Appunti per gli studenti del corso di Geometria
Corso di Laurea in Fisica
A.A. 2025/26

A. Lotta

1. Nozioni di base

La geometria analitica permette di trattare questioni geometriche utilizzando il linguaggio algebrico. Così $\mathbb{R}$ fornisce un modello universale di "retta euclidea", mentre $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ il modello categorico di "piano euclideo". In ultima analisi, ciò che tale identificazione permette è la manipolazione e lo studio degli oggetti geometrici mediante le operazioni fondamentali tra numeri reali: somma, prodotto, differenza, passaggio al reciproco. In tale procedimento sono di basilare importanza le proprietà formali di queste operazioni. Su di esse si basa la nozione generale di *campo*, che sarà discussa in questo paragrafo.

Consideriamo un insieme X su cui sia definita un'**operazione interna** $*$, ovvero una legge che associa univocamente ad ogni coppia ordinata (x, y) di elementi di X, un ben preciso elemento dello stesso insieme, denotato con $x * y$. In tal caso diremo che $(X, *)$ è una **struttura algebrica.**

A priori, quindi, l'ordine in cui si considerano gli oggetti x e y è importante ai fini di determinare il risultato dell'operazione $*$ applicata ad essi. In altri termini, gli oggetti $x * y$ e $y * x$ possono differire.

Se $*$ è un'operazione su X, la coppia $(X, *)$ prende il nome di **struttura algebrica**.

Osservazione 1.1. Osserviamo che un'operazione interna $*$ sull'insieme X altro non è che una funzione (applicazione) $* : X \times X \to X$, il cui dominio è il prodotto cartesiano $X \times X = \{(x, y) : x \in X,\ y \in Y\}$ e il cui insieme di arrivo è lo stesso insieme X.

Più in generale, si possono considerare strutture algebriche $(X, *_1, \ldots, *_k)$ dove $*_1, \ldots, *_k$ sono $k \geq 1$ operazioni diverse su X.

Esempio 1.2. Sono operazioni interne la somma $+$ ed il prodotto $\cdot$ standard sugli insiemi numerici $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Quindi, ad esempio, sono strutture algebriche $(\mathbb{Z}, +)$, $(\mathbb{Q}, \cdot)$, ma anche $(\mathbb{Z}, +, \cdot)$ oppure $(\mathbb{R}, +, \cdot)$ (strutture con due operazioni).

Fissiamo ora un po' di terminologia di uso corrente. Si dice che l'operazione $*$ sull'insieme X gode della proprietà **commutativa** se:

$$\forall x, y \in X \quad x * y = y * x.$$

L'operazione $*$ è detta **associativa** se per ogni $x, y, z \in X$ risulta:

$$x * (y * z) = (x * y) * z.$$

In questo caso, è possibile scrivere, senza ambiguità:

$$x * y * z.$$

Diremo che struttura algebrica $(X, *)$ ammette **elemento neutro**, se esiste un elemento $e \in X$ tale che:

$$x * e = x = e * x$$

per ogni $x \in X$.

Un tale oggetto, se esiste, è sempre univocamente determinato. Infatti, ammettendo che e ed e' siano entrambi elementi neutri, risulta necessariamente:

$$e = e * e' = e'.$$

La prima uguaglianza è giustificata dal fatto che e' è elemento neutro, la seconda dal fatto che e gode della stessa proprietà.

I più familiari esempi di elementi neutri sono il numero 0 nel caso degli insiemi numerici $(\mathbb{N}, +)$, $(\mathbb{Q}, +)$, ecc. ed il numero 1, laddove si consideri la moltiplicazione invece della somma.

Esempio 1.3. Dato un insieme A, possiamo considerare sull'insieme delle parti (cioè l'insieme di tutti i sottoinsiemi) $\mathcal{P}(A)$ di A le operazioni fondamentali di intersezione e unione che danno luogo a due differenti strutture algebriche: $(\mathcal{P}(A), \cap)$ e $(\mathcal{P}(A), \cup)$ e in entrambi i casi sussiste la proprietà associativa. Si noti che la prima di queste strutture ammette l'intero insieme A come elemento neutro, mentre per la seconda tale ruolo è svolto dall'insieme vuoto $\emptyset$.

Esempio 1.4. Dato un insieme A, consideriamo l'insieme A^A di tutte le funzioni (applicazioni) da A in A. Questo insieme è dotato della fondamentale operazione di composizione

$$(f, g) \mapsto f \circ g.$$

Ricordiamo che $f \circ g : A \to A$ è la funzione definita ponendo

$$(f \circ g)(x) := f(g(x))$$

per ogni $x \in A$.

L'operazione di composizione $\circ$ tra funzioni è *sempre associativa*, ma *non* è (di norma) commutativa.

La funzione identica $Id_A : A \to A$ (definita com'è noto da $Id_A(x) := x$ per ogni $x \in A$), è l'elemento neutro di $(A^A, \circ)$, avendosi infatti, per ogni funzione $f : A \to A$, che

$$f \circ Id_A = f = Id_A \circ f.$$

In presenza di un elemento neutro e per una struttura algebrica $(X, *)$, un elemento $x \in X$ è detto **invertibile** se in corrispondenza di esso esiste un altro $x' \in X$, detto *inverso di* x, per cui si abbia:

$$x * x' = e = x' * x.$$

Si dice anche che x **ammette un inverso** in X.

Considerato ad esempio l'insieme dei numeri reali, ogni numero $x \neq 0$ è un elemento invertibile di $(\mathbb{R}, \cdot)$, con inverso $\frac{1}{x}$. Invece ogni numero reale x è un elemento invertibile di $(\mathbb{R}, +)$, con inverso $-x$.

Esempio 1.5. Consideriamo ancora un insieme non vuoto A e la struttura algebrica $(A^A, \circ)$. Data una funzione $f : A \to A$, essa è un elemento invertibile di $(A^A, \circ)$ se e solo se esiste una funzione $g : A \to A$ tale che:

$$f \circ g = Id_A, \; g \circ f = Id_A.$$

Questo signfica che per ogni $x \in A$ deve aversi

$$g(f(x)) = x$$

e anche:

$$f(g(y)) = y,$$

sempre per ogni $y \in A$.

Allo studente è già noto che condizione necessaria e sufficiente affinchè ciò sia possibile è che f sia una funzione *bigettiva*, ovvero sia ingettiva che surgettiva. La funzione $g : A \to A$ è la funzione inversa $g = f^{-1}$ della bigezione f, che associa ad ogni $y \in A$ l'unico elemento x dello stesso insieme, tale che $f(x) = y$.

In tal caso quindi abbiamo a disposizione una caratterizzazione completa degli elementi invertibili della struttura algebrica $(A^A, \circ)$ in esame.

Teorema 1.6. *Sia $(X, *)$ un struttura algebrica, dotata di elemento neutro e, la cui operazione $*$ è associativa. Se $x \in X$ è un elemento invertibile, allora l'inverso x' di x è univocamente determinato.*

Dimostrazione: Supponiamo che x' e x'' siano entrambi inversi di x. Allora in particolare abbiamo

$$x * x' = e.$$

Applicando l'operazione $*$, "componendo" ambo i membri a sinistra con x'' si ottiene:

$$x'' * (x * x') = x'' * e$$

ovvero

$$x'' * (x * x') = x''$$

che si può riscrivere, stante la proprietà associativa:

$$(x'' * x) * x' = x''.$$

D'altra parte, per definizione di inverso, sappiamo anche che $x'' * x = e$ e quindi l'ultima uguaglianza si riscrive:

$$e * x' = x''$$

da cui

$$x' = x''.$$

□

Introduciamo ora una delle più importanti tipologie di struttura algebrica, rilevante non solo nell'ambito della matematica pura, ma anche in molte altre discipline.

Definizione 1.7. Si chiama **gruppo** ogni struttura algebrica $(G, *)$ soddisfacente le seguenti condizioni:

1) L'operazione $*$ è associativa.
2) $(G, *)$ ammette elemento neutro.
3) Tutti gli elementi di G sono invertibili.

Solitamente l'elemento neutro di un gruppo G si denota con e, mentre l'inverso di un elemento x si denota con x^{-1}, salvo diversa convenzione.

Definizione 1.8. Un gruppo $(G, *)$ si dice **abeliano** se l'operazione $*$ è commutativa.

Sono gruppi:

- $(X, +)$ dove X è uno degli insiemi numerici $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.
- $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$ e dove $*$ indica l'insieme numerico in questione privato di 0.

  Tutti questi gruppi sono abeliani.

  Invece *non* sono gruppi:
- $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$ e $(\mathbb{Z}, \cdot)$.

  Si osservi che l'operazione di moltiplicazione è effettivamente un'operazione interna su $\mathbb{R}^*$: il prodotto $a \cdot b$ di due numeri reali diversi da zero, è anch'esso diverso da zero (legge di annullamento del prodotto). Pertanto possiamo considerare l'operazione $\cdot : \mathbb{R}^* \times \mathbb{R}^* \to \mathbb{R}^*$ e la struttura algebrica $(\mathbb{R}^*, \cdot)$ su tale insieme. Lo stesso discorso vale per $\mathbb{Q}^*$.

Esempio 1.9. Dato un insieme qualsiasi non vuoto A, l'insieme $Sym(A)$ di tutte le *bigezioni* $f : A \to A$ è un gruppo rispetto all'operazione $\circ$ di composizione. Si tratta di un esempio significativo di gruppo, in cui l'operazione interna *non* gode della familiare proprietà commutativa, tipica dei gruppi costruiti con gli insiemi numerici. Esso prende il nome di *gruppo delle permutazioni* di A. Ad esempio, se $A = \{1, \ldots, n\}$ è un insieme finito, si tratta del gruppo delle *permutazioni su n oggetti*, un importante gruppo finito studiato ad es. nell'ambito del calcolo combinatorio.

Con riferimento all'esempio precedente, si osservi che, in analogia con quanto osservato per il gruppo $(\mathbb{R}^*, \cdot)$, l'operazione $\circ$ si può considerare come operazione interna su $Sym(A)$, grazie al fatto che la composizione $f \circ g$ di due bigezioni è ancora una bigezione (la composizione di due applicazioni ingettive è ingettiva e la composizione di due funzioni surgettive è pure surgettiva). Tale operazione risulta indotta dall'operazione di composizione della struttura algebrica più grande $(A^A, \circ)$. La definizione seguente generalizza questo tipo di situazione.

Definizione 1.10. Data una struttura algebrica $(X, *)$, un sottoinsieme $E \subset X$ è detto **chiuso** o **stabile** (rispetto all'operazione $*$) se per ogni $x, y \in E$, si ha che

$$x * y \in E.$$

Si osservi che ogni sottoinsieme stabile E di $(X, *)$ eredita in modo naturale per restrizione da X l'operazione interna $*$, che opera allo stesso modo, ma solo su coppie di elementi di E, producendo sempre elementi di E. Evidentemente, se $*$ è associativa (risp. commutativa), tale è questa operazione indotta.

Dunque l'insieme $Sym(A)$ di tutte le bigezioni $F : A \to A$ di un insieme su sè stesso è stabile per l'operazione di composizione $\circ$ su A^A, ovvero è un sottoinsieme stabile della struttura algebrica $(A^A, \circ)$.

Come altro esempio, abbiamo che $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$ è un insieme stabile di $(\mathbb{R}, +)$, mentre $\{1, 0, -1\}$ non lo è.

Esempio 1.11. 1) L'insieme costituito da tutte e sole le rotazioni del piano Euclideo E intorno ad un punto fissato O è un gruppo, rispetto all'operazione di composizione di applicazioni.

2) Data una retta r del piano euclideo E, detta $\sigma : E \to E$ la simmetria assiale rispetto a r, allora $G = \{Id_E, \sigma\}$ costituisce un gruppo, costituito da soli due elementi. Ricordiamo che per ogni punto P, la sua immagine $\sigma(P)$ è il punto simmetrico di P rispetto a r.

Basta notare che σ è una bigezione che ammette se stessa come inversa: infatti abbiamo

$$\sigma \circ \sigma = Id_A$$

perchè denotando, per ogni punto P, con P' il simmetrico di esso rispetto alla retta, allora $(P')' = P$ (il simmetrico del simmetrico è il punto stesso).

Si osservi che un altro esempio di gruppo costituito da due soli elementi è $(\{0,1\},+)$, la cui operazione è definita ponendo:

$$0+0 := 0, \quad 0+1 := 1 = 1+0, \quad 1+1 := 0.$$

Tale gruppo si denota con $\mathbb{Z}_2$ (gruppo additivo dell'aritmetica modulo 2). Il ruolo che svolge la funzione identica nell'esempio precedente $G = \{Id_E, \sigma\}$ è qui svolto da 0, mentre la simmetria assiale è rimpiazzata da 1. Evidentemente i due gruppi sono identificabili (il termine preciso è isomorfi). Anzi, qualsiasi gruppo $(G,*)$ costituito da esattamente due elementi distinti ha lo stesso tipo di struttura: detto x l'elemento diverso dall'elemento neutro, x deve necessariamente ammettere se stesso come inverso.

3) Si chiama *gruppo banale* ogni struttura algebrica del tipo $(G = \{e\}, *)$, costituita da un solo elemento, la cui operazione è definita semplicmente ponendo

$$e * e = e.$$

Notazione e convenzioni riguardanti i gruppi:

Dato un gruppo $(G,+)$ la cui operazione è denotata con notazione **additiva** $+$, si usano di solito le seguenti **convenzioni**:

-L'elemento neutro si denota preferibilmente con 0.

-L'inverso di $a \in G$ si denota con $-a$ e si chiama **l'opposto di** a.

Inoltre, dati due oggetti qualsiasi x e y, in tal caso si scriverà:

$$x - y$$

in luogo di

$$x + (-y).$$

Dato un gruppo $(G,\cdot)$ la cui operazione è denotata mediate notazione **moltiplicativa** $\cdot$, si usano le seguenti convenzioni:

-L'elemento neutro si denota preferibilmente con 1 (talvolta è utilizzato anche il simbolo e).

-L'inverso di $a \in G$ si denota con a^{-1}.

Osservazione 1.12. Osserviamo che, dato un gruppo abeliano $(G, +)$ le convenzioni precedenti sulla notazione additiva permettono di effettuare operazioni nel gruppo in modo naturale, operando con le usuali regole del calcolo letterale. Ad es. se a, b, c sono elementi qualsiasi, allora si ha

$$a + b = c$$

se e solo se

$$a - c = -b.$$

Enunciamo ora una proprietà basilare dei gruppi, nota come *legge di cancellazione*.

Teorema 1.13. *Siano a, b, c elementi di un gruppo $(G, *)$ e si assuma che*

$$a * c = b * c$$

oppure che

$$c * a = c * b.$$

Allora $a = b$.

Dimostrazione: Denotato con e l'elemento neutro di G, assumendo ad esempio che

$$a * c = b * c,$$

allora componendo ambo i membri con l'inverso c^{-1} di c segue:

$$(a * c) * c^{-1} = (b * c) * c^{-1}$$

da cui

$$a * (c * c^{-1}) = b * (c * c^{-1})$$

ovvero

$$a * e = b * e$$

e quindi in definitiva

$$a = b.$$

□

Proposizione 1.14. *Dato un gruppo $(G, *)$, per ogni a, b in G l'inverso di $a * b$ è dato da:*

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

Dimostrazione: Tenendo presente che l'inverso di $a * b$ è univocamente determinato, basta fare la seguente verifica:

$$(a * b) * (b^{-1} * a^{-1}) = e = (b^{-1} * a^{-1}) * (a * b).$$

Infatti, sfruttando la proprietà associativa, abbiamo:

$$(a * b) * (b^{-1} * a^{-1}) = a * (b * b^{-1}) * a^{-1} = (a * e) * a^{-1} = a * a^{-1} = e.$$

In modo completamente analogo si controlla la validità della seconda uguaglianza richiesta. $\square$

Esempio 1.15. Discutiamo ora un esempio importante di gruppo abeliano, che incontreremo e generalizzeremo presto nel contesto degli spazi vettoriali; sul prodotto cartesiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ si definisce un'operazione naturale di somma $+$ ponendo:

$$(x, y) + (x', y') := (x + x', y + y'),$$

per ogni (x, y) e (x', y') in $\mathbb{R}^2$.

Risulta che $(\mathbb{R}^2, +)$ è un gruppo abeliano. L'elemento neutro è la coppia $(0, 0)$; per ogni $A = (x, y)$, esiste l'elemento inverso dato dalla coppia $(-x, -y)$. La verifica dell'associatività dell'operazione $+$ è semplice, ed è basata sulla proprietà associativa dell'addizione in $\mathbb{R}$; si ha infatti:

$$[(x, y) + (x', y')] + (x'', y'') = (x + x', y + y') + (x'', y'') = (x + x' + x'', y + y' + y'')$$

e d'altra parte

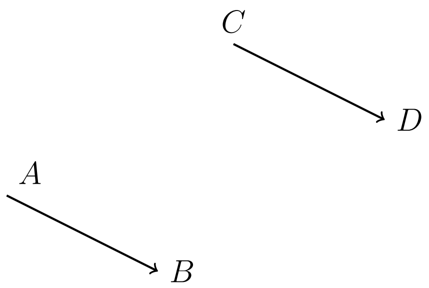
$$(x, y) + [(x', y') + (x'', y'')] = (x, y) + (x' + x'', y' + y'') = (x + x' + x'', y + y' + y'').$$

È altresì immediato constatare che $+$ gode della proprietà commutativa.

A questa operazione si può dare un significato geometrico, se interpretiamo ogni coppia (x, y) in $\mathbb{R}^2$ come un punto di un piano Cartesiano, $P(x, y)$ di ascissa x e ordinata y in un fissato sistema di riferimento.

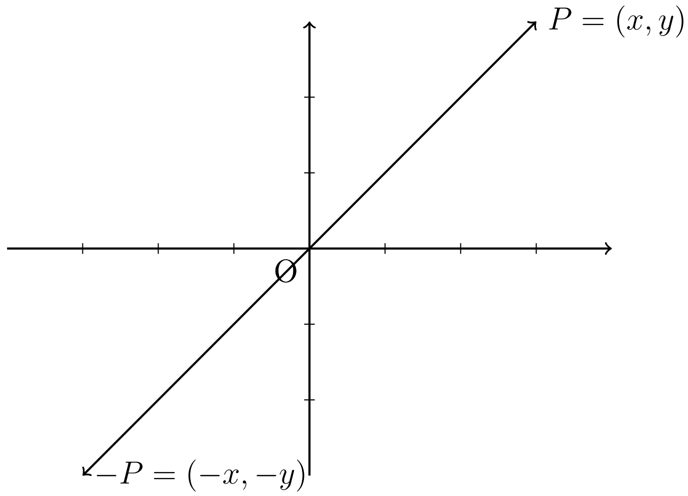
Ricordiamo che nel piano della geometria elementare vi è una nozione puramente sintetica di vettore: dati due punti distinti A e B, il *vettore applicato* $\overrightarrow{AB}$ è il segmento orientato AB. Nel caso generale $A \neq B$ esso ha tre caratteristiche: la *direzione*, che è quella individuata dalla retta per A e B, il verso, che è quello sulla stessa retta secondo cui A precede B, e la *lunghezza*, che è la lunghezza del segmento stesso (fissata una unità di misura) che è un numero strettamente positivo. Se invece $A = B$ tale vettore è detto *vettore nullo* ed ha lunghezza 0, mentre non gli si attribuisce nè direzione, nè verso. In ogni caso il punto A è il *punto di applicazione* del vettore $\overrightarrow{AB}$, mentre il punto B si dice punto di arrivo o punto finale dello stesso.

Dato un vettore $\overrightarrow{AB}$ e un punto C, esiste un unico vettore $\overrightarrow{CD}$ che ha lo stesse caratteristiche di $\overrightarrow{AB}$ (direzione, lunghezza e verso): esso si dice *equipollente ad* $\overrightarrow{AB}$. I vettori equipollenti $\overrightarrow{AB}$ e $\overrightarrow{CD}$ sono due rappresentanti dello stesso vettore *libero*.



La somma $\overrightarrow{AB} + \overrightarrow{AC}$ di due vettori $\overrightarrow{AB}$ e $\overrightarrow{AC}$ applicati nello stesso punto A, si definisce come il vettore $\overrightarrow{AD}$ dove D è il punto finale del vettore applicato in B equipollente ad $\overrightarrow{AC}$.
Questa regola prende il nome di *regola del parallelogramma* in quanto, nel caso generico in cui A, B e C non sono allineati, detto D il punto finale di $\overrightarrow{AB} + \overrightarrow{AC}$, allora A, C, D e B sono i vertici di un parallelogramma.

Tenendo presente ciò possiamo dare un significato geometrico all'operazione di somma in $\mathbb{R}^2$. Osserviamo innanzitutto che l'elemento neutro $(0,0)$ corrisponde all'origine O del nostro riferimento. Inoltre, l'opposto $-P$ di un punto coincide con il simmetrico di P rispetto a O.



Proposizione 1.16. *1) Siano A, B di $\mathbb{R}^2$ e sia C un altro punto. Allora l'unico punto D tale che*

$$\overrightarrow{AB} \text{ è equipollente a } \overrightarrow{CD}$$

è il punto

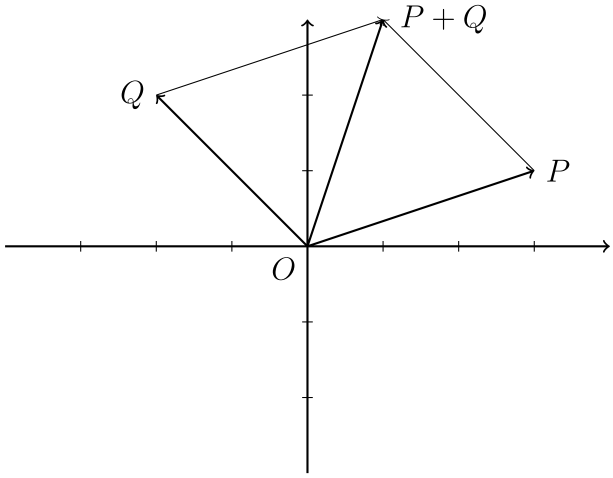
$$D = C + B - A.$$

2) Dati i punti A, B, C, D si ha:

$$\overrightarrow{AB} \text{ è equipollente a } \overrightarrow{CD} \iff B - A = D - C.$$

3) Dati i punti P e Q, la somma $P+Q$ coincide con il secondo estremo del vettore somma $\overrightarrow{OP}+\overrightarrow{OQ}$, effettuata secondo la regola del parallelogramma, ovvero:

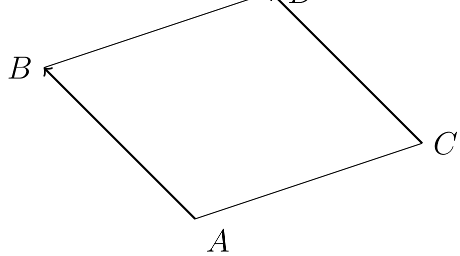
$$\overrightarrow{OP}+\overrightarrow{OQ}=\overrightarrow{O(P+Q)}.$$



Si osservi che in questa uguaglianza, la somma $P+Q$ è quella effettuata nel gruppo $(\mathbb{R}^2,+)$ che abbiamo introdotto, mentre la somma al primo membro è quella tra vettori applicati.

Notiamo anche che la proprietà 2) suggerisce che ogni differenza $B-A$ tra punti nel nostro gruppo si può interpretare come un vettore libero, di cui un rappresentante è il vettore applicato $\overrightarrow{AB}$.

Dimostrazione: Per ogni coppia di punti (A,B) distinti, denotiamo con $r_{A,B}$ la retta passante per questi punti. Nel provare 1), possiamo supporre che $A\neq B$, perchè nel caso $A=B$ il risultato è evidente, perchè in tal caso necessariamente $D=C$. Consideriamo dapprima il caso generico in cui C non è allineato con A e B. Dato il vettore $\overrightarrow{AB}$, sappiamo che il punto D in questione è univocamente determinato dalla condizione che A, B, D, C, siano i vertici di un parallelogramma:

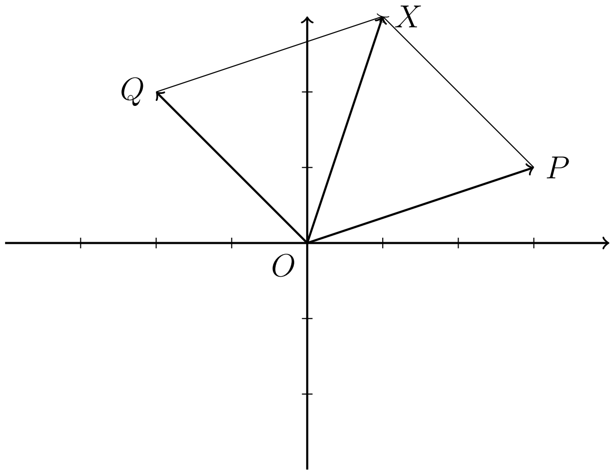


Affermiamo che tale punto è $X := C + B - A$. Per far ciò, posto $A(a_1, a_2)$, $B(b_1, b_2)$ e $C(c_1, c_2)$, basta controllare che la retta $r_{A,B}$ è parallela alla retta $r_{C,X}$ e che la retta $r_{B,X}$ è parallela alla retta $r_{A,C}$. Infatti, abbiamo che il punto X ha coordinate $(b_1 + c_1 - a_1, b_2 + c_2 - a_2)$ e un rapido confronto dei coefficienti angolari delle rette in questione ci dà conferma in merito. Ad esempio, tranne che nel caso in cui $a_1 = b_1$, il coefficiente angolare di $r_{A,B}$ è $m = \frac{b_2 - a_2}{b_1 - a_1}$; mentre quello della retta $r_{C,X}$ è $m' = \frac{b_2 + c_2 - a_2 - c_2}{b_1 + c_1 - a_1 - c_1} = m$. Nel caso particolare $a_1 = b_1$, la retta $r_{A,B}$ è comunque parallela a $r_{C,X}$ perchè entrambe sono parallele all'asse delle ordinate. La verifica per le rette $r_{B,X}$ e $r_{A,C}$ è analoga.

