f(v + v', w) = f(v, w) + f(v', w)$

S2 $\forall v, w \in V, \forall \lambda \in \mathbb{C} \quad f(\lambda v, w) = \lambda f(v, w)$

S3 $\forall v, w, w' \in V \quad f(v, w + w') = f(v, w) + f(v, w')$

S4 $\forall v, w \in V, \forall \lambda \in \mathbb{C} \quad f(v, \lambda w) = \overline{\lambda} f(v, w)$

S5 $\forall v, w \in V,\ f(v, w) = \overline{f(w, v)}$

L'ultima richiesta ci dice in particolare che, se $v \in V,\ f(v, v) = \overline{f(v, v)}$ e quindi è un numero reale.

Se $\mathcal{B}$ è una base di V possiamo definire la matrice di f rispetto a $\mathcal{B}$ come la matrice A data da $a_{ij} = f(b_i, b_j)$ e ripetere gran parte di quanto detto a proposito delle forme bilineari. La matrice A soddisfa $A = \overline{A}^t$; matrici di questo genere vengono chiamate *matrici hermitiane.*

Definizione 30. *Se f è una forma sesquilineare su V si dice che f è un* prodotto hermitiano *se, per ogni $0 \neq v \in V,\ f(v, v) > 0$.*

Definizione 31. *Uno* spazio euclideo complesso *è una coppia ordinata $(V, \langle\ ,\ \rangle)$ dove V è uno spazio vettoriale complesso e $\langle\ ,\ \rangle$ è un prodotto hermitiano su V; il* prodotto hermitiano standard *su $\mathbb{C}^n$ è definito da*

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = \sum_{i=1}^{n} z_i \overline{w}_i$$

Sugli spazi euclidei complessi si può ripetere con variazioni minime il procedimento di Gram-Schmidt per trovare le basi ortonormali.

La base standard di $\mathbb{C}^n$ è una base ortonormale per il prodotto hermitiano standard.

5.4 Operatori su spazi euclidei

Definizione 32. *Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo reale e T un operatore su V. Si dice che T è*

- simmetrico *se, per ogni $v, w \in V$* $\langle T(v), w\rangle = \langle v, T(w)\rangle$.
- ortogonale *se, per ogni $v, w \in V$* $\langle T(v), T(w)\rangle = \langle v, w\rangle$.

Definizione 33. *Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo complesso e T un operatore su V. Si dice che T è*

- hermitiano *se, per ogni $v, w \in V$* $\langle T(v), w\rangle = \langle v, T(w)\rangle$.
- unitario *se, per ogni $v, w \in V$* $\langle T(v), T(w)\rangle = \langle v, w\rangle$.

Le proprietà citate nelle ultime due definizioni sono soddisfatte a condizione che lo siano su una base di V. Si tratta di un facile esercizio che la lettrice è invitata a risolvere per conto suo.

Proposizione 30. *Sia T un operatore su uno spazio euclideo complesso $(V, \langle\ ,\ \rangle)$, $\mathcal{B}$ una base di V, $A = \mathcal{M}_{\mathcal{B}}(\langle\ ,\ \rangle)$, $B = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(T)$, allora*

1. *T è hermitiano se e solo se $B^t A = A\overline{B}$, in particolare se $\mathcal{B}$ è ortonormale allora $A = I$ e quindi la condizione diventa $B^t = \overline{B}$, che equivale a $B = \overline{B}^t$,cioè la matrice B è hermitiana..*
2. *T è unitario se e solo se $B^t A\overline{B} = A$, in particolare se $\mathcal{B}$ è ortonormale allora $A = I$ e quindi la condizione diventa $B^t\overline{B} = I$, che equivale a $B\overline{B}^t = I$,cioè la matrice B è* unitaria

Dimostrazione. Sia $n = \dim V$.

1. Per l'osservazione che segue le definizioni T è hermitiano se e solo se, per ogni $i, j = 1, \ldots, n$ $\langle T(b_i), b_j\rangle = \langle b_i, T(b_j)\rangle$ ma

$$\langle T(b_i), b_j\rangle = \langle \sum_{k=1}^{n} b_{ki} b_k, b_j\rangle = \sum_{k=1}^{n} b_{ki}\langle b_k, b_j\rangle = \sum_{k=1}^{n} b_{ki} a_{kj} = (b^t a)_{ij}$$

mentre

$$\langle b_i, T(b_j)\rangle = \langle b_i, \sum_{k=1}^{n} b_{kj} b_k\rangle = \sum_{k=1}^{n} \overline{b}_{kj}\langle b_i, b_k\rangle = \sum_{k=1}^{n} a_{ik}\overline{b}_{kj} = (a\overline{b})_{ij}.$$