Nel caso in cui A, B e C siano allineati, la verifica che $D = X$ va fatta direttamente con la definizione di equipollenza: si tratta di controllare che il segmento AB sia congruente a CX e che C preceda X nel verso sulla retta $r_{A,B}$ secondo cui A precede B. Quest'ultima verifica consiste nel confrontare le ascisse dei punti (o, in caso siano uguali, le ordinate). Ad esempio, supponiamo che $a_1 < b_1$; bisogna controllare che anche l'ascissa di C sia minore dell'ascissa di X. Infatti, si ha: $c_1 < c_1 + (b_1 - a_1)$. La prima verifica si lascia al lettore, e si fa agevolmente calcolando le distanze $d(A, B)$ e $d(C, X)$. Dunque, in ogni caso resta provato che $D = C + B - A$.

La 2) consegue come immediata conseguenza: infatti, stante la 1) abbiamo che $\overrightarrow{AB}$ è equipollente a $\overrightarrow{CD}$ se e solo se e solo se $D = C + B - A$ e tale uguaglianza, operando nel gruppo $(\mathbb{R}^2, +)$, si può riscrivere come $B - A = D - C$.

La 3) è un caso particolare, in quanto effettuando la somma dei vettori $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ con la regola del parallelogramma si ottiene un vettore $\overrightarrow{OX}$ dove X è tale che $\overrightarrow{OQ}$ e $\overrightarrow{PX}$ sono equipollenti:



Pertanto, applicando 1) il punto X in questione è $X = P + (Q - O) = P + Q$. □

Possiamo ora introdurre il concetto di campo, i cui esempi più noti sono il campo dei numeri razionali $\mathbb{Q}$, quello dei numeri reali $\mathbb{R}$ e il campo dei numeri complessi $\mathbb{C}$.

Definizione 1.17. Si chiama **campo** ogni struttura algebrica $(K, +, \cdot)$ le cui operazioni soddisfano i seguenti requisiti:

a) $(K, +)$ è un gruppo abeliano, il cui elemento neutro si denota con 0.

b) L'operazione di moltiplicazione $\cdot$ è associativa, commutativa e $(K, \cdot)$ ammette l'elemento neutro, denotato con 1, tale che $1 \neq 0$.

c) Ogni $x \in K$, tale che $x \neq 0$ ammette un inverso rispetto al prodotto, denotato con x^{-1}.

d) Sussiste la *proprietà distributiva* del prodotto rispetto alla somma:

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

per ogni $x, y, z \in K$.

Proposizione 1.18. *Sia* $(K, +, \cdot)$ *un campo. Sussistono le seguenti proprietà:*

1) Per ogni $a \in K$ *si ha:* $0 \cdot x = 0 = x \cdot 0$.

2) Vale la legge di annullamento del prodotto*:*

$$a \cdot b = 0 \ \Rightarrow a = 0 \text{ oppure } b = 0.$$

3) L'insieme $K^* = K - \{0\}$ *è stabile rispetto all'operazione di moltiplicazione, e* $(K^*, \cdot)$ *è un gruppo abeliano.*

Dimostrazione: 1) Abbiamo, utilizzando la prima delle proprietà distributiva (assioma d) della definizione di campo):

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0,$$

che porta a $x \cdot 0 = 0$ in forza della legge di cancellazione nel gruppo $(K, +)$.

2) Siano $a, b \in K$ tali che

$$a \cdot b = 0.$$

Vogliamo provare che $a = 0$ oppure $b = 0$. Se $a = 0$, non c'è nulla da provare; assumiamo quindi $a \neq 0$; in tal caso l'assioma c) garantisce che a ammette l'inverso a^{-1} rispetto alla moltiplicazione; utilizzando tale inverso si ottiene:

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0,$$

da cui, stante la 1)

$$a^{-1} \cdot (a \cdot b) = 0,$$

e quindi, in forza della proprietà associativa del prodotto:

$$(a^{-1} \cdot a) \cdot b = 0,$$

ovvero

$$1 \cdot b = 0$$

e resta provato che $b = 0$.

3) La legge di annullamento del prodotto si può reinterpretare esattamente come la circostanza che il prodotto di due oggetti a, b entrambi diversi da zero, è anch'esso diverso da zero:

$$a \neq 0 \text{ e } b \neq 0 \Rightarrow a \cdot b \neq 0.$$

Ciò significa proprio che K^* è stabile rispetto al prodotto. Ciò detto, in forza degli assiomi b) e c), è immediato riconoscere che $(K^*, \cdot)$ è un gruppo abeliano. Infatti, l'operazione $\cdot$ ristretta a K^* è associativa e commutativa, e ammette ancora 1 come elemento neutro: si noti infatti che $1 \in K^*$ in quanto l'assioma 2) garantisce che $1 \neq 0$. Infine, osserviamo che dato $x \in K^*$, allora anche il suo inverso x^{-1} previsto nell'assioma b) appartiene a K^*, perchè se fosse $x^{-1} = 0$, dalla relazione

$$x \cdot x^{-1} = 1$$

seguirebbe

$$x \cdot 0 = 1$$

ovvero, in forza di 1)

$$0 = 1$$

il che viola l'assioma b). Pertanto ogni $x \in K^*$ risulta invertibile $(K^*, \cdot)$ con lo stesso x^{-1} come inverso, e con ciò resta verificato che questa struttura algebrica è un gruppo abeliano. $\square$

Coerentemente con le convenzioni generali sui gruppi, dato un campo $(K, +, \cdot)$ l'opposto di un elemento a, rispetto al gruppo $(K, +)$, si denoterà sempre con $-a$. Se $a \neq 0$, il suo inverso rispetto alla moltiplicazione, ovvero il suo inverso in $(K^*, \cdot)$, si denoterà sempre con a^{-1}. Utilizzando la Prop. 1.14, sappiamo che dati a, b elementi qualsiasi diversi da 0, si ha:

$$(a \cdot b)^{-1} = a^{-1} \cdot b^{-1} = b^{-1} \cdot a^{-1}.$$

Un'altra proprietà naturale che si deduce dagli assiomi di campo, e che lega tra loro l'addizione e la moltiplicazione, è la seguente.

Proposizione 1.19. *Se $(K, +, \cdot)$ è un campo, allora per ogni $x \in K$ risulta:*

$$-x = (-1) \cdot x.$$

Dimostrazione: Stante l'unicità dell'opposto di x nel gruppo $(K, +)$, e la commutatività di $+$, basta verificare che

$$(-1) \cdot x + x = 0.$$

Ma ciò è conseguenza della proprietà distributiva e della proposizione precedente:

$$(-1) \cdot x + x = (-1) \cdot x + 1 \cdot x = (-1 + 1) \cdot x = 0 \cdot x = 0.$$

$\square$

Osservazione 1.20. Nei campi numerici, dato un numero $a \neq 0$, siamo abituati a scrivere
$$\frac{1}{a}$$
per denotare l'inverso a^{-1} di a. Questa notazione è ammissibile anche in un campo astratto K. Inoltre, dati $a, b \in K$ con $b \neq 0$ si può anche scrivere $\frac{a}{b}$ in luogo di $a \cdot b^{-1}$, ovvero in luogo di $a \cdot \frac{1}{b}$. Con questa notazione, dati a, b, c, d con $c \neq 0$ e $d \neq 0$, vale:
$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d},$$
dove il secondo membro ha senso in forza della legge di annullamento del prodotto. Si lascia la verifica come esercizio.

2. spazi vettoriali

L'algebra lineare è la disciplina matematica il cui oggetto principale di studio è la teoria degli *spazi vettoriali* e delle *applicazioni lineari*. Gli spazi vettoriali sono oggetti di natura algebrica su cui non solo si può fondare l'ordinaria geometria elementare e le sue generalizzazioni, ma sono strutture basilari utilizzate in tutti i campi della Matematica, della Fisica e numerose altre discipline. Uno dei temi più importanti dell'algebra lineare è lo studio dei sistemi di equazioni lineari.

Sia $\mathbb{K}$ un campo. Denoteremo, in accordo con quanto discusso nel paragrafo precedente, con 0 e 1 gli elementi neutri di $\mathbb{K}$ rispettivamente per la somma ed il prodotto, che denotiamo sempre con $+$ e $\cdot$. Se $x \in \mathbb{K}$, allora $-x$ denoterà l'opposto di x (rispetto alla somma) mentre se $x \in \mathbb{K}^*$, usiamo x^{-1} per denotare l'inverso di x (rispetto al prodotto).

In precedenza abbiamo introdotto il concetto di operazione interna su un insieme. Introduciamo ora un altro tipo di operazione, che coinvolge il campo $\mathbb{K}$.

Definizione 2.1. Sia V un insieme non vuoto. Chiameremo *prodotto esterno* o *moltiplicazione per gli scalari di* $\mathbb{K}$ ogni applicazione del tipo
$$\cdot : \mathbb{K} \times V \to V.$$
Per ogni coppia (λ, v) con $\lambda \in \mathbb{K}$ e $v \in V$, scriveremo
$$\lambda \cdot v$$
oppure più semplicemente, omettendo il simbolo $\cdot$:
$$\lambda v$$
in luogo di $\cdot((\lambda, v))$.

Nel seguito quindi ammetteremo strutture algebriche su insiemi che presentino operazioni di tipo misto, interne e/o esterne.

Definizione 2.2. Si chiama *spazio vettoriale sul campo* $\mathbb{K}$ ogni struttura algebrica $(V, +, \cdot)$, dove
$$+ : V \times V \to V, \qquad \cdot : \mathbb{K} \times V \to V$$
sono rispettivamente un'operazione interna ed una esterna, dette operazioni di *somma* e *prodotto per scalari*, verificanti le seguenti proprietà (assiomi di spazio vettoriale):

A) $(V, +)$ è un gruppo abeliano.

B) Il prodotto per scalari soddisfa, per ogni $\lambda, \mu \in \mathbb{K}$ e $v \in V$:
   1) $\lambda(\mu v) = (\lambda\mu)v$
   2) $1v = v$.

C) Le due operazioni sono legate dalle seguenti relazioni, valide per ogni $\lambda, \mu \in \mathbb{K}$ e per ogni $v, w \in V$:
   3) $(\lambda + \mu)v = \lambda v + \mu v$
   4) $\lambda(v + w) = \lambda v + \lambda w$.

Notiamo esplicitamente che negli assiomi di spazio vettoriale sono coinvolte, riconoscibili senza ambiguità, anche le operazioni di somma e prodotto del campo $\mathbb{K}$. In questo contesto gli elementi di $\mathbb{K}$ vengono di solito chiamati *scalari*.

Spesso uno spazio vettoriale $(V, +, \cdot)$ sarà denotate semplicemente con il simbolo V, sottintendendo le due operazioni di cui è dotato. Diremo anche V è un $\mathbb{K}$-spazio vettoriale. Gli elementi di V si diranno *vettori* di V.

Nel caso in cui $\mathbb{K} = \mathbb{R}$, parleremo di spazio vettoriale *reale*, mentre se $\mathbb{K} = \mathbb{C}$, la terminologia è spazio vettoriale *complesso*.

Se $\lambda \in \mathbb{K}$ e $v \in V$, il vettore λv è detto *prodotto di v per lo scalare λ*. L'elemento neutro di V rispetto all'operazione di somma verrà denotato con 0_V o con 0 e sarà chiamato il *vettore nullo* di V.

Ricordiamo inoltre che l'assioma A) garantisce la validità delle proprietà commutativa e associativa dell'operazione di somma; inoltre per ogni vettore $v \in V$ esiste un unico vettore $-v \in V$ tale che
$$v + (-v) = 0_V;$$
tale vettore sarà chiamato il *vettore opposto* di v.

Esempio 2.3. Ogni campo $\mathbb{K}$ è dotato di una struttura canonica di spazio vettoriale su se stesso; le operazioni sono esattamente quelle di somma e prodotto del campo; quest'ultima infatti si può interpretare, in questo caso, anche come operazione esterna.

Gli assiomi di campo garantiscono infatti la validità delle proprietà A) e B) e C).

Spesso tale spazio si chiama *spazio vettoriale 1-dimensionale canonico* su $\mathbb{K}$ o anche *retta* sul campo $\mathbb{K}$.

Esempio 2.4. Sappiamo che il piano cartesiano reale $\mathbb{R}^2$ ha una struttura canonica di gruppo abeliano, discussa nell'Esempio 1.15 Si definisce anche in modo naturale una moltplicazione esterna $\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ ponendo

$$\lambda \cdot (x, y) := (\lambda x, \lambda y)$$

per ogni $\lambda \in \mathbb{R}$ e per ogni $(x, y) \in \mathbb{R}^2$. Risulta quindi che $(\mathbb{R}^2, +, \cdot)$ è uno spazio vettoriale reale.

Per interpretare geometricamente le operazioni qui introdotte, è sufficiente rappresentare ogni (x, y) mediante il vettore applicato $\vec{OP}$, dove P è il punto di coordinate (x, y). Dunque ogni elemento di $\mathbb{R}^2$ ha due interpretazioni: punto o vettore. Quando usiamo la seconda interpretazione, stiamo dando una veste geometrica ai vettori dello *spazio vettoriale* $\mathbb{R}^2$. Come già visto in precedenza, la somma di due vettori corrisponde alla somma dei corrispondenti vettori applicati secondo la regola del parallelogramma. Dato poi un vettore $v = (x, y) \in \mathbb{R}^2$ rappresentato da $\vec{OP}$ e uno scalare $\lambda \in \mathbb{R}$ non nullo, allora $\lambda \cdot v$ corrisponde al vettore geometrico $\overrightarrow{OQ}$, dove il punto Q è allineato con O e P, il rapporto delle misure dei segmenti OQ e OP è λ, e il verso di $\overrightarrow{OQ}$ è concorde con quello di $\overrightarrow{OP}$ se $\lambda > 0$, mentre è il verso opposto a quello di $\overrightarrow{OP}$ se $\lambda < 0$.

Esempio 2.5. L'esempio precedente ammette una naturale generalizzazione. Dato un campo qualsiasi $\mathbb{K}$, possiamo introdurre una struttura di $\mathbb{K}$-spazio vettoriale sul prodotto cartesiano $\mathbb{K}^2 := \mathbb{K} \times \mathbb{K}$ introducendo due operazioni

$$+ : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}^2, \quad \cdot : \mathbb{K} \times \mathbb{K}^2 \to \mathbb{K}^2,$$

definite allo stesso modo di $\mathbb{R}^2$ come segue:

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2), \qquad \lambda(x, y) := (\lambda x, \lambda y).$$

La verifica della validità degli assiomi di spazio vettoriale per $(\mathbb{K}^2, +, \cdot)$ non presenta particolari difficoltà. In particolare, l'associatività e la commutatività di $+$ vengono ereditate dalle analoghe proprietà della somma nel campo $\mathbb{K}$ (che viene utilizzata contemporaneamente e separamente sulle due coordinate delle coppie di scalari coinvolte).

Risulta che $(\mathbb{K}^2, +)$ è un gruppo, in quanto ammette la coppia:

$$(0, 0)$$

come elemento neutro, e ogni $(x, y) \in \mathbb{K}^2$ ammette come opposto:

$$(-x, -y).$$

Dunque $(\mathbb{K}^2, +)$ è effettivamente un gruppo abeliano. Si lascia al lettore la verifica completa delle restanti proprietà B) e C) nella definizione di spazio vettoriale. A titolo di esempio, controlliamo la prima delle B): se $\lambda, \mu \in \mathbb{K}$ e $(x, y) \in \mathbb{K}$, allora per definizione della moltiplicazione esterna risulta:

$$\lambda(\mu(x, y)) = \lambda(\mu x, \mu y) = (\lambda(\mu x), \lambda \mu y) = (\lambda \mu x, \lambda \mu y),$$

e d'altra parte:

$$(\lambda\mu)(x,y) = ((\lambda\mu)x, (\lambda\mu)y) = (\lambda\mu x, \lambda\mu y).$$

In tale verifica ci siamo avvalsi della proprietà associativa del prodotto nel campo.

Esempio 2.6. Generalizzando quanto fatto nell'esempio precedente, possiamo introdurre una struttura di $\mathbb{K}$-spazio vettoriale sul prodotto cartesiano di più copie del campo $\mathbb{K}$, ovvero sull'insieme

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ volte}} = \{(x_1, \dots, x_n) \,|\, x_i \in \mathbb{K} \text{ per ogni } i\}.$$

Le operazioni di tale spazio vettoriale sono definite quindi come segue:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n),$$

$$\lambda(x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Il vettore nullo è:

$$(0, 0, \dots, 0),$$

mentre per ogni vettore $(x_1, \dots, x_n) \in \mathbb{K}^n$ il suo opposto è:

$$(-x_1, \dots, -x_n).$$

In simboli:

$$-(x_1, \dots, x_n) = (-x_1, \dots, -x_n).$$

Lo spazio vettoriale $\mathbb{K}^n$ sarà chiamato *spazio vettoriale n-dimensionale canonico* (talvolta *numerico*) sul campo $\mathbb{K}$.

3. Prodotti diretti

Siano V_1 e V_2 spazi vettoriali sullo stesso campo $\mathbb{K}$. Definiamo sul prodotto cartesiano $V_1 \times V_2$ due operazioni

$$+ : (V_1 \times V_2) \times (V_1 \times V_2) \to V_1 \times V_2, \qquad \cdot : \mathbb{K} \times (V_1 \times V_2) \to V_1 \times V_2$$

nel modo seguente

$$(v_1, v_2) + (w_1, w_2) := (v_1 + w_1, v_2 + w_2), \ \lambda(v_1, v_2) := (\lambda v_1, \lambda v_2). \tag{1}$$

È facile verificare che tali operazioni godono delle proprietà A), B) e C) della Definizione 2.2. In particolare, riguardo A), l'elemento neutro per la somma è la coppia $(0_{V_1}, 0_{V_2})$; per ogni vettore $(v_1, v_2) \in V_1 \times V_2$, il suo opposto è la coppia $(-v_1, -v_2)$.

Definizione 3.1. Lo spazio vettoriale $V_1 \times V_2$ dotato delle operazioni (1) si chiama *prodotto diretto* di V_1 e V_2.

4. Spazi vettoriali di funzioni

Introduciamo ora un'altra importante tipologia di spazi vettoriali. Siano X un insieme non vuoto e V uno spazio vettoriale sul campo $\mathbb{K}$. Denoteremo con

$$V^X$$

l'insieme di tutte le funzioni $F : X \to V$. Su tale insieme si definiscono in modo naturale un'operazione interna di somma e una moltiplicazione esterna per scalari, come segue. Date due funzioni $F : X \to V$ e $G : X \to V$, allora la *funzione somma* $F + G$ si definisce come la funzione denotata con $F + G : X \to V$ che opera come segue:

$$(F + G)(x) := F(x) + G(x) \quad \forall x \in X.$$

Invece se $F : X \to V$ e $\lambda \in \mathbb{K}$, allora il *prodotto* $\lambda \cdot F$ si definisce come la funzione $\lambda \cdot F : X \to V$ tale che:

$$(\lambda \cdot F)(x) := \lambda F(x) \quad \forall x \in X.$$

Risulta allora che $(V^X, +, \cdot)$ è uno spazio vettoriale sul campo $\mathbb{K}$. Si osserva che le proprietà di associatività della somma tra funzioni discende dall'analoga proprietà della somma tra vettori di V. Lo stesso vale per quel che riguarda la proprietà commutativa.

Il vettore nullo di tale spazio è la funzione costante di valore 0_V, che denotiamo con 0_{V^X}. Dunque si tratta della funzione definita da:

$$0_{V^X}(x) := 0_V \quad \forall x \in X,$$

che associa ad ogni x il vettore nullo di V.

Infatti, risulta che

$$F + 0_{V^X} = F$$

in quanto per ogni $x \in X$, calcolando il valore in x della funzione somma $F + 0_{V^X}$ si ottiene:

$$(F + 0_{V^X})(x) = F(x) + 0_{V^X}(x) = F(x) + 0_V = F(x).$$

Per ogni $F : X \to V$ il vettore opposto $-F$ è la funzione così definita:

$$(-F)(x) := -F(x) \quad \forall x \in X,$$

che associa cioè ad ogni punto x in X il vettore opposto di $F(x)$ in V. Anche in questo caso la verifica di ciò consiste nel giustificare un'uguaglianza tra funzioni:

$$F + (-F) = 0_{V^X}.$$

Questa uguaglianza sussiste in quanto per ogni $x \in X$ si ha:

$$(F + (-F))(x) = F(x) + (-F)(x) = F(x) - F(x) = 0_V.$$

La verifica della validità di tutti gli altri assiomi di spazio vettoriale per tale struttura algebrica $(V^X, +, \cdot)$ è un esercizio molto istruttivo per prendere confidenza con la nozione di spazio vettoriale.

5. Matrici

Discutiamo ora un altro esempio fondamentale di spazio vettoriale. Sia $\mathbb{K}$ un campo e siano m, n interi positivi.

Definizione 5.1. Una **matrice** A di tipo $m \times n$ (oppure (m, n)) a coefficienti in $\mathbb{K}$ è una tabella rettangolare costituita da $m \cdot n$ elementi di $\mathbb{K}$ disposti in m righe ed n colonne, e rappresentata nel modo seguente:

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix} \tag{2}$$

Si scrive anche in modo compatto $A = (a_j^i)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ o più semplicemente $A = (a_j^i)$. Nella scrittura $A = (a_j^i)$ converremo sempre che l'indice in alto (in basso) indichi la riga (risp. la colonna) occupata dall'elemento a_j^i. L'insieme di tutte le matrici di tipo $m \times n$ ad elementi in $\mathbb{K}$ verrà denotato col simbolo $M_{m,n}(\mathbb{K})$.

Se $m = n$, tale insieme si denota più semplicemente con $M_n(\mathbb{K})$ e i suoi elementi sono detti *matrici quadrate* di ordine n.

Una matrice $(a_1 \ \ldots \ a_n)$ di tipo $1 \times n$ è chiamata *vettore riga* (di lunghezza n) ed è identificata con il vettore $(a_1, \ldots, a_n)$ dello spazio $\mathbb{K}^n$. Analogamente, ogni matrice $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$ è detta *vettore colonna* ed è identificabile anch'essa con $(a_1, \ldots, a_m) \in \mathbb{K}^m$. Nel seguito tali identificazioni saranno utilizzate spesso senza ulteriori commenti.

Fissato $i \in \{1, \ldots, m\}$, l'i-ma riga di A è dunque la matrice di tipo $1 \times n$

$$A^{(i)} := \begin{pmatrix} a_1^i & a_2^i & \cdots & a_n^i \end{pmatrix}$$

mentre, fissato $j \in \{1, \ldots, m\}$, la j-ma colonna di A è la matrice di tipo $m \times 1$:

$$A_{(j)} := \begin{pmatrix} a_j^1 \\ a_j^2 \\ \vdots \\ a_j^m \end{pmatrix}.$$

Possiamo rappresentare A per colonne nel modo seguente:

$$A = \begin{pmatrix} A_{(1)} & A_{(2)} & \cdots & A_{(n)} \end{pmatrix}$$

oppure per righe:

$$A = \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(m)} \end{pmatrix}.$$

Teorema 5.2. *L'insieme $M_{m,n}(\mathbb{K})$ ammette una struttura canonica di spazio vettoriale su $\mathbb{K}$, le cui operazioni di somma e prodotto esterno sono definite come segue: per ogni $A = (a^i_j)$, $B = (b^i_j)$ e $\lambda \in \mathbb{K}$ si pone:*

$$A + B := (a^i_j + b^i_j), \quad \lambda A := (\lambda a^i_j).$$

Lasciamo al lettore la cura di verificare, per le operazioni descritte sopra, la validità degli assiomi A), B) e C) della definizione di spazio vettoriale. Ci limitiamo a mettere in evidenza due fatti salienti: il vettore nullo dello spazio vettoriale $M_{m,n}(\mathbb{K})$ in questione è la *matrice nulla* ovvero la matrice, denotata semplicemente con 0, i cui elementi sono tutti uguali allo zero di $\mathbb{K}$:

$$0 := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Per ogni $A = (a^i_j) \in M_{m,n}(\mathbb{K})$, il vettore opposto di A in $M_{m,n}(\mathbb{K})$ è la matrice

$$-A = (-a^i_j).$$

D'ora in avanti considereremo sempre l'insieme $M_{m,n}(\mathbb{K})$ munito della struttura di spazio vettoriale appena descritta.

6. Proprietà elementari degli spazi vettoriali

Prima di procedere oltre, mettiamo in evidenza alcune proprietà basilari degli spazi vettoriali che si deducono dagli assiomi.

Proposizione 6.1. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Allora per ogni $v \in V$ e per ogni $\lambda \in \mathbb{K}$, risulta:*

a) $\lambda 0_V = 0_V$
b) $0v = 0_V$
c) $-v = (-1)v$.
d) $\lambda v = 0_V \Rightarrow \lambda = 0 \vee v = 0_V$.

Dimostrazione: a) Abbiamo $\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V$ da cui, per la legge di cancellazione: $0_V = \lambda 0_V$.
b) Possiamo scrivere $0v = (0 + 0)v = 0v + 0v$ da cui ancora per la legge di cancellazione $0_V = 0v$.

c) Per l'unicità dell'opposto di v nel gruppo abeliano $(V, +)$, si tratta di provare che

$$v + (-1)v = 0_V.$$

Infatti,

$$v + (-1)v = 1v + (-1)v = (1 - 1)v = 0v = 0_V$$

dove l'ultimo passaggio è giustificato dalla proprietà b) già provata.

d) Si supponga $\lambda v = 0_V$. Se $\lambda = 0$, la tesi è provata. Supponiamo quindi $\lambda \neq 0$; allora esiste l'inverso λ^{-1} di λ in $\mathbb{K}$. Dalla relazione $\lambda v = 0_V$ moltiplicando ambo i membri per λ^{-1} otteniamo

$$\lambda^{-1}(\lambda v) = \lambda^{-1} 0_V$$

da cui per la a):

$$(\lambda^{-1}\lambda)v = 0_V$$

che possiamo riscrivere

$$1v = 0_V$$

ovvero $v = 0_V$.

□

7. Sottospazi

Sia V uno spazio vettoriale sul campo $\mathbb{K}$.

Definizione 7.1. Un sottoinsieme non vuoto W di V si dice *sottospazio vettoriale* di V se è stabile rispetto ad entrambe le operazioni $+$ e $\cdot$ di V, il che significa che sono soddisfatte le seguenti condizioni:

a) $\forall u, v \in W : \quad u + v \in W.$

b) $\forall u \in W \, \forall \lambda \in \mathbb{K} : \quad \lambda u \in W.$

La terminologia "sottospazio" è chiarita dalla seguente proposizione.

Proposizione 7.2. *Sia $W \subset V$ un sottospazio vettoriale del $\mathbb{K}$-spazio vettoriale V. Allora si ha che $0_V \in W$. Inoltre W eredita in modo canonico una struttura di spazio vettoriale su $\mathbb{K}$, le cui operazioni sono ottenute dalle operazioni dello spazio vettoriale ambiente V per restrizione.*

Dimostrazione: Sia $u \in W$; allora, utilizzando la proprietà b) di sottospazio, abbiamo che anche il vettore opposto $-u$ appartiene anch'esso a W, in quanto, ricordando quanto stabilito nella Prop. 6.1, possiamo scrivere $-u = (-1)u \in W$. In particolare, fissato un tale elemento $u \in W$ (è possibile perchè W non è vuoto), usando la proprietà a) si evince che $0_V \in W$; infatti:

$$0_V = u + (-u) \in W.$$

Con ciò resta provata la prima affermazione. Le condizioni a) e b) permettono di restringere le operazioni di somma e prodotto esterno su W, ottenendo le operazioni indotte

$$+ : W \times W \to W, \ \cdot : \mathbb{K} \times W \to W$$

con cui si opera allo stesso modo. Come sappiamo, l'operazione di somma così ottenuta continua a essere associativa e commutativa sull'insieme stabile W. Evidentemente 0_V è anche elemento neutro di $(W, +)$ e ogni $u \in W$ risulta invertibile in $(W, +)$ grazie al fatto, già stabilito, che $-u \in W$. Dunque $(W, +)$ è un gruppo abeliano. Infine, le proprietà B) e C) nella definizione di spazio vettoriale sono ancora manifestamente soddisfatte anche dalle operazioni qui coinvolte, ristrette a W. $\square$

Esempio 7.3. Pensando a $\mathbb{R}^2$ come ad un piano cartesiano, abbiamo che ogni retta $r \subset \mathbb{R}^2$ passante per l'origine è un sottospazio vettoriale di $\mathbb{R}^2$.

Infatti, supponiamo che r abbia equazione:

$$r : ax + by = 0,$$

dove le costanti reali a, b non sono entrambe zero. Dunque r è il luogo geometrico:

$$r = \{(x, y) \in \mathbb{R}^2 : \ ax + by = 0\}.$$

Dati $(x_1, y_1) \in r$ e $(x_2, y_2) \in r$, bisogna verificare innazitutto che

$$(x_1, y_1) + (x_2, y_2) \in r.$$

Infatti, per definizione $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ e risulta

$$a(x_1 + x_2) + b(y_1 + y_2) = (ax_1 + by_1) + (ax_2 + by_2) = 0 + 0 = 0.$$

Analogamente, se $\lambda \in \mathbb{R}$ e $(x_1, y_1) \in r$, allora anche $\lambda(x_1, y_1) = (\lambda x_1, \lambda y_1)$ appartiene alla retta perchè

$$a(\lambda x_1) + b(\lambda y_1) = \lambda(ax_1 + by_1) = \lambda \cdot 0 = 0.$$

Questo esempio ci fa intuire che la stessa nozione di retta, così come quella di piano, e tutta la geometria elementare si possono fondare, in modo indipendente dal metodo assiomatico classico, sulla nozione di spazio vettoriale reale.

Esempio 7.4. Una matrice quadrata $A = (a^i_j)$ di ordine n a coefficienti in un campo assegnato $\mathbb{K}$ si dice *simmetrica* se:

$$a^i_j = a^j_i, \text{ per ogni } i, j \in \{1, \dots, n\}.$$

Ciò significa che gli elementi della matrice sono disposti in modo simmetrico rispetto alla cosiddetta *diagonale principale*, formata dagli ingressi del tipo $a^1_1, \dots, a^n_n$, aventi indice riga e colonna uguali.

Così ad es. la generica matrice simmetria 3x3 è del tipo:

$$A = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_2^1 & a_2^2 & a_3^2 \\ a_3^1 & a_3^2 & a_3^3 \end{pmatrix}.$$

Non è difficile verificare che l'insieme $Sym_n(\mathbb{K})$ di tutte le matrici simmetriche di un fissato ordine n è un sottospazio vettoriale di $M_n(\mathbb{K})$. Infatti, tale insieme è non vuoto in quanto la matrice nulla 0 è banalmente simmetrica. Date due matrici simmetriche $A = (a_j^i)$ e $B = (b_j^i)$, allora $A + B$ è simmetrica in quanto

$$a_j^i + b_j^i = a_i^j + b_i^j.$$

Analogamente, per ogni $\lambda \in \mathbb{K}$ e $A = (a_j^i) \in Sym_n(\mathbb{K})$, risulta che anche λA è simmetrica avendosi

$$\lambda a_j^i = \lambda a_i^j.$$

In particolare, si ottiene anche $Sym_n(\mathbb{K})$ è canonicamente uno spazio vettoriale su $\mathbb{K}$.

8. Somma e intersezione di sottospazi

Discutiamo ora due metodi standard per costruire sottospazi vettoriali di uno spazio vettoriale assegnato.

Proposizione 8.1. *Siano U e W due sottospazi di un assegnato $\mathbb{K}$-spazio vettoriale V.*

a) Il sottoinsieme di V definito da

$$U + W := \{u + w |\, u \in U,\, w \in W\}$$

è un sottospazio vettoriale di V, detto somma *di U e W.*

b) L'intersezione insiemistica $U \cap W$ è un sottospazio vettoriale di V.

Dimostrazione:

a) Naturalmente $U + W \neq \emptyset$ perchè entrambi i sottospazi non sono vuoti. Considerati due vettori $u + w$ e $u' + w'$ di $U + W$, dove $u, u \in U$ e $w, w' \in W$, allora:

$$(u + w) + (u' + w') = (u + u') + (w + w') \in U + W$$

perchè sia U che W sono stabili rispetto alla somma. Analogamente, dato uno scalare λ si ha

$$\lambda(u + w) = (\lambda u) + (\lambda w) \in U + W.$$

b) L'intersezione in questione è non vuota, in quanto il vettore nullo appartiene a entrambi i sottospazi. Dati due vettori u e w in $U \cap W$allora il vettore $u + w$ appartiene ad U, in quanto U è sottospazio, e i vettori in questione vi appartengono; per lo stesso motivo $u + v$ appartiene anche a W, dunque appartiene a $U \cap W$. Lo stesso argomento mostra che, se $\lambda \in \mathbb{K}$ e $u \in U \cap W$, anche λu appartiene a $U \cap W$.

□

Si osservi che risulta:

$$U \subset U + W, \quad W \subset U + W,$$

perchè ogni vettore u di U può essere scritto nella forma $u = u + 0_V$ e analogamente per W.

9. Combinazioni lineari di vettori

Definizione 9.1. Siano V un $\mathbb{K}$-spazio vettoriale e

$$v_1, \ldots, v_k$$

una sequenza finita di vettori di V. Si chiama *combinazione lineare* dei vettori $v_1, \ldots, v_k$ ogni vettore del tipo

$$\lambda_1 v_1 + \cdots + \lambda_k v_k$$

dove $\lambda_1, \ldots, \lambda_r$ sono arbitrari scalari, detti *coefficienti* della combinazione lineare.

Esempio 9.2. Ad esempio tutte le combinazioni lineari dei vettori

$$(1, 2, 0), (0, 0, 1), (\sqrt{2}, 0, \sqrt{3})$$

di $\mathbb{R}^3$ sono i vettori del tipo:

$$(\alpha + \sqrt{2}\gamma, 2\alpha, \beta + \sqrt{3}\gamma)$$

al variare di α, β, γ in $\mathbb{R}$.

Invece, date le funzioni $\sin : \mathbb{R} \to \mathbb{R}$, $\cos : \mathbb{R} \to \mathbb{R}$, pensate come vettori di $\mathbb{R}^{\mathbb{R}}$, ogni funzione del tipo:

$$f(x) = a \sin(x) + b \cos(x),$$

con a e b costanti reali, è una loro combinazione lineare.

Sussiste la seguente caratterizzazione dei sottospazi, che utilizza la nozione di combinazione lineare di due vettori:

Proposizione 9.3. *Sia W un sottoinsieme non vuoto di un $\mathbb{K}$-spazio vettoriale V. Allora sono proprietà equivalenti:*

a) W è sottospazio vettoriale di V;

b) Per ogni $u, v \in W$ e $\lambda, \mu \in \mathbb{K}$ si ha $\lambda u + \mu v \in W$.

Si lascia la verifica come esercizio.

10. Sottospazio generato da un numero finito di vettori

Definizione 10.1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $v_1, \dots, v_m$, $m \geq 1$, una successione finita di vettori.

Si chiama *sottospazio vettoriale di V generato da $v_1, \dots, v_m$* il sottoinsieme di V

$$L(v_1, \dots, v_m) := \{\lambda_1 v_1 + \cdots + \lambda_m v_m | \ \lambda_1, \dots, \lambda_m \in \mathbb{K}\}$$

costituito da tutte le combinazione lineari dei vettori $v_1, \dots, v_m$.

Notiamo che $L(v_1, \dots, v_m)$ è effettivamente un sottospazio di V. Chiaramente $L(v_1, \dots, v_m)$ non è vuoto; notiamo in particolare che $0 \in L(v_1, \dots, v_m)$ in quanto

$$0 = 0v_1 + \cdots + 0v_m.$$

Date due combinazioni lineari $u = \lambda_1 v_1 + \cdots + \lambda_m v_m$ e $w = \mu_1 v_1 + \cdots + \mu_m v_m$, la loro somma è ancora una combinazione lineare degli stessi vettori, infatti:

$$\begin{aligned}
&u + w = (\lambda_1 v_1 + \cdots + \lambda_m v_m) + (\mu_1 v_1 + \cdots + \mu_m v_m) \\
&(\lambda_1 v_1 + \mu_1 v_1) + \cdots + (\lambda_m v_m + \mu_m v_m) = \\
&(\lambda_1 + \mu_1) v_1 + \cdots + (\lambda_m + \mu_m) v_m.
\end{aligned}$$

Si noti l'utilizzo delle proprietà commutativa e associativa della somma e dell'assioma C) di spazio vettoriale. Analogamente, per ogni $\lambda \in \mathbb{K}$ abbiamo

$$\lambda u = \lambda(\lambda_1 v_1 + \cdots + \lambda_m v_m) = (\lambda\lambda_1) v_1 + \cdots + (\lambda\lambda_m) v_m \tag{3}$$

per cui anche λu appartiene a $L(v_1, \dots, v_m)$.

Notiamo esplicitamente che ciascun vettori v_i della sequenza appartiene a $L(v_1, \dots, v_m)$, potendosi scrivere

$$v_i = 0v_1 + \cdots + 0v_{i-1} + 1\, v_i + 0v_{i+1} + \cdots + 0v_m.$$

Sussiste anzi la seguente proprietà:

Proposizione 10.2. *Data una sequenza di vettori $v_1, \dots, v_m$ dello spazio vettoriale V ed un sottospazio vettoriale W di V contenente tutti i vettori $v_1, \dots, v_m$, si ha che*

$$L(v_1, \dots, v_m) \subset W.$$

Dunque il sottospazio generato da $v_1, \dots, v_m$ è il più piccolo sottospazio che contiene gli stessi vettori.

Dimostrazione: Basta osservare che ogni vettore del tipo $\lambda_1 v_1 + \cdots + \lambda_m v_m$ deve appartenere a W in quanto, stante l'ipotesi su W, per definizione di sottospazio vettoriale ogni $\lambda_i v_i$ vi appartiene e W è stabile rispetto all'operazione di somma. □

11. Applicazioni lineari

Nello studio degli spazi vettoriali è essenziale lavorare con funzioni che permettano di mettere in relazione spazi diversi non solo come insiemi, ma tenendo conto della struttura algebrica aggiuntiva.

Definizione 11.1. Siano V e V' due spazi vettoriali sullo stesso campo $\mathbb{K}$. Un' applicazione $F : V \to V'$ si dice $\mathbb{K}$-*lineare* o semplicemente *lineare* se ha le proprietà seguenti:

$$\begin{array}{lll} 1) & \forall u, v \in V : & F(u+v) = F(u) + F(v). \\ 2) & \forall v \in V \, \forall \lambda \in \mathbb{K} : & F(\lambda v) = \lambda F(v). \end{array}$$

Ogni applicazione lineare deve soddisfare il seguente vincolo:

Proposizione 11.2. *Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Allora*

$$F(0_V) = 0_{V'}.$$

Dimostrazione: Infatti,

$$F(0_V) = F(0_V + 0_V) = F(0_V) + F(0_V).$$

Utilizzando la legge di cancellazione nel gruppo $(V', +)$ segue che $0_{V'} = F(0_V)$. □

Esempio 11.3. Sono lineari tutte le funzioni $\mathbb{R} \to \mathbb{R}$ del tipo:

$$x \in \mathbb{R} \mapsto ax,$$

dove $a \in \mathbb{R}$ è una costante assegnata. Si noti che invece un polinomio di primo grado

$$x \in \mathbb{R} \mapsto ax + b$$

con $b \neq 0$ non è lineare secondo la definizione che abbiamo introdotto, stante la proposizione precedente.

Esempio 11.4. Ogni funzione polinomiale $F : \mathbb{R}^2 \to \mathbb{R}$ del tipo

$$F(x, y) := ax + by$$

con a, b costanti assegnate, è lineare.

La verifica di questo fatto ricalca quanto fatto nell'Es. 7.3 per giustificare che ogni retta per l'origine è un sottospazio vettoriale di $\mathbb{R}^2$. Infatti, dati i vettori (x_1, y_1) e (x_2, y_2) risulta:

$$F((x_1, y_1) + (x_2, y_2)) = F(x_1 + x_2, y_1 + y_2) = a(x_1 + x_2) + b(y_1 + y_2) =$$
$$= (ax_1 + by_1) + (ax_2 + by_2) = F(x_1, y_1) + F(x_2, y_2).$$

E ancora in modo simile:

$$F(\lambda(x_1, y_1)) = F(\lambda x_1, \lambda y_1) = a(\lambda x_1) + b(\lambda y_1) = \lambda(ax_1 + by_1) = \lambda F(x_1, y_1).$$

Osservazione 11.5. Notiamo che ogni applicazione lineare $F : V \to V'$ soddisfa anche:

$$F(u - v) = F(u) - F(v),$$

per ogni $u, v \in V$.

Infatti: $F(u - v) = F(u + (-1)v) = F(u) + (-1)F(v) = F(u) - F(v)$.

Ad ogni applicazione lineare si associa in modo naturale un insieme notevole di vettori, definito come segue:

Definizione 11.6. Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Si chiama *nucleo di F* il seguente sottoinsieme di V:

$$Ker(F) := \{v \in V \,|\, F(v) = 0_{V'}\}.$$

Il seguente risultato è basilare:

Teorema 11.7. *Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Allora:*

1) $Ker(F)$ è un sottospazio vettoriale di V.

2) F è ingettiva $\iff$ $Ker(F) = \{0_V\}$.

Dimostrazione: a) Sappiamo che $0_V \in Ker(F)$, per cui $Ker(F)$ è non vuoto. Siano $u, v \in \mathrm{Ker}(F)$; allora

$$F(u + v) = F(u) + F(v) = 0_{V'} + 0_{V'} = 0_{V'},$$

e ciò mostra che $u + v \in \mathrm{Ker}(F)$. Inoltre, dato un arbitrario scalare λ abbiamo:

$$F(\lambda u) = \lambda F(u) = \lambda 0_{V'} = 0_{V'}.$$

b) Cominciamo con l'osservare che l'inclusione $\{0_V\} \subset Ker(F)$ è sempre vera. Supponiamo dapprima che F sia ingettiva. Sia $u \in Ker(F)$; allora $F(u) = 0_{V'}$, ovvero stante la Prop. 11.2:

$$F(u) = F(0_V),$$

da cui $u = 0_V$ per l'ingettività di F. Resta così provato che $Ker(F) = \{0_V\}$.

Supponiamo ora $Ker(F) = \{0_V\}$; siano $u, v \in V$ tali che

$$F(u) = F(v).$$

Utlizzando le proprietà del gruppo $(V', +)$, segue:

$$F(u) - F(v) = 0_{V'}$$

e quindi

$$F(u - v) = 0_{V'},$$

pertanto $u - v \in Ker(F)$ e per l'ipotesi segue necessariamente che $u - v = 0_V$, ovvero $u = v$. $\square$

Tornando all'esempio delle rette per l'origine in $\mathbb{R}^2$, ogni retta del piano di assegnata equazione $r : ax + by = 0$ coincide con il nucleo dell'applicazione lineare $F : \mathbb{R}^2 \to \mathbb{R}$ di cui nell'esempio precedente. Quindi il fatto che ogni retta in questione è uno spazio vettoriale è fondato su un principio generale.

Si noti anche la seguente caratterizzazione delle applicazioni lineari, che coinvolge combinazioni lineari di vettori:

Proposizione 11.8. *Sia $F : V \to V'$ un'applicazione tra due $\mathbb{K}$-spazi vettoriali. Sono proprietà equivalenti:*

a) F è lineare;

b) Per ogni $\lambda, \mu \in \mathbb{K}$ e $u, v \in V$ si ha: $F(\lambda u + \mu v) = \lambda F(u) + \mu F(v)$.

La dimostrazione è un esercizio: per provare che b) implica a), basta particolarizzare b) nel caso $\lambda = \mu = 1$ e nel caso $\mu = 0$.

Esempio 11.9. (Proiezioni) Fissato $n \geq 1$, le n funzioni

$$p_i : \mathbb{K}^n \to \mathbb{K}, \quad p_i(x_1, \ldots, x_n) := x_i, \quad i \leq i \leq n$$

che associano ad ogni vettore di $\mathbb{K}^n$ la sua coordinata i-ma sono tutte lineari. Esse si chiamano *proiezioni naturali* o semplicemente *proiezioni*.

La verifica di ciò è semplice: dati $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ e due scalari λ, μ allora:

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1, \ldots, \lambda x_i + \mu y_i, \ldots, \lambda x_n + \mu y_n)$$

sicchè

$$p_i(\lambda x + \mu y) = \lambda x_i + \mu y_i = \lambda p_i(x) + \mu p_i(y).$$

12. Monomorfismi, epimorfismi e isomorfismi

Definizione 12.1. Si consideri un'applicazione lineare $F : V \to V'$ tra due $\mathbb{K}$-spazi vettoriali.

Se F è ingettiva, prende il nome di **monomorfismo**, mentre se è surgettiva prende il nome di **epimorfismo**; infine, se F è bigettiva, essa prende il nome di **isomorfismo**.

Sappiano che i monomorfismi sono caratterizzati dall'avere nucleo banale; gli epimorfismi si caratterizzano, stante la definizione, mediante un altro sottospazio notevole:

Definizione 12.2. Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Si chiama **immagine di** F il sottoinsieme $Im(F) := F(V)$ di V'. Pertanto:

$$Im(F) = \{F(v) | \, v \in V\}.$$

Dunque F è un epimorfismo se e solo se $Im(F) = V'$.

Proposizione 12.3. *Per ogni applicazione lineare $F : V \to V'$, $Im(F)$ è un sottospazio vettoriale di V'.*

La dimostrazione è semplice: dati due vettori $u' = F(u)$ e $v' = F(v)$ di $Im(F)$ e due scalari λ e μ risulta:

$$\lambda u' + \mu v' = \lambda F(u) + \mu F(v) = F(\lambda u + \mu v) \in Im(F).$$

Due $\mathbb{K}$-spazi vettoriali V e V' si diranno **isomorfi** se esiste almeno un isomorfismo $F : V \to V'$. In tal caso, scriveremo

$$V \cong V'$$

o anche, volendo mettere in evidenza un particolare isomorfismo:

$$F : V \xrightarrow{\sim} V'.$$

Due spazi vettoriali isomorfi sono da considerarsi due diverse manifestazioni del medesimo oggetto matematico; dal punto di vista dell'algebra lineare, essi avranno le stesse proprietà e si potranno opportunamente identificare.

Vale il seguente risultato, la cui dimostrazione è un utile esercizio:

Teorema 12.4. *Sia $F : V \to V'$ un isomorfismo. Allora anche l'applicazione inversa $F^{-1} : V' \to V$ è lineare, e quindi è anch'essa un isomorfismo.*

Concludiamo questo paragrafo con la seguente osservazione:

Proposizione 12.5. *Sia $F : V \to V'$ un'applicazione lineare tra due $\mathbb{K}$-spazi vettoriali. Allora F trasforma combinazioni lineari di vettori in combinazioni lineari delle rispettive immagini, mantenendo gli stessi coefficienti. Ovvero per ogni $v_1, \ldots, v_k \in V$ e per ogni $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ si ha:*

$$F(\lambda_1 v_1 + \cdots + \lambda_k v_k) = \lambda_1 F(v_1) + \cdots + \lambda_k F(v_k).$$

Per provarlo è sufficiente applicare più volte la b) della Proposizione 11.8. Una dimostrazione più rigorosa si può fare agevolmente per induzione sul numero k dei vettori coinvolti.

13. Ulteriori proprietà delle applicazioni lineari

Proposizione 13.1. *Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Sia U un sottospazio vettoriale di V. Allora:*

a) la restrizione $F|_U : U \to V'$ è lineare.

b) L'immagine $F(U)$ di U è sottospazio vettoriale di V'.

Dimostrazione: a) La dimostrazione è immediata, in quanto la restrizione di F opera come F sui vettori di U: quindi per ogni $u, w \in U$ e λ, μ scalari abbiamo ancora

$$F|_U(\lambda u + \mu w) = F(\lambda u + \mu w) = \lambda F(u) + \mu F(w) = \lambda F|_U(u) + \mu F|_U(w).$$

b) Ricordiamo che, per definizione:

$$F(U) = \{F(u) |\, u \in U\},$$

e chiaramente tale insieme coincide con $Im(F|_U)$, avendosi:

$$\{F(u) |\, u \in U\} = \{F|_U(u) |\, u \in U\}.$$

□

La seguente proprietà è fondamentale nella teoria degli spazi vettoriali.

Teorema 13.2. *Siano V, V' e V'' spazi vettoriali sullo stesso campo $\mathbb{K}$ e siano $F : V \to V'$, $G : V' \to V''$ applicazioni lineari. Allora anche $G \circ F : V \to V''$ è lineare.*

Dimostrazione: Siano $u, v \in V$ e λ, μ scalari. Allora si ha:

$$\begin{aligned}&(G \circ F)(\lambda u + \mu v) = G(F(\lambda u + \mu v)) = G(\lambda F(u) + \mu F(v))\\ &= \lambda G(F(u)) + \mu G(F(v)) = \lambda (G \circ F)(u) + \mu (G \circ F)(v).\end{aligned}$$

□

Occupiamoci ora di ulteriori importanti proprietà delle applicazioni lineari, che coinvolgono le operazioni di somma di funzioni e moltiplicazione di scalari per funzioni, già considerate nel §4.

Proposizione 13.3. *Siano V e W due $\mathbb{K}$-spazi vettoriali.*

a) Se $F : V \to W$ e $G : V \to W$ sono due applicazioni lineari, allora anche $F + G : V \to W$ è lineare.

b) Se $F : V \to W$ è un'applicazione lineare e $\lambda \in \mathbb{K}$, allora anche λF è lineare.

Dimostrazione: a) Date le applicazioni lineari $F : V \to W$ e $G : V \to W$, due vettori u, v di V e scalari λ, μ, si tratta di verificare che:

$$(F + G)(\lambda u + \mu v) = \lambda (F + G)(u) + \mu (F + G)(v).$$

Infatti, per definzione della funzione somma $F + G$, e la linearità delle due funzioni abbiamo:

$$\begin{aligned}&(F + G)(\lambda u + \mu v) = F(\lambda u + \mu v) + G(\lambda u + \mu v) =\\ &\lambda F(u) + \mu F(v) + \lambda G(u) + \mu G(v) =\\ &\lambda (F(u) + G(u)) + \mu (F(v) + G(v)) =\\ &\lambda (F + G)(u) + \mu (F + G)(v).\end{aligned}$$

La verifica di b) è simile:

$$(\lambda F)(\alpha u + \beta v) = \lambda(F(\alpha u + \beta v)) = \lambda(\alpha F(u) + \beta F(v)) =$$
$$\lambda\alpha F(u) + \lambda\beta F(v) = \alpha\lambda F(u) + \beta\lambda F(v) = \alpha(\lambda F)(u) + \beta(\lambda F)(v).$$

□

Fissati i due spazi vettoriali V e W sullo stesso campo $\mathbb{K}$ denoteremo con il simbolo

$$Hom(V, W)$$

l'insieme di tutte le applicazioni lineari da V in W, ovvero

$$Hom(V, W) := \{F : V \to W \mid F \text{ è lineare}\}.$$

Si tratta di un sottoinsieme dello spazio vettoriale W^V di tutte le funzioni $F : V \to W$: esso è certamente non vuoto, perchè l'applicazione nulla $0 : V \to W$ è lineare. Le proprietà a) e b) appena stabilite garantiscono quindi che:

Corollario 13.4. *Dati due $\mathbb{K}$-spazi vettoriali, $Hom(V, W)$ è un sottospazio vettoriale di W^V.*

In particolare anche $Hom(V, W)$ è canonicamente un $\mathbb{K}$-spazio vettoriale, in cui le operazioni di somma tra funzioni lineari e moltiplicazione per scalari sono esattamente quelle discusse sopra. Il vettore nullo di tale spazio è la funzione nulla $0 : V \to W$.