2. T è unitario se e solo se, per ogni $i, j = 1, \ldots, n$ $\langle T(b_i), T(b_j)\rangle = \langle b_i, b_j\rangle = a_{ij}$ ma

$$\langle T(b_i), T(b_j)\rangle = \langle \sum_{k=1}^{n} b_{ki} b_k, \sum_{s=1}^{n} b_{sj} b_s\rangle = \sum_{k,s=1}^{n} b_{ki}\langle b_k, b_s\rangle \overline{b}_{sj} =$$

$$\sum_{k,s=1}^{n} b_{ki} a_{ks} \overline{b}_{sj} = (b^t a \overline{b})_{ij}.$$

E la dimostrazione è conclusa □

Proposizione 31. *Sia T un operatore su uno spazio euclideo $(V, \langle\ ,\ \rangle)$, $\mathcal{B}$ una base di V, $A = \mathcal{M}_{\mathcal{B}}(\langle\ ,\ \rangle)$, $B = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$, allora*

1. *T è simmetrico se e solo se $B^t A = AB$, in particolare se $\mathcal{B}$ è ortonormale allora $A = I$ e quindi la condizione diventa $B^t = B$, cioè la matrice B è simmetrica.*

2. *T è ortogonale se e solo se $B^t AB = A$, in particolare se $\mathcal{B}$ è ortonormale allora $A = I$ e quindi la condizione diventa $B^t B = I$, cioè la matrice B è* ortogonale

Dimostrazione. Un po' più facile di quella appena conclusa. □

Se T è un operatore hermitiano su V e $\lambda \in \mathbb{C}$ è un suo autovalore, allora esiste $0 \neq v \in V$ tale che $T(v) = \lambda v$ ma allora

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

Siccome $\langle v, v\rangle \neq 0$ segue che $\lambda \in \mathbb{R}$.

Lemma 1. *Ogni operatore simmetrico su uno spazio euclideo ha un autovalore.*

Dimostrazione. Sia V lo spazio euclideo in questione e poniamo $n = \dim V$. Siano T un operatore simmetrico su V, $\mathcal{B}$ una base ortonormale di V e $B = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(T)$; per il punto 1 della proposizione 31, $A = A^t$.

Equipaggiamo $\mathbb{C}^n$ con il prodotto hermitiano standard e sia L l'unico operatore su $\mathbb{C}^n$ la cui matrice rispetto alla base standard è A. Allora $A = A^t = \overline{A}^t$ e, grazie al caso particolare del punto 1 della proposizione 30, L è un operatore hermitiano.

Per l'osservazione appena esposta i suoi autovalori, che certamente esistono (in $\mathbb{C}$) per il teorema fondamentale dell'algebra, sono in realtà numeri reali. Essi sono anche gli autovalori di T in quanto T e L hanno lo stesso polinomio caratteristico, cioè $\det(tI - A)$. □

Teorema 16 (spettrale). *Sia V uno spazio euclideo e T un operatore su V, le seguenti condizioni sono equivalenti*

1. *T è un operatore simmetrico.*

2. *Esiste una base ortonormale di V formata da autovettori di T.*

Dimostrazione. $2 \Rightarrow 1$ Se $\mathcal{B}$ è come in 2 allora la matrice di T rispetto a tale base è diagonale, e quindi simmetrica. Visto che $\mathcal{B}$ è ortonormale, il caso particolare del punto 1 della proposizione 31 dice che T è simmetrico.

$1 \Rightarrow 2$ si dimostra per induzione su $n = \dim V$. Se $n = 1$ ogni vettore di norma 1 è una base ortonormale di autovettori; si può obbiettare di non aver usato l'ipotesi, ma questo è dovuto semplicemente al fatto che in dimensione uno ogni operatore è simmetrico.

Supponiamo dunque che $n > 1$ e che il teorema sia vero per gli spazi vettoriali di dimensione $n - 1$. Per il lemma 1 esistono $\lambda \in \mathbb{R}$ e $0 \neq v \in V$ tali che $T(v) = \lambda v$; dividendo v per la sua norma si trova un autovettore unitario $b_1 \in V_\lambda$. Poniamo $W = \mathcal{L}(b_1)^\perp$. $= \{v \in V \text{ tali che } \langle v, b_i \rangle = 0\}$ Per la proposizione 29 $\dim W = n - 1$, inoltre, se $v \in W$ allora

$$\langle T(v), b_i \rangle = \langle v, T(b_i) \rangle = \langle v, \lambda b_i \rangle = \lambda \langle v, b_i \rangle = 0$$

e quindi T, ristretto a W è un operatore, evidentemente simmetrico, sullo spazio euclideo W equipaggiato con la restrizione del prodotto scalare di V.