Concludiamo questo paragrafo discutendo le seguenti proprietà utili che coinvolgono somme e composizioni di applicazioni lineari:

Proposizione 13.5. *Siano $F : V \to V'$ e $G : V \to V'$ due applicazioni $\mathbb{K}$-lineari e $T : U \to V$, $L : V' \to W$ due altre applicazioni lineari, dove U e W sono altri $\mathbb{K}$-spazi vettoriali. Allora:*

$$(F + G) \circ T = F \circ T + G \circ T, \quad L \circ (F + G) = L \circ F + L \circ G.$$

Dimostrazione: La dimostrazione è ancora una semplice applicazione della definizione di somma di due funzioni; ad esempio, per ogni $u \in U$ si ha:

$$\begin{aligned} ((F + G) \circ T)(u) &= (F + G)(T(u)) = F(T(u)) + G(T(u)) = \\ &= (F \circ T)(u) + (G \circ T)(u) = (F \circ T + G \circ T)(u). \end{aligned}$$

Similmente, per ogni $v \in V$ ricaviamo che:

$$\begin{aligned} (L \circ (F + G))(v) &= L((F + G)(v)) = L(F(v) + G(v)) = L(F(v)) + L(G(v)) = \\ &= (L \circ F)(u) + (L \circ G)(v) = (L \circ F + L \circ G)(v). \end{aligned}$$

□

14. La base canonica di $\mathbb{K}^n$

Lo spazio vettoriale $\mathbb{K}^n$ contiene un insieme di vettori particolarmente importante; esso costituisce il prototipo della nozione di "base" di uno spazio vettoriale qualsiasi che verrà studiata in seguito.

Definizione 14.1. Sia $n \geq 1$. Si chiama *base canonica* di $\mathbb{K}^n$ la sequenza di n vettori

$$e_1, \ldots, e_n$$

così definiti:

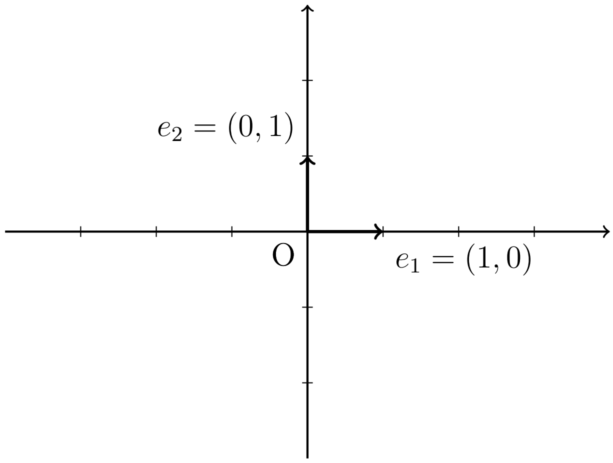
$$e_1 := (1, 0, \ldots, 0),\ e_2 := (0, 1, 0, \ldots, 0), \ldots, e_n := (0, 0, \ldots, 0, 1). \tag{4}$$

Dunque per ogni $j \in \{1, \ldots, n\}$, il vettore e_j ha tutte le coordinate nulle tranne la j-ma che coincide con 1.

Ad esempio, la base canonica di $\mathbb{K}^2$ è costituita dai vettori

$$e_1 := (1, 0),\ e_2 := (0, 1).$$

Nel caso $n = 1$, corrispondente al campo $\mathbb{K}$ pensato come spazio vettoriale, abbiamo l'unico vettore $e_1 = 1$.



Illustriamo ora una proprietà fondamentale della base canonica:

Teorema 14.2. *Ogni vettore $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ si può scrivere in modo unico come combinazione lineare dei vettori della base canonica, nel modo seguente:*

$$x = x_1 e_1 + \cdots + x_n e_n = \sum_{i=1}^{n} x_i e_i, \tag{5}$$

i coefficienti della combinazione lineare essendo le coordinate stesse del vettore.

Dimostrazione: La verifica di (5) è semplice: sviluppando il secondo membro abbiamo

$$\begin{aligned}
&x_1e_1 + \cdots + x_ne_n = \\
&x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \cdots + x_n(0, 0, \ldots, 0, 1) = \\
&(x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \cdots + (0, 0, \ldots, 0, x_n) = \\
&(x_1 + 0 + \cdots + 0, 0 + x_2 + 0 + \cdots + 0, \cdots, 0 + \cdots + 0 + x_n) = \\
&(x_1, x_2, \ldots, x_n).
\end{aligned}$$

Riguardo l'unicità della decomposizione, supponiamo che x si possa anche scrivere

$$x = \lambda_1e_1 + \cdots + \lambda_ne_n$$

mediante altri coefficienti; osserviamo che in base a quanto appena provato, il secondo membro di questa uguaglianza altri non è che il vettore $(\lambda_1, \ldots, \lambda_n)$, quindi abbiamo

$$(x_1, x_2, \ldots, x_n) = (\lambda_1, \ldots, \lambda_n)$$

e pertanto necessariamente $\lambda_i = x_i$ per ogni $i = 1, \ldots, n$. □

Come applicazione di questo risultato, determiniamo esplicitamente tutte le applicazioni lineari $\mathbb{K}^n \to \mathbb{K}$, dove $n \geq 1$ è un intero fissato.

Teorema 14.3. *Sia $n \geq 1$ un intero. Allora le applicazioni lineari $F : \mathbb{K}^n \to \mathbb{K}$ sono tutte e sole le funzioni del tipo*

$$F(x_1, \ldots, x_n) := a_1x_1 + \cdots + a_nx_n \tag{6}$$

L dove $a_1, \ldots, a_n$ sono delle costanti appartenenti al campo $\mathbb{K}$. Data inoltre una tale funzione, risulta $a_i = F(e_i)$.

Dimostrazione: La dimostrazione del fatto che ogni funzione F del tipo (6) è effettivamente un'applicazione lineare può farsi direttamente, oppure si può far uso delle proiezioni $p_i : \mathbb{K}^n \to \mathbb{K}$, $i = 1, \ldots, n$ e osservare che

$$F = a_1p_1 + \cdots + a_np_n.$$

Infatti per ogni si ha $a_ip_i(x) = a_ip_i(x) = a_ix_i$ e quindi

$$F(x) = a_1x_1 + \cdots a_nx_n = a_1p_1(x) + \cdots a_np_n(x) = (a_1p_1 + \cdots a_np_n)(x).$$

Siccome le p_i sono lineari possiamo concludere che anche F è lineare.

Proviamo ora il viceversa. Sia assegnata ora un'applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$. Utilizzando il fatto che

$$x = x_1e_1 + \cdots + x_ne_n$$

otteniamo, applicando la Proposizione 12.5:

$$F(x) = x_1F(e_1) + \cdots + x_nF(e_n)$$

da cui segue che F è della forma richiesta (6), ponendo $a_i := F(e_i)$ per ogni i.

Infine, data una qualsiasi funzione del tipo (6), è immediato dalla definizione che $F(e_i) = a_i$: infatti, l'unica coordinata non nulla di e_i è la i-ma e quindi

$$F(e_i) = a_1 \cdot 0 + \cdots + a_i \cdot 1 + \cdots + a_n \cdot 0 = a_i.$$

□

Vale la pena osservare esplicitamente che, stante il risultato appena provato, ogni funzione lineare $F : \mathbb{K}^n \to \mathbb{K}$ è completamente determinata dagli n scalari $a_i = F(e_i)$. La stessa cosa si può riformulare anche sotto forma di criterio di uguaglianza tra due funzioni lineari $F, G : \mathbb{K}^n \to \mathbb{K}$:

$$F = G \iff F(e_i) = G(e_i) \quad \forall i = 1, \ldots, n.$$

Consideriamo ora lo spazio $Hom(\mathbb{K}^n, \mathbb{K}^m)$, determinandone in modo esplicito tutti gli elementi:

Teorema 14.4. *Dati due interi $n, m \geq 1$ le applicazioni lineari $F : \mathbb{K}^n \to \mathbb{K}^m$ sono tutte e sole le funzioni del tipo*

$$F(x) := (F_1(x), \cdots, F_m(x)), \quad \forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n, \tag{7}$$

dove le $F_1, \ldots, F_m$ sono tutte funzioni lineari $F_i : \mathbb{K}^n \to \mathbb{K}$.

Così, ad esempio le seguenti sono funzioni lineari $F \in Hom(\mathbb{R}^3, \mathbb{R}^2)$, e $G \in Hom(\mathbb{R}^2, \mathbb{R}^3)$:

$$F(x, y, z) = (x + 3y + z, y - \sqrt{3}z), \quad G(x, y) = (0, x + \frac{1}{3}y, x - y).$$

Nel caso di questa F abbiamo $F_1(x, y, z) = x + 3y + z$ e $F_2(x, y, z) = y - \sqrt{3}z$, tali funzioni $\mathbb{R}^3 \to \mathbb{R}$ essendo lineari in base al Teorema 14.3.

Dimostrazione: Ogni funzione del tipo (7) è lineare, grazie alla linearità di tutte le F_i; infatti dati $x, y \in \mathbb{K}^n$ e scalari λ, μ risulta:

$$\begin{aligned}
&F(\lambda x + \mu y) = (F_1(\lambda x + \mu y), \ldots, F_m(\lambda x + \mu y)) = \\
&(\lambda F_1(x) + \mu F_1(y), \ldots, \lambda F_m(x) + \mu F_m(y)) = \\
&(\lambda F_1(x), \ldots, \lambda F_m(x)) + (\mu F_1(y), \ldots, \mu F_m(y)) = \\
&\lambda(F_1(x), \ldots, F_m(x)) + \mu(F_1(y), \ldots, F_m(y)) = \\
&\lambda F(x) + \mu F(y).
\end{aligned}$$

Viceversa, sia assegnata una qualunque applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}^m$; per ogni $x \in \mathbb{K}^n$, denotiamo con $F_1(x), \ldots, F_m(x)$ le m coordinate del vettore $F(x) \in \mathbb{K}^m$, che sono univocamente determinate. Quindi restano definite m funzioni scalari $F_1, \ldots, F_m : \mathbb{K}^m \to \mathbb{K}$: ciascuna F_i associa ad x la i-ma coordinata del vettore $F(x)$. Per costruzione, abbiamo pertanto:

$$F(x) = (F_1(x), \cdots, F_m(x))$$

per ogni $x \in \mathbb{K}^n$. Resta da controllare che tali funzioni F_i sono lineari. Basta notare che esse non sono altro che le funzioni composte:

$$F_i = p_i \circ F,$$

dove per ogni $i = 1, \ldots, m$, $p_i : \mathbb{K}^m \to \mathbb{K}$ è la i-ma proiezione

$$p_i(y_1, \ldots, y_m) := y_i.$$

Per concludere basta utilizzare il fatto che la composizione di applicazioni lineari è lineare. $\square$

Data una funzione $F : \mathbb{K}^n \to \mathbb{K}^m$ data dalla (7), scriveremo

$$F = (F_1, \ldots, F_m)$$

e chiameremo le F_i le *funzioni componenti* di F.

Notiamo ancora esplicitamente che una funzione lineare $F : \mathbb{K}^n \to \mathbb{K}^m$ è completamente determinata dai valori che assume sulla base canonica $e_1, \ldots, e_n$ di $\mathbb{K}^n$; in altri termini, se $G : \mathbb{K}^n \to \mathbb{K}^m$ è un'altra applicazione lineare, allora

$$F = G \iff F(e_i) = G(e_i) \quad i = 1, \ldots, n.$$

Questo segue immediatamente dall'analogo criterio di confronto per le funzioni coordinate F_i e G_i.

15. Il Prodotto righe per colonne tra matrici

Si definisce ora un'operazione fondamentale tra matrici, detta prodotto righe per colonne; come vedremo, essa è strettamente correlata con l'operazione di composizione di applicazioni lineari.

Consideriamo dapprima il caso più semplice che coinvolge un vettore riga e un vettore colonna.

Definizione 15.1. Siano $A = (a_1 \ \ldots \ a_n)$ un vettore riga e $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ un vettore colonna della stessa lunghezza $n \geq 1$, a coefficienti nello stesso campo $\mathbb{K}$. Si definisce il loro **prodotto** AB come lo scalare

$$AB := a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \in \mathbb{K}.$$

È chiaro che questa operazione di prodotto è direttamente connessa con la nozione di applicazione lineare. Consideriamo infatti una generica applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$. Sappiamo che una tale funzione è del tipo (6), ed è completamente determinata dagli n scalari $a_1, \ldots, a_n$. Possiamo rappresentare quindi tale funzione in modo sintetico mediante il vettore riga $(a_1 \ \cdots \ a_n)$, sapendo che, noto tale vettore, la funzione è ricostruita da

$$F(x) = a_1x_1 + \cdots + a_nx_n.$$

D'altra parte il secondo membro è semplicemente il prodotto tra $(a_1 \ \cdots \ a_n)$ e il vettore colonnna $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Questo fatto suggerisce di definire, per ogni vettore riga $A = (a_1 \ \cdots \ a_n)$ la seguente applicazione lineare

$$L_A = L_{(a_1 \ \cdots \ a_n)} : \mathbb{K}^n \to \mathbb{K}$$

definita ponendo, per ogni $x = (x_1, \ldots, x_n)$:

$$L_A(x) := (a_1 \ \cdots \ a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Così facendo ogni applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$ è rappresentabile con questa notazione, cioè $F = L_{(a_1 \ \cdots \ a_n)}$ dove il vettore riga $(a_1 \cdots a_n)$ corrispondente è univocamente determinato e $a_i = F(e_i) = L_{(a_1 \ \cdots \ a_n)}(e_i)$.

La definizione di prodotto si estende al caso di due matrici di tipo generale nel modo seguente:

Definizione 15.2. Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,q}(\mathbb{K})$ due matrici tali che il numero di colonne di A concida con il numero di righe di B. Si definisce **prodotto righe per colonne** o semplicemente **prodotto** di A per B la matrice di tipo $m \times q$ il cui elemento di posto (i, j) è dato dal prodotto della riga i-ma di A per la colonna j-ma di B. In simboli:

$$AB := (A^{(i)}B_{(j)}). \tag{8}$$

Esempio 15.3. Posto $A = \begin{pmatrix} \sqrt{2} & 3 & -1 \\ 0 & 1 & 5 \end{pmatrix}$ e $B = \begin{pmatrix} 4 & \sqrt{2} & 0 \\ -2 & 1 & -2 \\ 0 & 1 & 0 \end{pmatrix}$, risulta che AB è di tipo 2×3 e si calcola come segue:

$$AB = \begin{pmatrix} \sqrt{2} \cdot (4) + 3 \cdot (-2) + (-1) \cdot 0 & \sqrt{2} \cdot \sqrt{2} + 3 \cdot 1 + (-1) \cdot 1 & \sqrt{2} \cdot 0 + 3 \cdot (-2) + (-1) \cdot 0 \\ 0 \cdot 4 + 1 \cdot (-2) + 5 \cdot 0 & 0 \cdot \sqrt{2} + 1 \cdot 1 + 5 \cdot 1 & 0 \cdot 0 + 1 \cdot (-2) + 5 \cdot 0 \end{pmatrix}$$

$$AB = \begin{pmatrix} 4\sqrt{2} - 6 & 6 & -6 \\ -2 & 6 & -2 \end{pmatrix}.$$

Notiamo anche che in questo caso il prodotto BA non ha senso.

Utilizzando questa nozione di prodotto, anche la definizione delle funzioni L_A si generalizza come segue:

Definizione 15.4. Sia $A \in M_{m,n}(\mathbb{K})$ una matrice fissata; ad essa si associa un'applicazione

$$L_A : \mathbb{K}^n \to \mathbb{K}^m$$

definita come segue:

$$\forall x \in \mathbb{K}^n \quad L_A(x) := Ax. \tag{9}$$

Qui i vettori di $\mathbb{K}^n$ e di $\mathbb{K}^m$ sono identificati sempre con i corrispondenti *vettori colonna* ed il prodotto Ax è il prodotto righe per colonne.

Proposizione 15.5. *Per ogni matrice $A \in M_{m,n}(\mathbb{K})$, l'applicazione $L_A : \mathbb{K}^n \to \mathbb{K}^m$ ad essa associata è lineare.*

Dimostrazione: Notiamo infatti che per ogni vettore colonna $x \in \mathbb{K}^n$ risulta

$$L_A(x) = Ax = \begin{pmatrix} A^{(1)}x \\ \vdots \\ A^{(m)}x \end{pmatrix},$$

e quindi

$$L_A(x) = (L_{A^{(1)}}(x), \cdots, L_{A^{(m)}}(x)).$$

Pertanto, in base al Teorema 14.4, tale applicazione è lineare, in quanto tutte le sue funzioni componenti sono lineari.

□

A questo punto possiamo riformulare il Teorema 14.4 come segue:

Teorema 15.6. *Tutte e sole le applicazioni lineari $F : \mathbb{K}^n \to \mathbb{K}^m$ sono le funzioni del tipo*

$$F = L_A$$

con $A \in M_{m,n}(\mathbb{K})$ matrice assegnata di tipo (m, n).

Per provarlo, resta da giustificare che ogni $F : \mathbb{K}^n \to \mathbb{K}^m$ lineare è del tipo $F = L_A$. Sappiamo che F è della forma

$$F(x) = (F_1(x), \cdots, F_m(x)), \quad \forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n,$$

dove ciascuna F_i è lineare; ma a sua volta, come già stabilito sopra, F_i è del tipo

$$F_i = L_{u^i} \qquad i = 1, \ldots, m \tag{10}$$

per un certo vettore riga $u^i \in \mathbb{K}^n$. Consideriamo qundi la matrice

$$A := \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$$

che ha m righe (una per ogni funzione F_i) e n colonne. Per costruzione si ha immediatamente che $F = L_A$.

Esempio 15.7. Data la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, allora la funzione L_A associata è $L_A : \mathbb{R}^3 \to \mathbb{R}^2$, data da

$$L_A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6_3 \end{pmatrix}$$

ovvero

$$L_A(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 4x_1 + 5x_2 + 6_3).$$

Facciamo qualche ulteriore osservazione sul prodotto righe per colonne.

Osservazione 15.8. a) Date le matrici $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,q}(\mathbb{K})$ le righe e le colonne della matrice prodotto AB si possono ricavare come segue:

$$(AB)^{(i)} = A^{(i)}B$$

$$(AB)_{(j)} = AB_{(j)}.$$

La verifica è semplice conseguenza della defnizione del prodotto; ad esempio:

$$(AB)^{(i)} = (A^{(i)}B_{(1)} \cdots A^{(i)}B_{(q)}) = A^{(i)}B.$$

b) Data $A \in M_{m,n}(\mathbb{K})$ le sue righe e colonne si possono ricostruire mediante moltiplicazione con i vettori della base canonica di $\mathbb{K}^m$ e di $\mathbb{K}^n$; precisamente si ha:

$$Ae_i = A_{(i)},$$

dove i vettori della base canonica di $\mathbb{K}^n$ sono identificati con vettori colonna, mentre

$$e_iA = A^{(i)},$$

dove ciascun vettore della base canonica di $\mathbb{K}^m$ è pensato come vettore riga.

Per giustificarlo, basta ricordare che per ogni vettore riga $u = (u_1 \dots u_n)$ si ha $(L_u)e_i = u_i$ e usare questo fatto per ogni di riga di A:

$$Ae_i = \begin{pmatrix} A^{(1)}e_i \\ \vdots \\ A^{(m)}e_i \end{pmatrix} = \begin{pmatrix} L_{A^{(1)}}e_i \\ \vdots \\ L_{A^{(m)}}e_i \end{pmatrix} = \begin{pmatrix} a_i^1 \\ \vdots \\ a_i^m \end{pmatrix} = A_{(i)}.$$

Analogamente per ogni vettore colonna $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ si ha:

$$e_i \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 \cdots & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = u_i.$$

Quindi

$$e_i A = (\, e_i A_{(1)} \, \cdots \, e_i A_{(n)} \,) = (a^i_1 \, \cdots \, a^i_n \,) = A^{(i)}.$$

In particolare, abbiamo ottenuto, con riferimento all'applicazione lineare L_A, che

(11) $$L_A(e_i) = A_{(i)}.$$

Riassumiamo quanto detto sin qui: data una generica applicazione lineare

$$F = (F_1, \ldots, F_m) = L_A : \mathbb{K}^n \to \mathbb{K}^m,$$

la matrice A contiene le seguenti informazioni:

• Le righe corrispondono alle funzioni lineari componenti $F_1, \ldots, F_m$, nel senso che ciascuna F_i è ricostruibile dalla corrispondente riga come $L_{A^{(i)}}$.

• Le colonne corrispondono alle immagini $F(e_i)$ dei vettori della base canonica dello spazio di partenza $\mathbb{K}^n$.

Inoltre, data $F = L_A$, è possibile calcolare $F(x)$ direttamente a partire dalle colonne di A, mediante la seguente relazione:

(12) $$L_A(x) = x_1 A_{(1)} + x_2 A_{(2)} + \cdots + x_n A_{(n)}.$$

che esprime ogni immagine $F(x)$ come combinazione lineare delle colonne della matrice.

Questa relazione è di dimostrazione immediata, sfruttando la (14.2) e tenendo conto della (11):

$$L_A(x) = L_A(x_1 e_1 + \cdots + x_n e_n) = x_1 F(e_1) + \cdots + x_n F(e_n) = x_1 A_{(1)} + \cdots + x_n A_{(n)}.$$

Proposizione 15.9. *Siano $A, B \in M_{m,n}(\mathbb{K})$ e $\lambda \in \mathbb{K}$. Allora:*

1) $A = B \iff L_A = L_B$.
2) $L_{A+B} = L_A + L_B$.
3) $L_{\lambda A} = \lambda L_A$.

Dimostrazione: 1) L'implicazione $\Rightarrow$ è ovvia. Supponiamo che $L_A = L_B$. Allora in particolare

$$Ae_i = Be_i$$

che si può riscrivere

$$A_{(i)} = B_{(i)}$$

e ciò comporta che $A = B$.

2) Si tratta di un'uguaglianza tra applicazioni lineari da $\mathbb{K}^n$ in $\mathbb{K}^m$; quindi basta verificare che esse assumino gli stessi valori quando applicate ai vettori della base canonica di $\mathbb{K}^n$; risulta in effetti:

$$L_{A+B}(e_i) = (A+B)e_i = (A+B)_{(i)} = A_{(i)} + B_{(i)} = L_A(e_i) + L_B(e_i) = (L_A + L_B)(e_i).$$

La verifica di 3) è simile e si lascia per esercizio. $\square$

Proposizione 15.10. *Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,q}(\mathbb{K})$. Allora*

$$L_{AB} = L_A \circ L_B.$$

Dimostrazione: Notiamo che $L_{AB} : \mathbb{K}^q \to \mathbb{K}^m$ e anche $L_A \circ L_B : \mathbb{K}^q \to \mathbb{K}^m$; anche in questo caso confrontiamo i valori assunti dalle due applicazioni lineari sui vettori della base canonica di $\mathbb{K}^q$; tenendo conto della (15.8) si ha:

$$L_{AB}(e_i) = (AB)_{(i)} = AB_{(i)} = L_A(L_B(e_i)) = (L_A \circ L_B)(e_i).$$

$\square$

Come applicazione delle Proposizioni appena stabilite, siamo in grado di ricavare le seguenti proprietà basilari delle operazioni tra matrici:

Corollario 15.11. *Le operazioni di somma e prodotto tra matrici soddisfano:*

1) *Siano $A, B \in M_{m,n}(\mathbb{K})$, $C \in M_{n,q}(\mathbb{K})$ e $D \in M_{s,m}(\mathbb{K})$. Allora*

$$(A+B)C = AC + BC,$$
$$D(A+B) = DA + DB.$$

2) *Siano $A \in M_{m,n}(\mathbb{K}), B \in M_{n,q}(\mathbb{K})$ e $C \in M_{q,s}(\mathbb{K})$. Allora*

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

Dimostrazione: Tutte queste uguaglianze tra matrici si possono provare utilizzando il criterio fornito dalla 1) della Prop. 15.9, mostrando l'uguaglianza delle corrispondenti applicazioni lineari. Nel caso 1) si ha, ad esempio:

$$L_{(A+B)C} = L_{A+B} \circ L_C = (L_A + L_B) \circ L_C = L_A \circ L_C + L_B \circ L_C = L_{AC} + L_{BC} = L_{AC+BC}.$$

Quanto a 2):

$$L_{A(BC)} = L_A \circ L_{BC} = L_A \circ (L_B \circ L_C) = (L_A \circ L_B) \circ L_C = L_{AB} \circ L_C = L_{(AB)C}.$$

□

16. Classificazione degli spazi $\mathbb{K}^n$ a meno di isomorfismo

In questo paragrafo proviamo alcuni risultati significativi, su cui verrà basata la nozione di dimensione di uno spazio vettoriale. Cominciamo col provare che se $n > m$, lo spazio $\mathbb{K}^n$ è "più grande" dello spazio $\mathbb{K}^m$; cioè che non è possibile trovare una copia isomorfa del primo nel secondo.

Ricordiamo che due spazi vettoriali V e V' sono isomorfi se esiste almeno un isomorfismo tra essi; in tal caso scriviamo $V \cong V'$.

Proposizione 16.1. *Siano V, V' e V'' spazi vettoriali sullo stesso campo $\mathbb{K}$. Allora*
a) $V \cong V$.
b) $V \cong V' \Rightarrow V' \cong V$.
c) $V \cong V' \wedge V' \cong V'' \Rightarrow V \cong V''$.

La prima affermazione si ottiene considerando l'applicazione identica $Id_V : V \to V$. È immediato che si tratta di un isomorfismo. Le c) e b) sono immediate conseguenza del fatto che la composizione di isomorfismi è un isomorfismo, e del fatto che l'inverso di un isomorfismo è anch'esso un isomorfismo.

Assegnato dunque un campo $\mathbb{K}$, la "relazione di isomorfismo" è una *relazione di equivalenza* sull'insieme di tutti i $\mathbb{K}$-spazi vettoriali. Ciò permette di "classificare" gli spazi vettoriali su $\mathbb{K}$ disponendoli nelle classi di equivalenza corrispondenti. Dato uno spazio, la classe cui appartiene contiene tutti e soli gli spazi ad esso isomorfi, che potranno essere opportunamente identificati con esso. Nel seguito ci occuperemo di tale classificazione nel caso di spazi vettoriali di dimensione finita, dando un preciso significato matematico all'attributo "dimensione".

Cominciamo dunque dalla classificazione, a meno di isomorfismo degli spazi $\mathbb{K}^n$.

Osserviamo che, intuitivamente ci aspettiamo che per m e n diversi, $\mathbb{K}^m$ e $\mathbb{K}^n$ siano spazi sostanzialmente diversi, dato il numero diverso di parametri scalari liberi che occorrono per individuarne univocamente i vettori. Anzi, pensando al caso $\mathbb{K} = \mathbb{R}$ e all'interpretazione geometrica di $\mathbb{R}^2$ e $\mathbb{R}^3$ come piano Euclideo e spazio Euclideo tridimensionale, è lecito ritenere che lo spazio di "dimensione più grande" contenga al suo interno qualche copia di quello "di dimensione più piccola". In effetti non è difficile convincersi che, se $n \leq m$, tali copie esistono, ovvero più precisamente che $\mathbb{K}^m$ contiene sempre qualche *sottospazio* vettoriale W *isomorfo* a $\mathbb{K}^n$, se $n \leq m$.

Ad esempio, un sottospazio vettoriale di $\mathbb{R}^3$ isomorfo a $\mathbb{R}^2$ è

$$W = \{(x, y, 0) : x, y \in \mathbb{R}\},$$

un isomorfismo essendo:

$$(x, y) \in \mathbb{R}^2 \mapsto (x, y, 0) \in W.$$

Si noti che questa applicazione è manifestamente una funzione lineare da $\mathbb{R}^2$ in $\mathbb{R}^3$, che possiamo considerare a valori in W, conservando il carattere di linearità.

Per costruire un tale W in generale, assumendo $n \leq m$ è sufficiente considerare l'applicazione lineare

$$F : \mathbb{K}^n \to \mathbb{K}^m$$

definita da

$$F(x_1, \ldots, x_n) := (x_1, \ldots, x_n, 0, \ldots, 0).$$

È chiaro che le funzioni componenti di F sono lineari (si tratta di $(p_1, \ldots, p_n, 0, \ldots, 0)$, dove le p_i sono proiezioni e le restanti funzioni nulle), per cui F è lineare. Tale funzione è ingettiva perchè se $(x_1, \ldots, x_n) \in Ker(F)$, allora

$$(x_1, \ldots, x_n, 0, \ldots, 0) = (0, \ldots, 0)$$

e quindi $x_1 = \cdots = x_n = 0$. Dunque il $Ker(F)$ è banale.

Il sottospazio $W := Im(F)$ di $\mathbb{K}^m$ è allora isomorfo a $\mathbb{K}^n$ mediante la ridotta di questa funzione:

$$F : \mathbb{K}^n \to Im(F),$$

che è automaticamente anche surgettiva, oltre che ingettiva.

L'intuizione suggerisce invece che questo tipo di costruzione non sia possibile se $n > m$. Proveremo ciò rigorosamente. Prima di procedere, è opportuno includere nella famiglia degli spazi vettoriali $\mathbb{K}^n$ anche lo spazio $\mathbb{K}^0$:

Definizione 16.2. Dato un campo $\mathbb{K}$, denoteremo con $\mathbb{K}^0$ il sottospazio banale $\{0_{\mathbb{K}}\}$ di $\mathbb{K}$, detto anch'esso spazio vettoriale banale o 0-dimensionale canonico sul campo $\mathbb{K}$.

Ciò premesso, sussiste il seguente:

Teorema 16.3. *Siano n, m interi con $n > m \geq 0$. Allora non esistono applicazioni lineari ingettive $F : \mathbb{K}^n \to \mathbb{K}^m$.*

Prima di procedere con la dimostrazione, cominciamo con qualche osservazione preliminare. Data un'applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$, $n \geq 2$,

$$F(x_1, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

tale che $a_1 = 0$ allora F è del tipo:

$$F(x) = a_2x_2 + \cdots + a_nx_n.$$

e diremo quindi che F *non dipende da* x_1; in tal caso F si può trattare come una funzione lineare $\bar{F} : \mathbb{K}^{n-1} \to \mathbb{K}$; più precisamente denotiamo con $\bar{F} : \mathbb{K}^{n-1} \to \mathbb{K}$ l'applicazione lineare definita da:

$$\bar{F}(y_1, \ldots, y_n) = a_2y_1 + \cdots + a_ny_n.$$

Allora per ogni $x \in \mathbb{K}^n$, posto $\bar{x} = (x_2, \dots, x_n)$ abbiamo che:

$$F(x) = \bar{F}(\bar{x}).$$

Notiamo anche che la condizione che F non dipenda da x_1 è equivalente a $F(e_1) = 0$ perchè a priori si ha sempre, come sappiamo, $a_1 = F(e_1)$.

Lemma 16.4. *Siano $F : \mathbb{K}^n \to \mathbb{K}^n$ e $G : \mathbb{K}^n \to \mathbb{K}$ due applicazioni lineari, $n \geq 2$. Allora l'applicazione lineare $F(e_1)G - G(e_1)F : \mathbb{K}^n \to \mathbb{K}$ non dipende da x_1.*

La dimostrazione di questo Lemma è immediata: si ha infatti

$$(F(e_1)G - G(e_1)F)(e_1) = F(e_1)G(e_1) - G(e_1)F(e_1) = 0.$$

Nel seguito, per ogni vettore $\bar{x} \in \mathbb{K}^{n-1}$, $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_{n-1})$, denotiamo per ogni $x_1 \in \mathbb{K}$ con $(x_1, \bar{x})$ il vettore $(x_1, \bar{x}_1, \cdots, \bar{x}_{n-1})$ di $\mathbb{K}^n$.

Lemma 16.5. *Data un'applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$, con $n \geq 2$, che dipende da x_1, cioè tale che $F(e_1) \neq 0$ e un vettore $\bar{x} \in \mathbb{K}^{n-1}$, esiste sempre $x_1 \in \mathbb{K}$ tale che:*

$$F(x_1, \bar{x}) = 0.$$

Dimostrazione: Infatti, abbiamo:

$$F(x_1, \bar{x}) = a_1 x_1 + a_2 \bar{x}_1 + \cdots + a_n \bar{x}_{n-1},$$

e quindi affinchè si abbia $F(x_1, \bar{x}) = 0$, si tratta di risolvere l'equazione:

$$a_1 x_1 + a_2 \bar{x}_1 + \cdots + a_n \bar{x}_{n-1} = 0,$$

che, essendo per ipotesi $a_1 \neq 0$, ammette l'unica soluzione:

$$x_1 := -\frac{1}{a_1} (a_2 \bar{x}_1 + \cdots + a_n \bar{x}_{n-1}).$$

□

Ciò premesso, discutiamo la dimostrazione del Teorema 16.3. Ragionando per induzione su $n \geq 1$, proviamo che ogni applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}^m$ con $0 \leq m < n$ non è ingettiva. Ciò si può riformulare come segue: per ogni $n \geq 1$, ogni applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}^m$ con $0 \leq m < n$ ha nucleo non banale. Il passo base ($n = 1$) consiste nell'esaminare un'applicazione lineare

$$F : \mathbb{K} \to \mathbb{K}^0$$

cioè

$$F : \mathbb{K} \to \{0_{\mathbb{K}}\}.$$

Tale applicazione è necessariamente costante, ovvero è la funzione nulla $F = 0$. Essa non è certamente ingettiva, in quanto il campo $\mathbb{K}$ contiene, oltre lo 0, almeno l'elemento $1 \neq 0$, per cui $F(1) = 0$.

Sia $n > 1$ e proviamo l'asserto, assumendolo vero per $n-1$. Anche in questo caso, se $m = 0$, non c'è nulla da provare in quanto l'unica applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}^0$ è la funzione nulla $F = 0$, che non è ingettiva. Supponiamo quindi $m > 0$. Sappiamo dal Teorema 14.4 che F è della forma

$$F(x) := (F_1(x), \cdots, F_m(x)), \quad \forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n,$$

dove le $F_i : \mathbb{K}^n \to \mathbb{K}$ sono lineari. Se $m = 1$, si tratta ancora di un'applicazione lineare $F : \mathbb{K}^n \to \mathbb{K}$, e la prova che il nucleo di F è non banale è immediata; infatti, se $F(e_1) = 0$ non c'è nulla da provare, altrimenti basta applicare il Lemma 16.5 a partire da un qualsiasi vettore non nullo $\bar{x} \in \mathbb{K}^{n-1}$.

Trattiamo quindi il caso $m > 1$. Se $F(e_1) = 0$, si ha la tesi; altrimenti deve esistere una funzione componente F_i per cui $F_i(e_1) \neq 0$. A meno di scambiare due delle funzioni componenti $F_1, \ldots, F_m$, senza con questo alterare il nucleo di F, possiamo supporre che $F_1(e_1) \neq 0$. Allora, posto $\alpha = F_1(e_1)$ abbiamo che tutte le funzioni:

$$\alpha F_i - F_i(e_1)F_1, \quad i = 2, \ldots, m$$

non dipendono da x_1 (Lemma 16.4). Denotate con $G_i : \mathbb{K}^n \to \mathbb{K}$ queste funzioni, possiamo applicare l'ipotesi induttiva all'applicazione lineare:

$$G = (\bar{G}_2, \ldots, \bar{G}_m) : \mathbb{K}^{n-1} \to \mathbb{K}^{m-1}.$$

Ciò è possibile perchè $0 \leq m-1 < n-1$. Quindi esiste un vettore $\bar{x} = (\bar{x}_2, \ldots, \bar{x}_n) \in \mathbb{K}^{n-1}$, $\bar{x} \neq 0$, tale che $G(\bar{x}) = 0$. Sfrutteremo questo vettore per determinare un vettore non nullo nel nucleo di F. Scegliamo a tal scopo $x_1 \in \mathbb{K}$ in modo che

$$F_1(x_1, \bar{x}) = 0,$$

il che è sempre possibile applicando il Lemma 16.5. Con questa scelta risulta che il vettore non nullo $(x_1, \bar{x}) \in \mathbb{K}^n$ appartiene al nucleo di F; infatti, poichè $\bar{x}$ appartiene a $Ker(G)$ risulta, per ogni $i = 2, \ldots, m$:

$$0 = \bar{G}_i(\bar{x}) = G_i(x_1, \bar{x}) = \alpha F_i(x_1, \bar{x}) - F_i(e_1)F_1(x_1, \bar{x}) = \alpha F_i(x_1, \bar{x}).$$

Pertanto possiamo concludere che $F_i(x_1, \bar{x}) = 0$ per ogni $i = 2, \ldots, m$, ancora tenendo conto del fatto che $\alpha \neq 0$. $\square$.

Corollario 16.6. *(Teorema di classificazione) Siano $\mathbb{K}$ un campo e $n, m \in \mathbb{N}$. Allora*

$$\mathbb{K}^n \cong \mathbb{K}^m \text{ se e solo se } n = m.$$

La dimostrazione è un'immediata applicazione del risultato precedente: infatti, se $\mathbb{K}^n \cong \mathbb{K}^m$, allora dev'essere $n = m$ perchè, fissato un isomorfismo $F : \mathbb{K}^n \to \mathbb{K}^m$, è impossibile che si abbia $n > m$, ma considerando l'isomorfismo inverso $F^{-1} : \mathbb{K}^m \to \mathbb{K}^m$, è anche impossibile che $m > n$.

Corollario 16.7. *Fissato un campo $\mathbb{K}$ e due interi $n > m$, non esiste alcun sottospazio vettoriale W di $\mathbb{K}^m$ isomorfo a $\mathbb{K}^n$.*

Infatti, supponiamo che un tale sottospazio W esista e sia $F : \mathbb{K}^n \to W$ un isomorfismo. Allora la funzione composta

$$i \circ F : \mathbb{K}^n \to \mathbb{K}^m$$

con $i : W \to \mathbb{K}^m$ l'inclusione canonica di W in $\mathbb{K}^m$, ovvero $i = Id|_W : W \to \mathbb{K}^m$, con Id applicazione identica di $\mathbb{K}^m$, sarebbe un'applicazione lineare ingettiva, il che è impossibile essendo $n > m$.

17. Spazi vettoriali finitamente generati

Definizione 17.1. Sia V un $\mathbb{K}$-spazio vettoriale. Una sequenza finita $v_1, \dots, v_m$ di vettori di V, $m \geq 1$, è detta **sistema di generatori** di V se ogni vettore v di V si può esprimere come combinazione lineare dei vettori della sequenza, ovvero se per ogni $v \in V$ esistono scalari $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tali che

$$v = \lambda_1 v_1 + \cdots \lambda_m v_m.$$

Osserviamo che la condizione in esame si può sintetizzare come segue; una sequenza $A = \{v_1, \dots, v_m\}$ è un sistema di generatori di V se e solo se:

$$V = L(v_1, \dots, v_k).$$

Mediante questa nozione introduciamo la classe di spazi vettoriali che saranno studiati sistematicamente nel seguito.

Definizione 17.2. Uno spazio vettoriale V sul campo $\mathbb{K}$ si dice **finitamente generato** oppure **di dimensione finita** se ammette un sistema di generatori (finito) $v_1, \dots, v_m$.

Esempio 17.3. Ogni spazio $\mathbb{K}^n$ è finitamente generato: ciò è ovvio se $n = 0$; se $n > 0$, sappiamo dalla (5) che ogni vettore è combinazione lineare dei vettori della base canonica $\{e_1, \dots, e_n\}$ introdotta nel §14.

Esempio 17.4. Consideriamo lo spazio vettoriale reale V costituito da *tutti* i polinomi di una variabile reale:

$$V = \{p : \mathbb{R} \to \mathbb{R} |\, p = p(x) \text{ è un polinomio}\}.$$

e il sottoinsieme:

$$W = \{p : \mathbb{R} \to \mathbb{R} |\, p = p(x) \text{ è un polinomio di grado } \leq 3\}.$$

Certamente W non è vuoto, e poichè la somma $p + q$ di due polinomi di grado al massimo 3 è anch'esso un polinomio dello stesso tipo, e analogamente per λp, con λ costante, risulta che W è un sottospazio vettoriale di V. In particolare, anche W è un $\mathbb{R}$-spazio vettoriale.

Risulta che
$$W = L(1, x, x^2, x^3)$$
per cui W è finitamente generato. Infatti ogni polinonio $p \in W$ è una funzione del tipo
$$p(x) = a_0 \cdot 1 + a_1 x + a_2 x^2 + a_3 x^3,$$
dove gli $a_i \in \mathbb{R}$ sono costanti.

Invece lo spazio V *non* è finitamente generato. Infatti, ammettiamo per assurdo che esista un sistema di generatori $p_1, \dots, p_m$ di W: allora ogni polinomio $p : \mathbb{R} \to \mathbb{R}$ sarebbe della forma
$$p = \lambda_1 p_1 + \cdots + \lambda_m p_m$$
per opportune costanti λ_i. Detto M il grado massimo dei polinomi $p_1, \dots, p_m$ seguirebbe da ciò che il grado di *ogni* polinomio p è al massimo M. Ciò è evidentemente impossibile, ad esempio il polinomio x^{M+1} ha grado $M + 1$.

Vediamo ora che c'è uno stretto legame tra la nozione di sistema di generatori e quella di epimorfismo.

In corrispondenza di una successione finita $A = \{v_1, \dots, v_m\}$ di vettori qualsiasi di un $\mathbb{K}$-spazio vettoriale V possiamo definire un'applicazione
$$\varphi_A : \mathbb{K}^m \to V$$
nel modo seguente:

$$\varphi_A(\lambda_1, \dots, \lambda_m) := \lambda_1 v_1 + \cdots + \lambda_m v_m. \tag{13}$$

Risulta che φ_A è un'appicazione lineare: la verifica è sostanzialmente la stessa che è stata discussa per giustificare che $L(v_1, \dots, v_m)$ è un sottospazio: infatti, dati i vettori $(\lambda_1, \dots, \lambda_m)$ e $(\mu_1, \dots, \mu_m)$ di $\mathbb{K}^m$, allora:
$$\begin{aligned}
&\varphi_A((\lambda_1, \dots, \lambda_m) + (\mu_1, \dots, \mu_m)) = \varphi_A((\lambda_1 + \mu_1, \dots, \lambda_m + \mu_m)) = \\
&(\lambda_1 + \mu_1) v_1 + \cdots + (\lambda_m + \mu_m) v_m = \\
&(\lambda_1 v_1 + \cdots + \lambda_m v_m) + (\mu_1 v_1 + \cdots + \mu_m v_m) = \\
&\varphi_A(\lambda_1, \dots, \lambda_m) + \varphi_A(\mu_1, \dots, \mu_m).
\end{aligned}$$
Il lettore può agevolmente completare la verifica della linearità di φ_A tenendo conto della (3).

Notiamo anche che, per costruzione:
$$Im(\varphi_A) = L(v_1, \dots, v_m).$$

Consegue che sussiste la seguente utile caratterizzazione dei sistemi di generatori:

Proposizione 17.5. *Data una squenza $A = \{v_1, \dots, v_m\}$ di vettori di uno spazio vettoriale V, allora:*
$$v_1, \dots, v_m \text{ è un sistema di generatori di } V \iff \varphi_A \text{ è surgettiva.}$$

Dimostrazione: Infatti, dire che $v_1, \ldots, v_m$ è un sistema di generatori di V significa, per definizione, che $V = L(v_1, \ldots, v_m) = Im(\varphi_A)$ il che è vero se e solo se φ_A è surgettiva. $\square$

18. vettori linearmente dipendenti e linearmente indipendenti

Sia V uno spazio vettoriale sul campo $\mathbb{K}$.

Definizione 18.1. Si dice che i vettori $v_1, \ldots, v_m$ di V sono *linearmente dipendenti* se esistono m scalari **non tutti nulli** tali che:

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0_V.$$

Si introduce a questo proposito la seguente terminologia:

Definizione 18.2. Una combinazione lineare

$$\lambda_1 v_1 + \cdots + \lambda_m v_m$$

di vettori di V verrà detta:

-*nulla* se coincide con il vettore nullo, cioè se $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0_V$.

-*banale* se tutti i coefficienti λ_i sono uguali a 0.

Con questa terminologia, i vettori $v_1, \ldots, v_m$ sono linearmente dipendenti se esiste una loro combinazione lineare nulla ma **non banale**.

Osserviamo esplicitamente che ogni combinazione lineare banale è sempre nulla, ma il viceversa può non essere vero: ad esempio, la seguente combinazione lineare

$$(1,1) - \frac{1}{3}(3,3)$$

dei vettori $(1,1)$ e $(3,3)$ di $\mathbb{R}^2$ è nulla, ma non è banale. Pertanto i vettori in questione sono linearmente dipendenti.

Esempio 18.3. I vettori $(1,0,-1), (0,2,3), (2,2,1)$ di $\mathbb{R}^3$ sono linearmente dipendenti.

Per stabilire ciò cerchiamo di determinarne una combinazione lineare nulla non banale, imponendo

$$\lambda_1(1,0,-1) + \lambda_2(0,2,3) + \lambda_3(2,2,1) = 0.$$

Si tratta di stabilire se il seguente sistema di equazioni lineari ammette una soluzione non banale $(\lambda_1, \lambda_2, \lambda_3)$:

$$\begin{cases} \lambda_1 + 2\lambda_3 = 0 \\ 2\lambda_2 + 2\lambda_3 = 0 \\ -\lambda_1 + 3\lambda_2 + \lambda_3 = 0 \end{cases} .$$

Risolvendo con il metodo di sostituzione, otteniamo

$$\begin{cases} \lambda_1 = -2\lambda_3 \\ \lambda_2 = -\lambda_3 \\ 0 = 0 \end{cases}.$$

Quindi tutte le soluzioni sono le terne $(-2\lambda_3, -\lambda_3, \lambda_3)$ al variare della variabile libera λ_3. Ponendo ad es. $\lambda_3 = 1$ otteniamo una soluzione non banale, corrispondente alla seguente combinazione lineare nulla non banale dei tre vettori:

$$-2(1, 0, -1) - (0, 2, 3) + (2, 2, 1) = 0.$$

Analogamente a quanto fatto in merito ai sistemi di generatori, possiamo utilizzare ancora la funzione φ_A per caratterizzare la nozione di dipenendenza lineare tra vettori:

Proposizione 18.4. *Si consideri una successione finita* $A = \{v_1, \dots, v_m\}$*,* $m \geq 1$ *di vettori dello spazio* V*. Allora:*

$$v_1, \dots, v_m \textit{ sono linearmente dipendenti } \iff \varphi_A \textit{ non è ingettiva.}$$

Dimostrazione: Per definizione, $v_1, \dots, v_m$ sono linearmente dipendenti se e solo se se esiste un vettore $(\lambda_1 \dots, \lambda_m) \in \mathbb{K}^m$, $(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$, tale che:

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0_V,$$

cioè

$$\varphi_A(\lambda_1, \dots, \lambda_m) = 0_V,$$

ovvero tale che

$$(\lambda_1 \dots, \lambda_m) \in Ker(\varphi_A).$$

In conclusione, abbiamo provato che $v_1, \dots, v_m$ sono linearmente dipendenti se e solo se $Ker(\varphi_A) \neq \{0_{\mathbb{K}^m}\}$, da cui l'asserto. $\square$

Definizione 18.5. Diremo che i vettori $v_1, \dots, v_m$ sono **linearmente indipendenti** se non sono linearmente dipendenti.

Dunque $v_1, \dots, v_m$ sono linearmente indipendenti se e solo se ogni relazione del tipo

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0_V$$

implica sempre che $\lambda_1 = \cdots = \lambda_m = 0$, ovvero:

$$\forall (\lambda_1, \cdots, \lambda_m) \in \mathbb{K}^m : \ \lambda_1 v_1 + \cdots + \lambda_m v_m = 0 \ \Rightarrow \ \lambda_1 = \cdots = \lambda_m = 0.$$

Ovviamente, data una sequenza $A = \{v_1, \dots, v_m\}$, sussiste la caratterizzazione:

$$v_1, \dots, v_m \text{ sono linearmente indipendenti } \iff \varphi_A \text{ è un monomorfismo.}$$

Osservazione 18.6. Una terminologia molto in uso in algebra lineare è la seguente: se $v_1, \dots, v_m$ sono vettori linearmente indipendenti, allora l'insieme $A = \{v_1, \dots, v_m\}$ è detto *libero*.

19. basi

Introduciamo ora una nozione cruciale per lo studio degli spazi vettoriali finitamente generati.

Definizione 19.1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Una sequenza finita di vettori $\{v_1, \dots, v_n\}$ si dice **base** di V se:

1) $v_1, \dots, v_n$ costituiscono un sistema di generatori di V.
2) $v_1, \dots, v_n$ sono linearmente indipendenti.

Esempio 19.2. La base canonica $\{e_1, \dots, e_n\}$ di $\mathbb{K}^n$ è effettivamente una base nel senso della definizione precedente; osserviamo che l'indipendenza dei vettori $e_1, \dots, e_n$ è ancora conseguenza immediata dell'identità fondamentale (5): infatti, data una combinazione lineare nulla

$$\lambda_1 e_1 + \cdots + \lambda_n e_n = 0_{\mathbb{K}^n}$$

il vettore al primo membro si riscrive:

$$(\lambda_1, \dots, \lambda_n) = 0_{\mathbb{K}^n},$$

e quindi necessariamente tutti i coefficienti λ_i sono nulli.