Per l'ipotesi induttiva esiste una base ortonormale $(b_2, \ldots, b_n)$ di W formata da autovettori di $T_{|W}$ e quindi di T. Allora la base $(b_1, \ldots, b_n)$ di V soddisfa tutte le richieste. □

5.5 Classificazione degli operatori unitari in dimensione 2 e 3.

Sia V uno spazio euclideo di dimensione 2 e $\mathcal{B} = (b_1, b_2)$ una sua base ortonormale. Sia T un operatore ortogonale su V, il che ammonta a dire che $\|T(b_1)\| = \|T(b_2)\| = 1$ e che $\langle T(b_1), T(b_2) \rangle = 0$, cioè che $T(b_2) \in \mathcal{L}(T(b_1))^\perp$;

se $T(b_1) = xb_1 + yb_2$ allora $\|T(b_1)\| = 1 \Rightarrow x^2 + y^2 = 1$ e, visto che $T(b_1))^\perp$ ha dimensione 1 per la proposizione 29 e $0 \neq -yb_1 + xb_2$ ci appartiene, esiste $k \in \mathbb{R}$ tale che $T(b_2) = -kyb_1 + kxb_2$; infine $1 = \|T(b_2)\| = k^2(x^2 + y^2) = k^2$ da cui segue che $k = \pm 1$.

In sintesi la matrice di T rispetto a $\mathcal{B}$ può essere di uno dei seguenti tipi

$$A = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \quad \text{oppure} \quad B = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}$$

con $x^2 + y^2 = 1$; notiamo che $\det(A) = 1$ mentre $\det(B) = -1$.

Nel primo caso esiste (tutt'altro che unico) un numero reale θ tale che $\cos\theta = x$ e $\sin\theta = y$; allora la matrice A diventa $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ e l'operatore T viene chiamato *rotazione di θ radianti*; non ha autovalori salvo nei due casi $T = \mathrm{id}_V$ e $T = -\mathrm{id}_V$, cioè $\theta = k\pi$, $\forall k \in \mathbb{Z}$.

Nel secondo caso il polinomio caratteristico di T è $P_T(t) = t^2 - 1$ e quindi ha esattamente i due autovalori 1 e -1 e quindi due autospazi V_1 e V_{-1} che hanno entrambi dimensione 1 ed è diagonalizzabile; se $V_1 = \mathcal{L}(v)$ e $V_{-1} = \mathcal{L}(w)$ allora

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

e quindi i due vettori sono ortogonali; in altri termini $V_{-1} = V_1^\perp$.

Ogni vettore in V può essere scritto in modo unico come $av + bw$ e allora $T(av + bw) = av - bw$ e T di chiama *riflessione rispetto alla retta* V_1.

Supponiamo ora che $\dim V = 3$; se λ è un autovalore di T allora esiste $0 \neq v \in V$ tale che $T(v) = \lambda v$, ma allora $0 \neq \langle v, v\rangle = \langle T(v), T(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda^2\langle v, v\rangle$ da cui segue che $\lambda = \pm 1$, ed in particolare che T è un isomorfismo. Dal punto 2 della proposizione 31 segue che $\det(A) = \pm 1$, dove A è la matrice di T rispetto a una qualunque base. Il polinomio caratteristico di T è della forma $P_T(t) = t^3 + \alpha t^2 + \beta t - \det(A)$ dove $\alpha, \beta \in \mathbb{R}$ ed il termine noto, essendo uguale a $P_T(0)$ è proprio $-\det(A)$.

Se $\det(A) = -1$ siccome $\lim_{t\to-\infty} P_T(t) = -\infty$, per il teorema (di analisi) dell'esistenza degli zeri, T ha un autovalore negativo, che deve essere -1; analogamente se $\det(A) = 1$ allora T ha 1 come autovalore; in sintesi T ha sempre un autovalore $\lambda = \pm 1$ tale che $\lambda\det(A) = 1$.

Sia b_1 un autovettore unitario di tale autovalore e sia $W = \mathcal{L}(b_1)^\perp$, che ha dimensione 2. Se $w \in W$ allora $\langle T(w), \lambda b_1\rangle = \langle w, b_1\rangle = 0$, cioè la restrizione

di T a W è un operatore, evidentemente ortogonale, su W, che ha dimensione 2, quindi, se (b_2, b_3) è una base ortonormale di W la matrice di T rispetto alla base ortonormale (b_1, b_2, b_3) di V è di una delle due forme

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & x & -y \\ 0 & y & x \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \lambda & 0 & 0 \\ 0 & x & y \\ 0 & y & -x \end{pmatrix}$$

con $x^2 + y^2 = 1$; in realtà il secondo caso non si può presentare perché il determinante di tale matrice non è concorde con λ.

Quindi esiste $\theta \in \mathbb{R}$ tale che la matrice di T rispetto a (b_1, b_2, b_3) è

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Se $\lambda = 1$ si dice che T è una *rotazione* di θ radianti attorno all'asse $\mathcal{L}(b_1)$.

Se $\lambda = -1$ si dice che T è una *rotazione* di θ radianti attorno all'asse $\mathcal{L}(b_1)$ seguita (o preceduta) dalla *riflessione* rispetto al piano W.

Indice analitico

Indice