Esempio 19.3. Una base dello spazio vettoriale $M_2(\mathbb{K})$ è costituita dalla sequenza di matrici:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Ragionando in modo simile al caso della base canonica di $\mathbb{K}^4$, ciò si giustifica tenendo conto del fatto che per ogni matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{K})$ si ha:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

In base a quanto stabilito nei paragrafi precedenti, considerata una sequenza di vettori $\mathfrak{B} = \{v_1, \dots, v_n\}$, abbiamo:

$$\mathfrak{B} \text{ è base di } V \iff \varphi_{\mathfrak{B}} : \mathbb{K}^n \to V \text{ è un isomorfismo.}$$

Un'applicazione molto importante di ciò e del Teorema di classificazione 16.6 è il risultato seguente, che stabilisce che due basi sono sempre necessariamente costituite dallo stesso numero di vettori:

Teorema 19.4. *Siano $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C} = \{w_1, \ldots, w_m\}$ due basi di un $\mathbb{K}$-spazio vettoriale V. Allora:*

$$n = m.$$

Dimostrazione: Infatti, in corrispondenza della base $\mathfrak{B}$ abbiamo l'isomorfismo $\varphi_{\mathfrak{B}} : \mathbb{K}^n \to V$, mentre in corrispondenza dell'altra base $\mathfrak{C}$ abbiamo l'isomorfismo $\varphi_{\mathfrak{C}} : \mathbb{K}^m \to V$. Ne segue che $\mathbb{K}^n \cong \mathbb{K}^m$, e quindi necessariamente $n = m$. □

20. Proprietà dei vettori linearmente indipendenti

In questo paragrafo approfondiamo alcune proprietà utili relative al concetto di (in)dipenza lineare tra vettori. Cominciamo con l'osservare che, se una sequenza di vettori $v_1, \ldots, v_n$ contiene il vettore nullo, allora sicuramente i vettori sono linearmente dipendenti.

Infatti, se $v_j = 0$, allora

$$0v_1 + \cdots + 0v_{j-1} + v_j + 0v_{j+1} + \cdots + 0v_n = 0.$$

Lo stesso dicasi se due dei vettori della sequenza concidono; se ad es. $v_i = v_j$ con $i < j$ due indici fissati, allora

$$v_i - v_j = 0$$

e quindi $v_i - v_j$ è una combinazione lineare nulla non banale di $v_1, \ldots, v_n$ (si sottintende qui che i coefficienti che moltiplicano i vettori della sequenza di indice diverso da i e j sono tutti uguali a 0).

Si osserva anche che ogni sottosequenza $v_{i_1}, \ldots, v_{i_k}$ di una sequenza di vettori linearmente indipendenti $v_1, \ldots, v_n$ conserva il carattere di lineare indipendenza. Ciò perchè ogni combinazione lineare $\lambda_1 v_{i_1} + \cdots + \lambda_k v_{i_k}$ dei vettori della sottosequenza è anche una combinazione lineare dei vettori della sequenza originale $v_1, \ldots, v_n$. In altri termini, ogni sottoinsieme di un insieme libero di vettori è libero.

Sussiste la seguente caratterizzazione:

Proposizione 20.1. *Siano assegnati i vettori $v_1, \ldots, v_m$ dello spazio vettoriale V. Le seguenti proprietà sono equivalenti:*

a) *$v_1, \ldots, v_m$ sono linearmente dipendenti.*

b) *Almeno uno dei vettori $v_1, \ldots, v_m$ è combinazione lineare degli altri.*

Dimostrazione: *a)* $\Rightarrow$ *b)*. Si assuma che $v_1, \ldots, v_m$ siano linearmente dipendenti; allora esiste una combinazione lineare nulla:

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0_V$$

non banale; supponiamo che $\lambda_j \neq 0$; allora utilizzando l'inverso di λ_j possiamo ricavare v_j dalla relazione precedente ottenendo

$$v_j = -\lambda_j^{-1}\lambda_1 v_1 - \cdots - \lambda_j^{-1}\lambda_{j-1}v_{j-1} - \lambda_j^{-1}\lambda_{j+1}v_{j+1} - \cdots - \lambda_j^{-1}\lambda_m v_m.$$

e ciò prova b).

b) $\Rightarrow$ *a)*. Per ipotesi esiste $j \in \{1, \ldots, m\}$ tale che

$$v_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \cdots + \alpha_m v_m$$

per opportuni scalari α_i. Si ottiene pertanto:

$$0_V = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} - v_j + \alpha_{j+1} v_{j+1} + \cdots + \alpha_m v_m$$

il che mostra l'esistenza di una combinazione lineare nulla non banale dei vettori $v_1, \ldots, v_m$, datosi che $-1 \neq 0$. $\square$

Il risultato seguente è molto utilizzato; fornisce una condizione sufficiente affinchè si possa "allungare" una sequenza libera, aggiungendo un altro vettore, conservando la lineare indipendenza:

Proposizione 20.2. *Siano $v_1, \ldots, v_s$ vettori linearmente indipendenti. Allora, se $v \notin L(v_1, \ldots, v_s)$, anche i vettori $v_1, \ldots, v_s, v$ sono linearmente indipendenti.*

Dimostrazione: Siano $\lambda_1, \ldots, \lambda_{s+1} \in \mathbb{K}$ tali che

$$\lambda_1 v_1 + \cdots \lambda_s v_s + \lambda_{s+1} v = 0.$$

Cominciamo con l'osservare che non può essere $\lambda_{s+1} \neq 0$ perchè altrimenti si ricaverebbe

$$v = -\lambda_{s+1}^{-1} \lambda_1 \, v_1 - \cdots - \lambda_{s+1}^{-1} \lambda_s \, v_s$$

contro il fatto che $v \notin L(v_1, \ldots, v_s)$. Dunque $\lambda_{s+1} = 0$ e si ottiene:

$$\lambda_1 v_1 + \cdots + \lambda_s v_s = 0,$$

da cui $\lambda_1 = \cdots = \lambda_s = 0$ perchè i vettori $v_1, \ldots, v_s$ sono linearmente indipendenti. $\square$

21. Esistenza di basi

In questo paragrafo proviamo che ogni spazio vettoriale finitamente generato è sempre dotato di basi.

Vi è un caso particolare che merita un'osservazione preliminare.

Definizione 21.1. Uno spazio vettoriale V è detto **banale** se $V = \{0_V\}$.

Un tale spazio è ovviamente finitamente generato. Ma in esso non esistono vettori linearmente indipendenti. Si converrà pertanto che ogni spazio banale *ha dimensione 0 e che l'insieme vuoto $\emptyset$ è la sua unica base.*

Veniamo al caso di spazi vettoriali finitamente generati non banali. Descriveremo un algoritmo mediante il quale è sempre possibile estrarre da un sistema di generatori assegnato $v_1, \ldots, v_m$, $m \geq 1$, una base $\{v_{i_1}, \ldots, v_{i_n}\}$ con $1 \leq i_1 < \cdots < i_n \leq m$ e $n \geq 1$. A tale scopo, introduciamo la seguente terminologia:

Definizione 21.2. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $v_1, \ldots, v_m$ un sistema di generatori di V. Sia $i \in \{1, \ldots, m\}$. Il vettore v_i si dirà **eliminabile** se esso è combinazione lineare dei vettori che lo precedono, ovvero se

$$v_i \in L(v_1, \ldots, v_{i-1}).$$

Altrimenti v_i si dirà **non eliminabile**.

Per $i = 1$, il sottospazio $L(v_1, \ldots, v_{i-1})$ va inteso $\{0_V\}$, e quindi per definizione v_1 è eliminabile se e solo se è il vettore nullo.

Teorema 21.3. *Sia V un $\mathbb{K}$-spazio vettoriale finitamente generato non banale. Dato un sistema di generatori $v_1, \ldots, v_m$, siano*

$$v_{i_1}, \ldots, v_{i_n}, \quad 1 \leq i_1 < \cdots < i_n \leq m$$

i vettori non eliminabili da esso. Allora $\mathfrak{B} = \{v_{i_1}, \ldots, v_{i_n}\}$ è una base di V.

La base $\mathfrak{B}$ così costruita si dice estratta dal sistema di generatori $v_1, \ldots, v_m$ mediante *l'algoritmo di eliminazione* o algoritmo degli *scarti successivi.*

Dimostrazione: Mostriamo innanzitutto che se V è uno spazio vettoriale non banale, ogni sistema di generatori $v_1, \ldots, v_m$ contiene qualche vettore non eliminabile. Infatti, se tali vettori fossero tutti eliminabili, avremmo

$$V = L(v_1, \ldots, v_m) \subset L(v_1, \ldots, v_{m-1})$$

perchè tutti i vettori v_i appartengono al sottospazio $L(v_1, \ldots, v_{m-1})$, stante l'eliminabilità di v_m; da qui l'inclusione (si applichi la Prop. 10.2). Per lo stesso motivo, essendo anche v_{m-1} eliminabile, si ha:

$$V = L(v_1, \ldots, v_m) \subset L(v_1, \ldots, v_{m-1}) \subset L(v_1, \ldots, v_{m-2}).$$

Ripetendo questo argomento più volte si ottiene in definitiva

$$V \subset L(v_1);$$

ma $v_1 = 0_V$ essendo eliminabile, per cui si avrebbe

$$V = L(0_V) = \{0_V\}$$

contro l'ipotesi.

Ciò premesso, proviamo che $\mathfrak{B}$ è una base ragionando per induzione sulla lunghezza m della lista di vettori assegnata.

Nel caso $m = 1$, abbiamo per ipotesi

$$V = L(v_1)$$

e l'unico vettore v_1 della sequenza dev'essere necessariamente non eliminabile, perchè altrimenti si avrebbe $v_1 = 0$ e quindi $V = \{0_V\}$. Ora, $\mathfrak{B} = \{v_1\}$ è una base in quanto $\{v_1\}$ è un insieme libero avendosi, per ogni scalare λ:

$$\lambda v_1 = 0 \Rightarrow \lambda = 0$$

perchè $v_1 \neq 0$.

Supponiamo ora $m > 1$ e l'asserto vero per $m-1$. Consideriamo il sottospazio $W = L(v_1, \ldots, v_{m-1})$ di V e la sequenza $v_1, \ldots, v_{m-1}$ che ne costituisce un sistema di generatori.

Osserviamo che, se W è banale, allora $v_1 = \cdots = v_{m-1} = 0_V$ e quindi

$$V = L(v_1, \ldots, v_{m-1}, v_m) = L(0_V, \ldots, 0_V, v_m) = L(v_m)$$

e, ragionando come sopra, v_m dev'essere non eliminabile e $\mathfrak{B} = \{v_m\}$ è una sua base.

Consideriamo quindi il caso in cui W non è banale. Per l'ipotesi induttiva applicata a W e al suo sistema di generatori $v_1, \ldots, v_{m-1}$, la sequenza:

$$\mathfrak{C} = \{v_{i_1}, \ldots, v_{i_k}\}$$

è una base di W, dove $v_{i_1}, \ldots, v_{i_k}$ sono i vettori non eliminabili della sequenza $v_1, \ldots, v_{m-1}$. In particolare, i vettori di $\mathfrak{C}$ sono linearmente indipendenti. A questo punto vi sono due possibilità:

1) v_m è eliminabile

2) v_m non è eliminabile.

Nel primo caso i vettori non eliminabili della sequenza $v_1, \ldots, v_m$ sono tutti e soli quelli non eliminabili della sottosequenza $v_1, \ldots, v_{m-1}$, ovvero si ha $\mathfrak{B} = \mathfrak{C}$. Risulta che $\mathfrak{B}$ è di fatto una base di V perchè in tal caso risulta che $W = V$; infatti, stante l'eliminabilità di v_m abbiamo:

$$V = L(v_1, \ldots, v_m) \subset L(v_1, \ldots, v_{m-1}) = W.$$

Nel caso 2), abbiamo $\mathfrak{B} = \{v_{i_1}, \ldots, v_{i_k}, v_m\}$. Applicando la Prop. 20.2, i vettori di $\mathfrak{B}$ risultano linearmente indipendenti, perchè $v_{i_1}, \ldots, v_{i_k}$ lo sono e inoltre $v_m \notin L(v_{i_1}, \ldots, v_{i_k})$, essendo v_m non eliminabile. Infine

$$V = L(v_1, \ldots, v_m) \subset L(v_{i_1}, \ldots, v_{i_k}, v_m)$$

perchè tutti i vettori $v_1, \ldots, v_{m-1}, v_m$ appartengono al sottospazio $L(v_{i_1}, \ldots, v_{i_k}, v_m)$, in quanto $v_1, \ldots, v_{m-1}$ sono tutti combinazioni lineari di $v_{i_1}, \ldots, v_{i_k}$ essendo $\mathfrak{C}$ base di W. Quindi resta provato che i vettori di $\mathfrak{B}$ costituiscono un sistema di generatori di V e con ciò che $\mathfrak{B}$ è una base di V. $\square$

22. Dimensione di uno spazio vettoriale

I Teoremi 19.4 e 21.3 permettono di introdurre la seguente nozione fondamentale:

Definizione 22.1. Si chiama **dimensione** di un $\mathbb{K}$-spazio vettoriale V finitamente generato il numero $n \in \mathbb{N}$ di elementi di una sua qualsiasi base, dove $n = 0$ nel caso in cui V sia banale. La dimensione di uno spazio V si denota con

$$\mathtt{dim}_{\mathbb{K}}(V)$$

o semplicemente con $\mathtt{dim}(V)$.

Il seguente risultato fornisce una interpretazione alternativa dello stesso concetto:

Teorema 22.2. *La dimensione di un $\mathbb{K}$-spazio vettoriale V finitamente generato è l'unico intero $n \in \mathbb{N}$ tale che*
$$V \cong \mathbb{K}^n.$$

Dimostrazione: Poniamo $n = \mathtt{dim}_{\mathbb{K}}(V)$. Se V è banale, allora $n = 0$ e si ha $V \cong \mathbb{K}^0$. Infatti, l'applicazione nulla
$$0 : V = \{0_V\} \to \{0_{\mathbb{K}}\}$$
è un isomorfismo. Se V non è banale, il Teorema 21.3 garantisce che $n > 0$ ed esiste almeno una base $\mathfrak{B} = \{v_1, \ldots, v_n\}$, $n \geq 1$, necessariamente costituita da n vettori, per definizione di dimensione. In corrispondenza di tale base vi è l'isomorfismo $\varphi_{\mathfrak{B}} : \mathbb{K}^n \to V$ discusso nel §19.

Infine, n è l'unico numero intero con questa proprietà in forza del teorema di classificazione 16.6 degli spazi $\mathbb{K}^n$: se esistesse un altro intero m per cui $V \cong \mathbb{K}^m$, avremmo di conseguenza $\mathbb{K}^m \cong \mathbb{K}^n$, il che comporta $m = n$. □

Dunque si può decidere qual è la dimensione di uno spazio vettoriale finitamente generato individuandone una base (per esempio mediante l'algoritmo di eliminazione) e contandone il numero di elementi, oppure individuando un isomorfismo con un opportuno spazio del tipo $\mathbb{K}^n$.

Ad esempio, abbiamo ovviamente $\mathtt{dim}_{\mathbb{K}}(\mathbb{K}^n) = n$, perchè la sua base canonica $\mathfrak{B} = \{e_1, \ldots, e_n\}$ consta esattamente di n vettori. D'altra parte, il fatto che $\mathbb{K}^n$ è isomorfo a sè stesso dà un'altra giustificazione della stessa cosa.

Esempio 22.3. Lo spazio vettoriale $M_2(\mathbb{K})$ delle matrici di tipo 2x2 a coefficienti in $\mathbb{K}$ ha dimensione 4, in quanto, come si è già osservato, esso ammette la base costituita dalle 4 matrici:
$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Osservazione 22.4. Osserviamo che, se è assegnato un isomorfismo $F : \mathbb{K}^n \to V$, $n > 0$, allora esso determina in modo naturale una base di V, ottenuta "trasportando in V" la base canonica: infatti i vettori
$$F(e_1), \ldots, F(e_n)$$
costituiscono automaticamente una base di V. Si lascia la verifica come esercizio. Nel caso ad esempio dello spazio $M_2(\mathbb{K})$, utilizzando l'isomorfismo
$$F(a, b, c, d) := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
si ottiene in tal modo la base E_1, E_2, E_3, E_4 discussa sopra.

La dimensione di uno spazio vettoriale si caratterizza ulteriormente mediante le nozioni di indipendenza lineare e di sistema di generatori, come segue:

Teorema 22.5. *Sia V uno spazio vettoriale finitamente generato di dimensione $n \geq 1$.*

(1) *La dimensione n è il massimo numero di vettori linearmente indipendenti; ciò significa che esistono n vettori indipendenti e inoltre se $w_1, \dots, w_m$ sono vettori linearmente indipendenti, allora $m \leq n$.*

(2) *La dimensione n è il numero minimo di vettori che occorrono per generare V; cioè esistono n vettori che generano V e inoltre se $w_1, \dots, w_m$ è un sistema di generatori qualsiasi, allora $m \geq n$.*

DIMOSTRAZIONE: Proviamo (1). Fissiamo una base $\mathfrak{B} = \{v_1, \dots, v_n\}$ di V. Essa è costituita da n vettori, che sono linearmente indipendenti. Sia poi $A = \{w_1, \dots, w_m\}$ un insieme libero. Allora l'applicazione lineare $\varphi_A : \mathbb{K}^m \to V$ studiata nel paragrafo 19 è ingettiva. Componendola con un assegnato isomorfismo $F : V \to \mathbb{K}^n$ si ottiene un'applicazione lineare ingettiva

$$F \circ \varphi_A : \mathbb{K}^m \to \mathbb{K}^n.$$

In base al Teorema 16.3 dev'essere $m \leq n$.

Per quanto riguarda (2), i vettori di $\mathfrak{B}$ sono n e generano V. Consideriamo un altro sistema di generatori $w_1, \dots, w_m$. Il Teorema 21.3 garantisce che da tale sistema si può estrarre una base $w_{i_1}, \dots, w_{i_k}$, $k \leq m$. Poichè qualunque base è costituita da n vettori dev'essere $k = n$, e pertanto $m \geq n$. $\square$

Corollario 22.6. *Sia V uno spazio vettoriale finitamente generato di dimensione $n \geq 1$. Si consideri una sequenza di n vettori $v_1, \dots, v_n$.*

(1) *Se $v_1, \dots, v_n$ sono indipendenti, allora $\{v_1, \dots, v_n\}$ è base di V.*

(2) *Se $v_1, \dots, v_n$ costituiscono un sistema di generatori, allora $\{v_1, \dots, v_n\}$ è base di V.*

DIMOSTRAZIONE: (1) Per assurdo, si assuma che $v_1, \dots, v_n$ siano indipendenti e che $V \neq L(v_1, \dots, v_n)$. Allora esisterebbe almeno un vettore v_{n+1} di V tale che $v_{n+1} \notin L(v_1, \dots, v_n)$. Per la Proposizione 20.2, seguirebbe che $v_1, \dots, v_n, v_{n+1}$ sono linearmente indipendenti, contro la (1) del Teorema precedente.

(2) Supponiamo per assurdo che $v_1, \dots, v_n$ generino V e siano linearmente dipendenti. Allora uno di tali vettori v_j sarebbe combinazione lineare degli altri; avremmo quindi

$$V = L(v_1, \dots, v_n) \subset L(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n)$$

contro la (2) del Teorema precedente. $\square$

ESEMPIO 22.7. Il lettore verifichi ad esempio che $\mathfrak{B} = \{(1, 2, 0), (1, -1, 0), (0, 0, 1)\}$ è una base di $\mathbb{R}^3$. Poichè tale spazio ha dimensione 3, basta controllare che i vettori in questione sono linearmente indipedenti (oppure che costituiscono un sistema di generatori).

Osservazione 22.8. Dati due $\mathbb{K}$-spazi vettoriali finitamente generati, allora essi hanno la stessa dimensione se e solo se isomorfi.

La dimostrazione è un semplice esercizio.

Un'altra applicazione del Teorema 22.5 è la seguente, che caratterizza gli spazi vettoriali che non hanno dimensione finita, mediante la possibilità di poter costruire sequenze libere di vettori di lunghezza arbitraria:

Proposizione 22.9. *Sia V un $\mathbb{K}$-spazio vettoriali. Le seguenti proprietà sono equivalenti:*

a) V non è finitamente generato.

b) Per ogni intero $n \geq 1$, esistono n vettori linearmente indipendenti $v_1, \ldots, v_n$ in V.

Dimostrazione: a) $\Rightarrow$ *b*) Assumendo che V non sia finitamente generato, proviamo la proprietà b) per induzione su n. Poichè in V esiste almeno un vettore $v_1 \neq 0$ (altrimenti si avrebbe $V = \{0_V\} = L(0_V)$, contro l'ipotesi), la b) è vera per $n = 1$. Supposta ora la b) vera per $n \geq 1$, abbiamo che in V esistono n vettori $v_1, \ldots, v_n$ linearmente indipendenti. Ora, siccome V non è finitamente generato, dev'essere $V \neq L(v_1, \ldots, v_n)$. Scelto un vettore v_{n+1} di V tale che $v_{n+1} \notin L(v_1, \ldots, v_n)$, abbiamo che $v_1, \ldots, v_n, v_{n+1}$ sono indipendenti, in forza sempre della Proposizione 20.2. Pertanto la proprietà b) è vera anche per $n + 1$.

b) $\Rightarrow$ a). Assumendo b), certamente $V \neq \{0_V\}$. Ammettiamo che V sia finitamente generato. Posto $n = \mathtt{dim}(V)$, stante l'ipotesi esisterebbe almeno una sequenza $v_1, \ldots, v_{n+1}$ costituita da $n + 1$ vettori linearmente indipendenti, in contraddizione con la 1) del Teorema 22.5. $\square$

23. Sottospazi di spazi di dimensione finita

Teorema 23.1. *Sia V uno spazio vettoriale finitamente generato sul campo $\mathbb{K}$e sia W un sottospazio di V. Allora:*

1) *W è finitamente generato.*
2) *Si ha $\mathtt{dim}_{\mathbb{K}}(W) \leq \mathtt{dim}_{\mathbb{K}}(V)$ ed inoltre*

$$\mathtt{dim}_{\mathbb{K}}(W) = \mathtt{dim}_{\mathbb{K}}(V) \iff W = V.$$

Dimostrazione: 1) Ragioniamo per assurdo, assumendo che W non sia finitamente generato. Allora, applicando a W la Prop. 22.9, per ogni $n \geq 1$ sarebbe possibile individuare una sequenza $w_1, \ldots, w_n$ di n vettori di W linearmente indipendenti, che è anche una sequenza di vettori indipendenti di V; dunque per la stessa Prop. 22.9, anche V non sarebbe finitamente generato, contro l'ipotesi.

2) Supponiamo che W abbia dimensione k. Possiamo supporre $k > 0$. Fissata una base $\mathfrak{B}' = \{w_1, \ldots, w_k\}$ di W, i vettori che la costituiscono sono linearmente indipendenti. Quindi $k \leq \mathtt{dim}_{\mathbb{K}}(V)$ per il Teorema 22.5. Se $k = \mathtt{dim}_{\mathbb{K}}(V)$, i vettori di $\mathfrak{B}'$ sono n formano necessariamente anche una base di V in virtù del Corollario 22.6; pertanto:

$$W = L(w_1, \ldots, w_n) = V.$$

□

24. Il teorema di completamento

Il risultato seguente mostra come si possano costruire basi di uno spazio vettoriale finitamente generato, a partire da un insieme libero qualsiasi.

Teorema 24.1. *(Completamento di un insieme libero)*
Sia V uno spazio vettoriale sul campo $\mathbb{K}$, di dimensione $n > 0$. Siano assegnati dei vettori linearmente indipendenti

$$u_1, \ldots, u_k, \quad k \geq 1.$$

Sia inoltre $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base di V. Allora esistono $n-k$ vettori $v_{j_1}, \ldots, v_{j_{n-k}}$ della base $\mathfrak{B}$ tali che

$$\mathfrak{B}' = \{u_1, \ldots, u_k, v_{j_1}, \ldots, v_{j_{n-k}}\}$$

è una base di V.

Dimostrazione: Consideriamo la sequenza costituita da tutti i vettori

$$u_1, \ldots, u_k, v_1, \ldots, v_n.$$

Tale sequenza è ancora un sistema di generatori, da cui si può estrarre una base $\mathfrak{B}'$ mediante l'algoritmo di eliminazione. Poichè $u_1, \ldots, u_k$ sono indipendenti, ciascuno di essi non è eliminabile. Dunque la base $\mathfrak{B}'$ è del tipo richiesto. □

25. Somme dirette

Definizione 25.1. Dati due sottospazi vettoriali U e W di uno spazio vettoriale V, la somma $U + W$ verrà detta *diretta* se

$$U \cap W = \{0_V\}. \tag{14}$$

In tal caso, in luogo di $U + W$ si scrive $U \oplus W$.

Sussiste la seguente caratterizzazione della condizione (14).

Proposizione 25.2. *Dati i sottospazi U e W di V, le seguenti proprietà sono equivalenti:*

a) *La somma $U + W$ è diretta.*

b) *Ogni vettore $v \in U + W$ si scrive in modo unico nella forma $v = u + w$ con $u \in U$ e $w \in W$.*

DIMOSTRAZIONE: *a)* $\Rightarrow$ *b)*. Stante la definizione di $U + W$, assumendo a) è chiaro che ogni $v \in U + W$ si può scrivere come enunciato in b); supponiamo che ciò possa farsi in due modi diversi: $v = u + w = u' + w'$, con ovvio significato dei simboli. Allora $u - u' = w' - w$ ed il vettore con cui coincidono ambo i membri appartiene sia ad U che a W in virtù delle proprietà dei sottospazi. Pertanto, per la a), si ha $u - u' = 0_V = w' - w$ e di qui la b).

b) $\Rightarrow$ *a)*. Sia $v \in U \cap W$; poichè

$$v = v + 0_V = 0_V + v,$$

applicando la *b)*, dal confronto di tali rappresentazioni di v si ottiene che necessariamente $v = 0_V$. $\square$

Un'utile proprietà in merito al completamento di un insieme libero ad una base è la seguente:

Corollario 25.3. *Siano assegnata una sequenza di vettori linearmente indipendenti*

$$u_1, \ldots, u_k, \quad k \geq 1.$$

dello spazio vettoriale V e siano $u_{k+1}, \ldots, u_n$ altri $n - k$ vettori che completano tale sequenza ad una base $\mathfrak{B} = \{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$. Allora risulta:

$$V = L(u_1, \ldots, u_k) \oplus L(u_{k+1}, \ldots, u_n).$$

DIMOSTRAZIONE: Occorre provare che la somma di tali sottospazi è diretta. Sia $v \in L(u_1, \ldots, u_k) \cap L(u_{k+1}, \ldots, u_n)$. Allora v si può scrivere in due modi:

$$v = \lambda_1 u_1 + \cdots + \lambda_k u_k,\ v = \lambda_{k+1} u_{k+1} + \cdots + \lambda_n u_n,$$

per opportuni scalari λ_i, da cui:

$$\lambda_1 u_1 + \cdots + \lambda_k u_k - \lambda_{k+1} u_{k+1} + \cdots - \lambda_n u_n = 0$$

e quindi $\lambda_1 = \cdots = \lambda_n = 0$ stante la lineare indipendenza dei vettori della base; in particolare $v = 0$. Infine, è chiaro che

$$V \subset L(u_1, \ldots, u_k) + L(u_{k+1}, \ldots, u_n)$$

perchè $V = L(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ e tutti i vettori u_i appartengono al sottospazio somma di $L(u_1, \ldots, u_k)$ e $L(u_{k+1}, \ldots, u_n)$.

$\square$

Osservazione 25.4. Applicando questa Proposizione, otteniamo che, data una base qualsiasi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ di uno spazio vettoriale V e un indice $1 \leq k \leq n$, allora si ha sempre:

$$V = L(v_1, \ldots, v_k) \oplus L(v_{k+1}, \ldots, v_n).$$

26. Calcolo della dimensione di alcuni spazi notevoli

Nel seguito calcoliamo la dimensione ed una base di alcuni spazi vettoriali interessanti.

Esempio 26.1. Fissato $n \geq 0$, l'insieme

$$\mathbb{R}_n[x] = \{p : \mathbb{R} \to \mathbb{R} |\, p = p(x) \text{ è un polinomio di grado } \leq n\},$$

è un sottospazio vettoriale dello spazio $\mathbb{R}^{\mathbb{R}}$ di tutte le funzioni $p : \mathbb{R} \to \mathbb{R}$.

Risulta che tale spazio è finitamente generato e ha dimensione $n+1$. Infatti ogni polinomio p di grado al massimo n è del tipo

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

dove gli a_i sono costanti. Ciò significa che p è combinazione lineare dei polinomi

$$1, x, x^2, \dots, x^n \tag{$*$}$$

con coefficienti esattamente $a_0, \dots, a_n$; tali polinomi costituiscono quindi un sistema generatori di $\mathbb{R}_n[x]$. I polinomi (*) costituiscono di fatto una base di $\mathbb{R}_n[x]$; infatti data una loro combinazione lineare nulla:

$$a_0 + a_1 x + \cdots + a_n x^n = 0,$$

dove 0 è il polinomio nullo (la funzione costante di valore 0), allora il principio di identità dei polinomi garantisce che necessariamente tutti i coefficienti a_i devono essere uguali a zero. Dunque gli $n+1$ vettori (*) formano una base del nostro spazio $\mathbb{R}_n[x]$, che pertanto ha dimensione $n+1$.

Esempio 26.2. Lo spazio vettoriale $M_{m,n}(\mathbb{K})$ delle matrici di tipo $m \times n$ a coefficienti in un campo $\mathbb{K}$ è isomorfo a $\mathbb{K}^{mn}$ e quindi ha dimensione mn. Un isomorfismo naturale si ottiene associando a ogni matrice A un vettore con mn elementi, costituito da tutti gli ingressi a^i_j della matrice, scritti in ordine lessicografico, ovvero scansionando le righe dalla prima all'ultima e disponendo gli elementi per ordine crescente rispetto alla colonna di appartenenza; si definisce così la bigezione:

$$F : M_{m,n}(\mathbb{K}) \to \mathbb{K}^{mn}$$

definita da:

$$F(A) = (a^1_1, \dots, a^1_n, a^2_1, \dots, a^2_n, \dots, a^m_1, \dots, a^m_n).$$

Possiamo anche scrivere (con piccolo abuso di notazione):

$$F(A) = (A^{(1)}, \dots, A^{(m)}).$$

Date due matrici $A, B \in M_{m,n}(\mathbb{K})$ e scalari λ, μ, poichè la riga i-ma della matrice $\lambda A + \mu B$ è $\lambda A^{(1)} + \mu B^{(1)}$, si ottiene:

$$F(\lambda A + \mu B) = (\lambda A^{(1)} + \mu B^{(1)}, \dots, \lambda A^{(m)} + \mu B^{(m)}) = \lambda(A^{(1)}, \dots, A^{(m)}) + \mu(B^{(1)}, \dots, B^{(m)})$$

il che mostra che F è lineare.

Una base di tale spazio è

$$\mathfrak{B} = \{E_{ij} : 1 \leq i, j \leq n\}$$

dove denotiamo con E_{ij} la matrice con m righe e n colonne, i cui elementi sono tutti nulli, eccetto l'elemento di posto (i, j) che è uguale a 1. Tale base si ottiene applicando ai vettori della base canonica $\{e_i\}$ di $\mathbb{K}^{mn}$ l'inverso $F^{-1} : \mathbb{K}^{mn} \to M_{mn}(\mathbb{K})$ dell'isomorfismo F, trasformando in modo naturale ogni vettore e_i in una matrice. Così, ad esempio:

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

è una base di $M_{2,3}(\mathbb{K})$.

Esempio 26.3. Dato un campo $\mathbb{K}$, calcoliamo la dimensione di $Hom(\mathbb{K}^n, \mathbb{K})$, ovvero lo spazio di tutte le applicazioni lineari $F : \mathbb{K}^n \to \mathbb{K}$.

Sappiamo che ogni tale F è del tipo:

$$F(x) = a_1x_1 + \cdots + a_nx^n \qquad (*)$$

dove gli scalari $a_1, \ldots, a_n$ sono univocamente determinati, essendo $a_i = F(e_i)$, dove come al solito $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{K}^n$. Notiamo ora che (*) si può interpretare in termini di combinazione lineare, come segue:

$$F = a_1p_1 + \cdots + a_np_n$$

dove per ogni i, $p_i : \mathbb{K}^n \to \mathbb{K}$ è la proiezione

$$p_i(x) := x_i,$$

che è anch'essa un vettore di $Hom(\mathbb{K}^n, \mathbb{K})$. Ciò mostra che tali proiezioni $p_1, \ldots, p_n$ formano un sistema di generatori di $Hom(\mathbb{K}^n, \mathbb{K})$.

Risulta poi che $p_1, \ldots, p_n$ sono linearmente indipendenti; infatti, data una combinazione lineare nulla:

$$a_1p_1 + \cdots + a_np_n = 0,$$

allora la funzione lineare al primo membro è quella costante di valore 0, ma ciò implica direttamente che $a_i = 0$ per ogni i. Conclusione: $\mathfrak{B} = \{p_1, \ldots, p_n\}$ è una base di $Hom(\mathbb{K}^n, \mathbb{K})$ e dunque:

$$\mathtt{dim}_{\mathbb{K}} Hom(\mathbb{K}^n, \mathbb{K}) = n.$$

Ad esempio, una base di $Hom(\mathbb{R}^3, \mathbb{R})$ è costituita dalle 3 proiezioni:

$$p_1(x_1, x_2, x_3) = x_1,\ p_2(x_1, x_2, x_3) = x_2,\ p_3(x_1, x_2, x_3) = x_3.$$

Esempio 26.4. Più in generale, risulta che lo spazio $Hom(\mathbb{K}^n, \mathbb{K}^m)$ ha dimensione mn. Risulta che esso è isomorfo a $M_{m,n}(\mathbb{K})$ mediante l'applicazione

$$\Phi : A \in M_{m,n}(\mathbb{K}) \mapsto L_A \in Hom(\mathbb{K}^n, \mathbb{K}^m).$$

Sappiamo infatti che questa funzione è bigettiva in virtù del Teorema 15.6; essa è lineare grazie alla Prop. 15.9; infatti:

$$\Phi(A + B) = L_{A+B} = L_A + L_B = \Phi(A) + \Phi(B),$$

mentre è immediato che

$$\Phi(\lambda A) = L_{\lambda A} = \lambda L_A = \lambda\Phi(A).$$

Concludiamo questo paragrafo trattando il caso del prodotto diretto di due spazi vettoriali.

Proposizione 26.5. *Siano U e W due $\mathbb{K}$-spazi vettoriali di dimensione finita. Allora*

$$\mathtt{dim}(U \times W) = \mathtt{dim}(U) + \mathtt{dim}(W).$$

Dimostrazione: Possiamo limitarci a considerare il caso in cui i due spazi sono entrambi non banali: se ad es. $U = \{0_U\}$, allora è evidente che $U \times W \cong W$ tramite l'isomorfismo:

$$(0_U, w) \mapsto w.$$

Cominciamo con lo studiare il caso in cui $U = \mathbb{K}^n$ e $W = \mathbb{K}^m$ con m e n interi positivi. In tal caso risulta:

$$\mathbb{K}^n \times \mathbb{K}^m \cong \mathbb{K}^{n+m}.$$

Per giustificare ciò, è sufficiente considerare l'applicazione

$$((x_1, \ldots, x_n), (y_1, \ldots, y_m)) \in \mathbb{K}^n \times \mathbb{K}^m \mapsto (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{K}^{n+m}.$$

Si lascia al lettore la semplice verifica che si tratta di un isomorfismo.

Venendo ora al caso generale, posto $n = \mathtt{dim}(U)$ e $m = \mathtt{dim}(W)$, siano assegnati due isomorfismi $F : U \to \mathbb{K}^n$ e $G : W \to \mathbb{K}^m$. A partire da questi, possiamo definire in modo standard un'applicazione:

$$F \times G : U \times W \to \mathbb{K}^n \times \mathbb{K}^m,$$

ponendo

$$(F \times G)(u, w) := (F(u), G(w)) \quad \forall (u, w) \in U \times W.$$

È un semplice esercizio verificare che si tratta di una bigezione, stante la bigettività di entrambe le funzioni coinvolte. Tale applicazione è lineare; infatti, dati due vettori

(u, w) e (u', w') di $U \times W$, allora:

$$\begin{aligned}(F \times G)(\lambda(u,w) + \mu(u',w')) &= (F \times G)(\lambda u + \mu u', \lambda w + \mu w')\\ &= (F(\lambda u + \mu u'), G(\lambda w + \mu w'))\\ &= (\lambda F(u) + \mu F(u'), \lambda G(w) + \mu G(w'))\\ &= \lambda(F(u), G(w)) + \mu(F(u'), G(w'))\\ &= \lambda(F \times G)(u,w) + \mu(F \times G)(w,w').\end{aligned}$$

Resta quindi stabilito che:

$$U \times V \cong \mathbb{K}^n \times \mathbb{K}^m,$$

da cui la conclusione. □

27. Il Teorema del rango

Proposizione 27.1. *Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali.*

a) Data una sequenza qualsiasi di vettori $v_1, \ldots, v_n$ di V si ha:

$$F(L(v_1, \ldots, v_n)) = L(F(v_1), \ldots, F(v_n)).$$

b) Dati due sottospazi vettoriali U e W di V, si ha: $F(U + W) = F(U) + F(W)$.

Dimostrazione: a) Ciò è conseguenza immediata del fatto che F trasforma combinazioni lineari dei vettori v_i in combinazioni lineari delle loro immagini $F(v_i)$, cfr. Prop. 12.5:

$$\begin{aligned}F(L(v_1, \ldots, v_n)) &= F(\{\lambda_1 v_1 + \cdots + \lambda_n v_n : \lambda_1, \ldots, \lambda_n \in \mathbb{K}\}) =\\ &= \{F(\lambda_1 v_1 + \cdots + \lambda_n v_n) : \lambda_1, \ldots, \lambda_n \in \mathbb{K}\}) =\\ &= \{\lambda_1 F(v_1) + \cdots + \lambda_n F(v_n) : \lambda_1, \ldots, \lambda_n \in \mathbb{K}\} =\\ &= L(F(v_1), \ldots, F(v_n)).\end{aligned}$$

b) Basta applicare la definizione di spazio somma:

$$\begin{aligned}F(U + W) &= F(\{u + w | \, u \in U, w \in W\}) = \{F(u + w) | \, u \in U, w \in W\} =\\ &= \{F(u) + F(w) | u \in U, w \in W\} = F(U) + F(W).\end{aligned}$$

□

Mostriamo ora che le applicazioni lineari conservano la proprietà degli spazi a cui sono applicati di essere finitamente generati:

Proposizione 27.2. *Sia $F : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali e sia W un sottospazio vettoriale di V. Allora, se W è finitamente generato, tale è $F(W)$ e si ha*

$$\mathtt{dim}_{\mathbb{K}} F(W) \leq \mathtt{dim}_{\mathbb{K}} W.$$

Se F è un monomorfismo, allora risulta:

$$\mathtt{dim}_{\mathbb{K}} F(W) = \mathtt{dim}_{\mathbb{K}} W.$$

Dimostrazione: Assunto che W sia finitamente generato, se $W = \{0_V\}$, allora $F(W) = \{0_V\}$ e quindi la tesi. Supponiamo quindi $\mathtt{dim}_{\mathbb{K}} W = n > 0$ e fissiamo una base $\mathfrak{B} = \{w_1, \ldots, w_n\}$ di W. Allora abbiamo:

$$F(W) = F(L(w_1, \ldots, w_n)) = L(F(w_1), \ldots, F(w_n))$$

per cui anche $F(W)$ è finitamente generato; siccome la dimensione di $F(W)$ è il numero minimo di vettori costituenti un sistema di generatori, deve aversi

$$\mathtt{dim}_{\mathbb{K}} F(W) \leq n.$$

Se F è un monomorfismo, consideriamo la funzione

$$F|_W : W \to F(W),$$

ridotta della restrizione di F a W. Essa è ancora lineare, e per costruzione surgettiva, ma anche ingettiva perchè lo è F. Dunque è un isomorfismo e pertanto W e $F(W)$ in tal caso hanno la stessa dimensione. □

Il risultato seguente è di importanza basilare in tutte le applicazioni dell'algebra lineare.

Teorema 27.3. *(Teorema del rango, o della dimensione, o rango-nullità). Siano V e V' spazi vettoriali sullo stesso campo $\mathbb{K}$. Si assuma che V sia finitamente generato. Sia*

$$F : V \to V'$$

un'applicazione lineare. Allora risulta:

$$\mathtt{dim}(V) = \mathtt{dim}(Ker(F)) + \mathtt{dim}(Im(F)). \tag{15}$$

Dimostrazione: Esaminiamo dapprima il caso in cui F è ingettiva. Allora, in base alla Prop. 27.2 sappiamo che $\mathtt{dim}(F(V)) = \mathtt{dim}(V)$. D'altra parte per l'ingettività, $Ker(F) = \{0_V\}$ e quindi $\mathtt{dim}(Ker(F)) = 0$ e pertanto la (15) è vera in questo caso.

Supponiamo ora che F non sia ingettiva e poniamo $k := \mathtt{dim}(Ker(F))$; dunque $k > 0$. Fissiamo una base $\mathfrak{B} = \{u_1, \ldots, u_k\}$ di $Ker(F)$. Posto $n = \mathtt{dim}(V)$, applicando il Teorema di completamento, sappiamo che esistono $n-k$ vettori $u_{k+1}, \ldots, u_n$ che completano $\mathfrak{B}$ ad una base

$$\mathfrak{B}' = \{u_1, \ldots, u_n\}$$

di V. Consideriamo quindi il sottospazio di V:

$$W = L(u_{k+1}, \ldots, u_n).$$

Chiaramente, tale sottospazio ha dimensione $n - k$ perchè i vettori $u_{k+1}, \ldots, u_n$ ne costituiscono una base, in quanto sono linearmente indipendenti. Sappiamo inoltre dal Cor. 25.3 che

$$V = L(u_1, \ldots, u_k) \oplus L(u_{k+1}, \ldots, u_n)$$

cioè che:

$$V = Ker(F) \oplus W.$$

Consideriamo quindi la restrizione di F a W:

$$F|_W : W \to V'.$$

Questa funzione, che è ancora lineare, è ingettiva, perchè, se $w \in W$ è un vettore appartenente a $Ker(F|_W)$, si ha $F(w) = 0_{V'}$, ma allora $w \in W \cap Ker(F) = \{0_V\}$, da cui $w = 0$. Pertanto, sempre per la Prop. 27.2, $F(W)$ ha la stessa dimensione di W, ovvero ha dimensione $n - k$. Ma $F(W) = Im(F)$; infatti, applicando b) della Prop. 27.1:

$$Im(F) = F(V) = F(Ker(F)+W) = F(Ker(F))+F(W) = \{0_{V'}\}+F(W) = F(W).$$

In conclusione

$$\mathtt{dim}(Im(F)) = \mathtt{dim}(F(W)) = n - k = n - \mathtt{dim}(Ker(F))$$

e la (15) è provata. $\square$

28. Il teorema di Grassmann

Discutiamo ora il caso rilevante della somma di due sottospazi, provando un altro risultato fondamentale:

Teorema 28.1. *(Formula di Grassmann) Sia V uno spazio vettoriale di dimensione finita. Siano U e W due sottospazi di V. Allora:*

$$\mathtt{dim}(U + W) = \mathtt{dim}(U) + \mathtt{dim}(W) - \mathtt{dim}(U \cap W). \tag{16}$$

Dimostrazione: Consideriamo l'applicazione

$$F : U \times W \to U + W$$

definita da:

$$F((u, w)) := u + w.$$

Per la definizione stessa dello spazio somma $U + W$, F è surgettiva. Verifichiamo che F è anche lineare: si considerino due vettori (u, w) e (u', w') di $U \times W$; allora

$$\begin{aligned} &F((u, w) + (u', w')) = F((u + u', v + v')) \\ &= (u + u') + (v + v') = (u + v) + (u' + v') \\ &= F((u, w)) + F((u', w')). \end{aligned}$$

Per ogni $\lambda \in \mathbb{K}$ abbiamo anche

$$F(\lambda(u, w)) = F((\lambda u, \lambda w)) = \lambda u + \lambda w = \lambda(u + w) = \lambda F(u, w).$$

Determiniamo ora il nucleo di F. Per ogni $(u, w) \in U \times W$ risulta

$$F(u, w) = 0 \iff u + w = 0 \iff w = -u;$$

d'altra parte se $w = -u$ allora ambo i membri sono vettori che devono appartenere sia a U che a W; se ne deduce che:

$$Ker(F) = \{(u, -u) |\, u \in U \cap W\}. \tag{17}$$

Da qui segue facilmente che:

$$Ker(F) \cong U \cap W.$$

Infatti, un isomorfismo tra questi spazi è dato semplicemente dall'applicazione

$$G : u \in U \cap W \mapsto (u, -u) \in Ker(F),$$

che è manifestamente bigettiva. Anche la verifica della linearità di questa funzione non presenta difficoltà:

$$G(\lambda u + \mu u') = (\lambda u + \mu u', -\lambda u - \mu u') = (\lambda u, -\lambda u) + (\mu u', -\mu u')$$
$$= \lambda(u, -u) + \mu(u', -u') = \lambda G(u) + \mu G(u').$$

Applicando infine a F il Teorema del rango, otteniamo:

$$\mathtt{dim}(U \times W) = \mathtt{dim}(Ker(F)) + \mathtt{dim}(Im(F)) = \mathtt{dim}(U \cap W) + \mathtt{dim}(U + W),$$

da cui segue la formula di Grassmann, avendosi $\mathtt{dim}(U \times W) = \mathtt{dim}(U) + \mathtt{dim}(W)$ per la Prop. 26.5. □

Corollario 28.2. *Sia V uno spazio vettoriale di dimensione finita e siano U, W sottospazi di V. Allora la somma $U + W$ è diretta se e solo se si ha:*

$$\mathtt{dim}(U + W) = \mathtt{dim}(U) + \mathtt{dim}(W).$$

29. Componenti di un vettore rispetto ad una base

Consideriamo un $\mathbb{K}$-spazio vettoriale V. Sussiste la seguente caratterizzazione delle basi di V:

Proposizione 29.1. *Sia $v_1, \ldots, v_n$ una sequenza di vettori di V. Allora sono proprietà equivalenti:*

a) $\mathfrak{B} = \{v_1, \ldots, v_n\}$ è una base di V.

b) Ogni vettore $v \in V$ si scrive in modo unico come combinazione lineare dei vettori $v_1, \ldots, v_n$.

Ha senso allora dare la seguente:

Definizione 29.2. Si chiameranno **componenti** (o coordinate) di un vettore $v \in V$ **rispetto alla base $\mathfrak{B}$** gli scalari $x_1, \ldots, x_n$ tali che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Chiameremo anche *vettore delle componenti* di v rispetto a $\mathfrak{B}$ il vettore $x = (x_1, \ldots, x_n)$ di $\mathbb{K}^n$.

Tale vettore viene spesso identificato anche con il corrispondente vettore colonna

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Naturalmente la nozione introdotta dipende strettamente dalla scelta della base. Studieremo in seguito come cambiano le componenti di un vettore al variare della base. Giova osservare anche che l'ordine assegnato ai vettori della base $\mathfrak{B}$ è essenziale nella definizione delle componenti.

L'equivalenza tra le condizioni a) e b) nell'enunciato precedente è conseguenza immediata del fatto che la a) è caratterizzata mediante l'applicazione lineare

$$\varphi_{\mathfrak{B}} : \mathbb{K}^n \to V,$$

studiata nel §17 e definita da

$$\varphi_{\mathfrak{B}}(x_1, \dots, x_n) = x_1 v_1 + \dots + x_n v_n,$$

ed equivale al fatto che questa sia un isomorfismo. Ma la bigettività di $\varphi_{\mathfrak{B}}$ si traduce nel fatto che:

$$\forall v \in V \;\; \exists | (x_1, \dots, x_n) \in \mathbb{K}^n : \quad v = \varphi_{\mathfrak{B}}(x_1, \dots, x_n) = x_1 v_1 + \dots + x_n v_n$$

che è proprio la condizione b) enunciata sopra.

Nel seguito, data una base $\mathfrak{B}$, denoteremo con

$$\psi_{\mathfrak{B}} : V \to \mathbb{K}^n$$

la funzione che associa ad ogni vettore v la n-pla $(x_1, \dots, x_n)$ delle sue componenti rispetto alla base in questione.

Per definizione, questa funzione altri non è che l'isomorfismo inverso di $\varphi_{\mathfrak{B}}$; in simboli:

$$\psi_{\mathfrak{B}} = \varphi_{\mathfrak{B}}^{-1}.$$

Notiamo che la linearità di questa funzione garantisce che:

- *il vettore delle componenti della somma $v + v'$ di due vettori v e v' coincide con la somma $x + x'$ dei vettori delle componenti di v e di v';*

- *per ogni scalare $\lambda \in \mathbb{K}$, dato un vettore v con vettore di componenti x, il vettore delle componenti di λv è λx.*

Esempio 29.3. Nel caso in cui $V = \mathbb{K}^n$ e scegliamo la base canonica $\mathfrak{B}_0 = \{e_1, \dots, e_n\}$, allora le componenti di ogni $x = (x_1, \dots, x_n)$ sono esattamente $(x_1, \dots, x_n)$. Ciò segue dalla identità fondamentale (5).

Dunque in tal caso $\psi_{\mathfrak{B}_0}$ è semplicemente l'isomorfismo identico: $\psi_{\mathfrak{B}_0} = Id_{\mathbb{K}^n}$.

30. Matrice associata ad un insieme di vettori

Sia assegnato uno spazio vettoriale V di dimensione $n > 0$ e si fissi una sua base $\mathfrak{B} = \{v_1, \dots, v_n\}$.

Definizione 30.1. Sia $u_1, \dots, u_k$ una successione finita di vettori arbitrari di V, $k \geq 1$. Si chiamerà **matrice associata** a tali vettori **rispetto alla base $\mathfrak{B}$**, la matrice con n righe e k colonne, la cui colonna j-ma è il vettore delle componenti di u_j rispetto a $\mathfrak{B}$, ovvero è la matrice

$$(\psi_{\mathfrak{B}}(u_1) \quad \psi_{\mathfrak{B}}(u_2) \quad \cdots \quad \psi_{\mathfrak{B}}(u_k)). \tag{18}$$

L'utilità della nozione di matrice associata ad un insieme di vettori è stabilita dal risultato seguente, che riconduce il calcolo della dimensione di un sottospazio di V al calcolo del rango di una matrice. Ricordiamo che il rango di una matrice $A \in M_{m,n}(\mathbb{K})$ è il massimo numero di righe (o equivalentemente, colonne) linearmente indipendenti. Se $r = rg(A)$, allora

$$r = \mathtt{dim}(L(A^{(1)}, \dots, A^{(m)})).$$

Possiamo supporre $r > 0$, altrimenti $A = 0$ e il risultato è ovvio. Denotata con s la dimensione dello spazio vettoriale generato dalla righe di A, e data una base $A^{(i_1)}, \dots, A^{(i_s)}$ di tale spazio, allora tali righe sono linearmente indipendenti, per cui $s \leq r$. D'altra parte, sappiamo che s è il massimo numero di vettori indipendenti nello stesso spazio, per cui dev'essere anche $r \leq s$. Analogamente, coinvolgendo le colonne:

$$r = \mathtt{dim}(L(A_{(1)}, \dots, A_{(n)})).$$

Ciò premesso, si ha:

Teorema 30.2. *Sia $u_1, \dots, u_k$ una successione finita di vettori di V, $k \geq 1$. Sia $\mathfrak{B}$ una base di V e sia A la matrice associata a questa sequenza di vettori rispetto alla base $\mathfrak{B}$. Allora*

$$\mathtt{dim}(L(u_1, \dots, u_k)) = rg(A).$$

Dimostrazione: Poichè $\psi_{\mathfrak{B}}$ è un isomorfismo, sappiamo che ogni sottospazio di V viene trasformato da $\psi_{\mathfrak{B}}$ in un sottospazio di $\mathbb{K}^n$ avente la stessa dimensione (Prop. 27.2); si ha quindi:

$$\begin{aligned} \mathtt{dim}(L(u_1, \dots, u_k)) &= \mathtt{dim}(\psi_{\mathfrak{B}}(L(u_1, \dots, u_k)) = \\ &= \mathtt{dim}\, L(\psi_{\mathfrak{B}}(u_1), \dots, \psi_{\mathfrak{B}}(u_k)) = \\ &= \mathtt{dim}\, L(A_{(1)}, \dots, A_{(k)}) = rg(A). \end{aligned}$$

□

Come applicazione, discutiamo un'altra caratterizzazione delle basi, attraverso la nozione di matrice invertibile. Ricordiamo che $GL(n, \mathbb{K})$ denota il gruppo delle matrici invertibili di ordine n. Queste matrici sono caratterizzate mediante la condizione $\det(A) \neq 0$ o, equivalentemente, $rg(A) = n$.

Corollario 30.3. *Siano V un $\mathbb{K}$-spazio vettoriale di dimensione n, $\mathfrak{B}' = \{u_1, \ldots, u_n\}$ una sequenza di n vettori, e $A \in M_n(\mathbb{K})$ la matrice associata a tali vettori rispetto ad una fissata base $\mathfrak{B}$. Allora:*

$$\mathfrak{B}' \text{ è una base di } V \iff A \in GL(n, \mathbb{K}).$$

Dimostrazione: Infatti, sappiamo che la sequenza di n vettori in questione forma una base se e solo se è un sistema di generatori, il che accade se e solo se il sottospazio $L(u_1, \ldots, u_n)$ coincide con V, il che accade se e solo se ha dimensione n, e ciò è vero se e solo se $rg(A) = n$ per il teorema precedente. □

31. Matrici associate ad un'applicazione lineare

Siano V e W due spazi vettoriali non banali sullo stesso campo $\mathbb{K}$, entrambi finitamente generati, aventi dimensioni n ed m rispettivamente.

Definizione 31.1. Siano fissate due basi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C}$ di V e di W rispettivamente. Sia $F : V \to W$ un'applicazione lineare. Si chiama **matrice associata ad F rispetto alle basi $\mathfrak{B}$ e $\mathfrak{C}$** la matrice associata alla successione di vettori $F(v_1), \ldots, F(v_n)$ di W rispetto alla base $\mathfrak{C}$.

Questa matrice si denoterà con il simbolo

$$M^{\mathfrak{C}}_{\mathfrak{B}}(F)$$

oppure talvolta con

$$M_{\mathfrak{B},\mathfrak{C}}(F).$$

Dunque, per definizione, si tratta della matrice di tipo $m \times n$ data da:

$$M^{\mathfrak{C}}_{\mathfrak{B}}(F) := (\psi_{\mathfrak{C}}(F(v_1)) \quad \psi_{\mathfrak{C}}(F(v_2)) \quad \cdots \quad \psi_{\mathfrak{C}}(F(v_n))). \tag{19}$$

Qui utilizziamo ancora l'isomorfismo $\psi_{\mathfrak{C}} : W \to \mathbb{K}^m$ che associa ad ogni vettore di W il vettore delle sue componenti rispetto alla base $\mathfrak{C}$. Mettiamo in evidenza il fatto che la matrice $M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ dipende dalla scelta delle basi scelte nei due spazi.

Esempio 31.2. Consideriamo gli spazi vettoriali $\mathbb{R}_3[x]$ e $\mathbb{R}_2[x]$ costituiti dai polinomi $p : \mathbb{R} \to \mathbb{R}$ di grado rispettivamente minore o uguale a 3 e minore o uguale a 2. L'applicazione

$$F : \mathbb{R}_3[x] \to \mathbb{R}_2[x]$$

definita da

$$F(p) = p',$$

che associa a ogni polinomio $p \in \mathbb{R}_3[x]$ la sua derivata prima, è ben definita ed è lineare in virtù delle proprietà note della derivata. Determiniamo la matrice associata a F rispetto alle basi $\mathfrak{B} = \{1, x, x^2, x^3\}$ e $\mathfrak{C} = \{1, x, x^2\}$. Intanto osserviamo che, per ogni polinomio $p = a_0 + a_1x + a_2x^2$ in $\mathbb{R}_2[x]$, il vettore delle sue componenti è (a_0, a_1, a_2), cioè:

$$\psi_{\mathfrak{C}}(p) = (a_0, a_1, a_2).$$

Abbiamo

$$F(1) = 0,\ F(x) = 1,\ F(x^2) = 2x,\ F(x^3) = 3x^2$$

e pertanto la matrice associata a F è la seguente:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(F) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

L'utilità di questa nozione è subito chiarita dal risultato seguente:

Teorema 31.3. *Siano fissate due basi $\mathfrak{B}$ e $\mathfrak{C}$ di due $\mathbb{K}$-spazi vettoriali V e W. Sia $F : V \to W$ un'applicazione lineare. Allora:*

$$\mathtt{dim}(Im(F)) = rg(M_{\mathfrak{B}}^{\mathfrak{C}}(F)).$$

Dimostrazione: Posto $\mathfrak{B} = \{v_1, \ldots, v_n\}$, si ha:

$$\mathtt{dim}(Im(F)) = \mathtt{dim}\, L(F(v_1), \ldots, F(v_n)) = rg(M_{\mathfrak{B}}^{\mathfrak{C}}(F)),$$

dove l'ultima uguaglianza è una diretta applicazione del Teorema 30.2 alla sequenza di vettori $F(v_1), \ldots, F(v_n)$. $\square$

Questo risultato autorizza a chiamare **rango di un'applicazione lineare** F la dimensione di $Im(F)$; si usa anche il simbolo $rg(F)$ per denotare questa dimensione. Ciò giustifica anche la denominazione del Teorema fondamentale 27.3.

Corollario 31.4. *Data un'applicazione lineare $F : V \to W$, fissate due basi $\mathfrak{B}$ e $\mathfrak{C}$ di V e di W rispettivamente, allora:*

$$F \text{ è un epimorfismo} \iff rg(M_{\mathfrak{B}}^{\mathfrak{C}}(F)) = \mathtt{dim}\, W,$$
$$F \text{ è un monomorfismo} \iff rg(M_{\mathfrak{B}}^{\mathfrak{C}}(F)) = \mathtt{dim}\, V.$$

È un'immediata applicazione del risultato precedente per quel che concerne la prima affermazione; riguardo la seconda, basta applicare il Teorema del rango: infatti F è un monomorfismo se e solo se $\mathtt{dim}(Ker(F)) = 0$ ll che accade se e solo se $\mathtt{dim}\, V = \mathtt{dim}(Im(F))$.

Esempio 31.5. Ad esempio, se $F : \mathbb{K}^4 \to \mathbb{K}^3$ è definita da:
$$F(x_1, x_2, x_3, x_4) := (x_1 + x_2, 0, x_3 + x_4),$$
allora il rango di F si può calcolare valutando il rango della matrice A ad essa associata che, come sappiamo, è:
$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$
Pertanto $rg(F) = 2$. Un base di $Im(F)$ è ad esempio $\{F(e_1), F(e_3)\} = \{(1, 0, 0), (0, 0, 1)\}$.

32. Applicazioni lineari tra spazi vettoriali della stessa dimensione

In questo paragrafo approfondiamo il caso di un'applicazione lineare che operi tra spazi vettoriali aventi la stessa dimensione.

Teorema 32.1. *Sia $F : V \to W$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali della stessa dimensione n. Allora le seguenti proprietà sono equivalenti:*
a) *F è un monomorfismo,*
b) *F è un epimorfismo,*
c) *F è un isomorfismo.*

Dimostrazione: È sufficiente dimostrare che a) e b) sono equivalenti. Si tratta di una semplice applicazione del Teorema del rango. Infatti, tenendo conto della (15) risulta che F è un monomorfismo se e solo se $\mathtt{dim}(Ker(F)) = 0$, il che accade se e solo se $\mathtt{dim}(V) = \mathtt{dim}(Im(F))$, ovvero, stante, l'ipotesi, se e solo se $\mathtt{dim}(W) = Im(F)$ e ciò equivale a dire che $Im(F) = W$, ovvero che F è un epimorfismo. □

Come applicazione di questo risultato abbiamo la seguente caratterizzazione degli isomorfismi, come quelle applicazioni lineari che trasformano basi in basi; più precisamente:

Teorema 32.2. *Sia $F : V \to W$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali della stessa dimensione $n > 0$. Allora le seguenti proprietà sono equivalenti:*
a) *F è un isomorfismo;*
b) *Per ogni base $\mathfrak{B} = \{v_1, \ldots, v_n\}$ di V, si ha che $\{F(v_1), \ldots, F(v_n)\}$ è una base di W;*
c) *Esiste una base $\mathfrak{B} = \{v_1, \ldots, v_n\}$ di V, tale che $\{F(v_1), \ldots, F(v_n)\}$ è una base di W.*

Dimostrazione: Riguardo l'implicazione $a) \Rightarrow b)$, se vale a), F è un epimorfismo e quindi
$$L(F(v_1), \ldots, F(v_n)) = W,$$
e quindi $F(v_1), \ldots, F(v_n)$ è un sistema di generatori di W; siccome W ha dimensione n, sappiamo che ciò basta per affermare che $\{F(v_1), \ldots, F(v_n)\}$ è base di

W. L'implicazione $b) \Rightarrow c)$ è ovvia; resta da provare che $c) \Rightarrow a)$. Sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base di V e assumiamo che essa venga trasformata in una base $\{F(v_1), \ldots, F(v_n)\}$ di W. Ai fini di provare che F è un isomorfismo, stante il risultato precedente, basta controllare che F è surgettiva. Infatti abbiamo:

$$Im(F) = L(F(v_1), \ldots, F(v_n)) = W.$$

□

Concludiamo questo paragrafo con un'altra caratterizzazione, operativamente conveniente, degli isomorfismi in termini di matrici:

Teorema 32.3. *Sia $F : V \to W$ un'applicazione lineare tra due $\mathbb{K}$-spazi vettoriali aventi la stessa dimensione $n \geq 1$. Siano $\mathfrak{B}$ e $\mathfrak{C}$ basi di V e di W rispettivamente. Allora*

$$F \text{ è un isomorfismo} \iff M^{\mathfrak{C}}_{\mathfrak{B}}(F) \in GL(n, \mathbb{K}).$$

Dimostrazione: Infatti, tenendo conto dei Teoremi 32.1 e 31.3:

$$F \text{ isomorfismo} \iff rg(F) = n \iff rg(M^{\mathfrak{C}}_{\mathfrak{B}}(F)) = n \iff M^{\mathfrak{C}}_{\mathfrak{B}}(F) \in GL(n, \mathbb{K}).$$

□

33. Proprietà della matrice associata a un'applicazione lineare

Il Teorema che segue spiega come si trasformano le componenti di un vettore applicandovi un'applicazione lineare:

Teorema 33.1. *Siano V, W spazi vettoriali entrambi non banali sullo stesso campo $\mathbb{K}$, entrambi finitamente generati. Si fissino una base $\mathfrak{B}$ di V ed una base $\mathfrak{C}$ di W.*

Sia $F : V \to W$ un'applicazione lineare e sia $A = M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ la matrice associata ad F rispetto a $\mathfrak{B}$ e $\mathfrak{C}$.

Per ogni vettore $v \in V$, denotato con x il vettore delle componenti di v rispetto alla base $\mathfrak{B}$, si ha che il vettore y delle componenti di $F(v)$ rispetto a $\mathfrak{C}$ è dato da:

$$y = Ax. \tag{20}$$

Dimostrazione: Infatti, posto $\mathfrak{B} = \{v_1, \ldots, v_n\}$, abbiamo, posto $x = (x_1, \ldots, x_n)$:

$$v = x_1 v_1 + \cdots + x_n v_n,$$

da cui

$$F(v) = x_1 F(v_1) + \cdots + x_n F(v_n).$$

Per definizione il vettore y delle componenti di $F(v)$ si ottiene calcolando $\psi_{\mathfrak{C}}(F(v))$; poichè anche questa applicazione è lineare si ricava:

$$y = \psi_{\mathfrak{C}}(F(v)) = x_1 \psi_{\mathfrak{C}}(F(v_1)) + \cdots + x_n \psi_{\mathfrak{C}}(F(v_n)).$$

Ma ciascun vettore $\psi_{\mathfrak{C}}(F(v_i))$ non è altro che la colonna i-ma di A; pertanto in definitiva

$$y = x_1 A_{(1)} + \cdots + x_n A_{(n)} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax,$$

dove si è applicata la (12). $\square$

34. Esistenza ed unicità di applicazioni lineari

In questo paragrafo dimostriamo un risultato importante che permette di costruire, fissato un $\mathbb{K}$-spazio vettoriale V di dimensione finita, applicazioni lineari $F : V \to W$, dove W è un arbitrario $\mathbb{K}$-spazio vettoriale.

Cominciamo col puntualizzare che un'applicazione lineare definita su uno spazio vettoriale di dimensione finita è completamente determinata dai valori che assume su una fissata base. Questo risultato generalizza quanto visto in precedenza per un'applicazione $F : \mathbb{K}^n \to \mathbb{K}^m$, dove si utilizzava la base canonica di $\mathbb{K}^n$.

Teorema 34.1. *Siano V e W due $\mathbb{K}$-spazi vettoriali, di cui il primo di dimensione $n \geq 1$, e sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base di V.*

Date due applicazioni lineari $F : V \to W$ e $G : V \to W$, si ha:

$$F = G \iff \forall i \in \{1, \ldots, n\} \quad F(v_i) = G(v_i).$$

Dimostrazione: È chiaro che la condizione enunciata è necessaria affinchè $F = G$. Siano $F, G : V \to W$ lineari e si assuma che $F(v_i) = G(v_i)$ per ogni $i \in \{1, \ldots, n\}$. Consideriamo un vettore $v \in V$; possiamo scrivere in modo unico tale vettore nella forma $v = x_1 v_1 + \cdots + x_n v_n$, da cui per la linearità di entrambe le applicazioni:

$$F(v) = F(x_1v_1+\cdots+x_nv_n) = x_1F(v_1)+\cdots+x_nF(v_n) = x_1G(v_1)+\cdots+x_nG(v_n) = G(v).$$

Stante l'arbitrarietà di v abbiamo che $F = G$. $\square$

Il risultato seguente mostra che si possono costruire applicazioni lineari *prescrivendone* i valori assunti sui vettori di una base:

Teorema 34.2. *(Esistenza ed unicità delle applicazioni lineari)*

Siano V e W spazi vettoriali sul campo $\mathbb{K}$ e si assuma V di dimensione finita $n > 0$. Si fissi una base $\mathfrak{B} = \{v_1, \ldots, v_n\}$ di V. Siano $w_1, \ldots, w_n$ vettori di W arbitrari. Allora esiste una ed una sola applicazione lineare $F : V \to W$ tale che

$$\forall i \in \{1, \ldots, n\} \quad F(v_i) = w_i. \tag{21}$$

Dimostrazione: L'unicità dell'applicazione F nell'enunciato è garantita dal Teorema precedente. Proviamo che una tale F esiste. Assegnati i vettori $w_1, \dots, w_n$, definiamo un'applicazione $F : V \to W$ nel modo seguente:

$F(v) := x_1w_1 + \cdots + x_nw_n$, dove $(x_1, \dots, x_n)$ sono le componenti di v rispetto a $\mathfrak{B}$.

Si osservi che la definizione è ben posta per l'unicità delle componenti di ciascun vettore v di V. Verifichiamo innanzitutto che questa funzione soddisfa (21). Infatti, notiamo che per ogni $i \in \{1, \dots, n\}$, risulta

$$v_i = 0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + v_n \tag{22}$$

ovvero per ogni i le componenti di v_i sono tutte nulle tranne la i-ma che è uguale a 1. Dunque direttamente dalla definizione di F, abbiamo $F(v_i) = w_i$.

Resta da provare che l'applicazione F è lineare. Posto $A = \{w_1, \dots, w_n\}$, vi è l'applicazione lineare

$$\varphi_A : \mathbb{K}^n \to W, \quad \varphi_A(x_1, \dots, x_n) = x_1w_1 + \cdots + x_nw_n.$$

Allora basta osservare che F coincide con la funzione composta

$$F = \varphi_A \circ \psi_{\mathfrak{B}}.$$

Infatti, per ogni vettore $v \in V$ di componenti $(x_1, \dots, x_n)$ si ha:

$$(\varphi_A \circ \psi_{\mathfrak{B}})(v) = \varphi_A(\psi_{\mathfrak{B}}(v)) = \varphi_A(x_1, \dots, x_n) = x_1w_1 + \cdots + x_nw_n = F(v).$$

Dunque F è lineare perchè composta di applicazioni lineari.

Si invita il lettore a fare anche una verifica diretta della linearità di F usandone direttamente la definizione, a titolo di esercizio. □

Sottolineamo che i vettori $w_1, \dots, w_n$ possono essere scelti in modo arbitrario; ad esempio, se $w_1 = \cdots = w_n = 0$, allora l'applicazione lineare risultante è necessariamente la funzione nulla $0 : V \to W$.

35. Relazione tra applicazioni lineari e matrici

Discutiamo ora l'importante risultato seguente, che stabilisce un legame essenziale tra le nozioni di applicazione lineare e quella di matrice. Ricordiamo che l'insieme $Hom(V, W)$ di tutte le applicazioni lineari tra due $\mathbb{K}$-spazi vettoriali assegnati è esso stesso un $\mathbb{K}$-spazio vettoriale.

Teorema 35.1. *Siano V, W spazi vettoriali entrambi non banali sullo stesso campo $\mathbb{K}$, entrambi finitamente generati ed aventi dimensioni n ed m rispettivamente. Si fissino una base $\mathfrak{B}$ di V ed una base $\mathfrak{C}$ di W. L'applicazione*

$$M^{\mathfrak{C}}_{\mathfrak{B}} : Hom(V, W) \to M_{m,n}(\mathbb{K})$$

che fa corrispondere ad ogni $F \in Hom(V, W)$ la matrice $M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ ad essa associata rispetto alle due basi, è un isomorfismo di spazi vettoriali.

Come conseguenza rilevante, abbiamo:

Corollario 35.2. *Se V e W sono spazi vettoriali non banali di dimensione n ed m, allora anche $Hom(V,W)$ è finitamente generato e*

$$\mathtt{dim}(Hom(V,W)) = mn.$$

Per dimostrare il Teorema 35.1, assumiamo $\mathfrak{B} = \{v_1, \ldots, v_n\}$. Proviamo innanzitutto che $M_{\mathfrak{B}}^{\mathfrak{C}}$ è ingettiva; siano $F : V \to W$ e $G : V \to W$ due applicazioni lineari. Se

$$M_{\mathfrak{B}}^{\mathfrak{C}}(F) = M_{\mathfrak{B}}^{\mathfrak{C}}(G),$$

allora in particolare le colonne di queste matrici coincidono:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(F)_{(i)} = M_{\mathfrak{B}}^{\mathfrak{C}}(G)_{(i)},$$

ma ciò significa che

$$\psi_{\mathfrak{C}}(F(v_i)) = \psi_{\mathfrak{C}}(G(v_i)),$$

e quindi, siccome $\psi_{\mathfrak{C}}$ è ingettiva:

$$F(v_i) = G(v_i).$$

Dunque possiamo concludere che $F = G$ in forza del Teorema 34.1.

Discutiamo ora la surgettività della nostra applicazione. Sia $A \in M_{m,n}(\mathbb{K})$ una matrice assegnata e consideriamo i seguenti vettori di W:

$$w_i := \psi_{\mathfrak{C}}^{-1}(A_{(i)}), \quad i = 1, \ldots, n.$$

Per costruzione, ciascun vettore w_i ha proprio il vettore colonna $A_{(i)}$ come vettore delle sue componenti rispetto alla base $\mathfrak{C}$. Applicando il Teorema 34.2, abbiamo che esiste una ed una sola applicazione lineare $F : V \to W$ tale che

$$F(v_i) = w_i, \quad i = 1, \ldots, n.$$

Risulta allora per costruzione che $M_{\mathfrak{B}}^{\mathfrak{C}}(F) = A$. Si ha infatti:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(F) = (\, \psi_{\mathfrak{C}}(F(v_1)) \ \cdots \ \psi_{\mathfrak{C}}(F(v_n)) \,) = (\, \psi_{\mathfrak{C}}(w_1) \ \cdots \ \psi_{\mathfrak{C}}(w_n) \,) = (\, A_{(1)} \ \cdots \ A_{(n)} \,) = A.$$

Resta da provare che l'applicazione $M_{\mathfrak{B}}^{\mathfrak{C}}$ è lineare. Si noti che ciò corrisponde a stabilire due proprietà utili delle matrici associate, ovvero:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(F+G) = M_{\mathfrak{B}}^{\mathfrak{C}}(F) + M_{\mathfrak{B}}^{\mathfrak{C}}(G), \ M_{\mathfrak{B}}^{\mathfrak{C}}(\lambda F) = \lambda M_{\mathfrak{B}}^{\mathfrak{C}}(F),$$

dove λ è uno scalare.

Siano $F, G \in Hom(V,W)$. Allora

$$\begin{aligned}
&M_{\mathfrak{B}}^{\mathfrak{C}}(F+G) = (\, \psi_{\mathfrak{C}}((F+G)(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}((F+G)(v_n)) \,) = \\
&(\, \psi_{\mathfrak{C}}(F(v_1) + G(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(F(v_n) + G(v_n)) \,) = \\
&(\, \psi_{\mathfrak{C}}(F(v_1)) + \psi_{\mathfrak{C}}(G(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(F(v_n)\,) + \psi_{\mathfrak{C}}(G(v_n)) \,) = M_{\mathfrak{B}}^{\mathfrak{C}}(F) + M_{\mathfrak{B}}^{\mathfrak{C}}(G)
\end{aligned}$$

dove abbiamo tenuto conto della definizione di $F+G$ e sfruttato la linearità di $\psi_{\mathfrak{C}}$. Analogamente:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(\lambda F) = (\, \psi_{\mathfrak{C}}((\lambda F)(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}((\lambda F)(v_n))\,) =$$

$$(\, \psi_{\mathfrak{C}}(\lambda F(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(\lambda F(v_n))\,) = (\, \lambda\psi_{\mathfrak{C}}(F(v_1)) \quad \cdots \quad \lambda\psi_{\mathfrak{C}}(F(v_n))\,) = \lambda M_{\mathfrak{B}}^{\mathfrak{C}}(F).$$

Ciò conclude la dimostrazione del Teorema 35.1. $\square$

36. Relazione tra composizione di applicazioni lineari e prodotto di matrici

Teorema 36.1. *Siano V, W ed U spazi vettoriali sullo stesso campo $\mathbb{K}$, tutti di dimensione finita. Siano $\mathfrak{B}$, $\mathfrak{C}$ e $\mathfrak{D}$ basi di V, W ed U rispettivamente. Siano $F: V \to W$ e $G: W \to U$ applicazioni lineari. Allora:*

$$M_{\mathfrak{B}}^{\mathfrak{D}}(G \circ F) = M_{\mathfrak{C}}^{\mathfrak{D}}(G) M_{\mathfrak{B}}^{\mathfrak{C}}(F). \tag{23}$$

Dimostrazione: Sfrutteremo la proprietà fondamentale (20) della matrice associata a un'applicazione lineare. Denotate rispettivamente con n, m e s le dimensioni di V, W e U, e posto:

$$A := M_{\mathfrak{B}}^{\mathfrak{C}}(F),\ B := M_{\mathfrak{C}}^{\mathfrak{D}}(G),\ \mathfrak{B} = \{v_1, \dots, v_n\},$$

per ogni vettore v di V con vettore di componenti x rispetto alla base $\mathfrak{B}$, abbiamo che il vettore delle componenti y di $F(v)$ è dato da $y = Ax$; quindi il vettore z delle componenti di $G(F(v))$ rispetto alla base $\mathfrak{D}$ è dato da

$$z = By = B(Ax) = (BA)x.$$

In particolare, per ogni vettore v_i di $\mathfrak{B}$, essendo $x = e_i$, si ottiene che $(G \circ F)(v_i)$ ha vettore di componenti:

$$(BA)e_i = (BA)_{(i)}.$$

Pertanto, stante la definizione, la matrice associata a $G \circ F$ è:

$$M_{\mathfrak{B}}^{\mathfrak{D}}(G \circ F) = ((BA)_{(1)} \cdots (BA)_{(n)}) = BA.$$

$\square$

Prima di procedere facciamo un'utile osservazione riguardante la matrice associata all'applicazione identica:

Esempio 36.2. Si ha, qualunque sia la base $\mathfrak{B}$ di V:

$$M_{\mathfrak{B}}^{\mathfrak{B}}(Id_V) = I_n, \quad n = \mathtt{dim}(V).$$

Infatti, posto $\mathfrak{B} = \{v_1, \dots, v_n\}$, ricordando che $\psi_{\mathfrak{B}}(v_i) = e_i$:

$$\begin{aligned} M_{\mathfrak{B}}^{\mathfrak{B}}(Id_V) &= \big(\psi_{\mathfrak{B}}(Id_V(v_1)) \quad \cdots \quad \psi_{\mathfrak{B}}(Id_V(v_n))\big) \\ &= \big(\psi_{\mathfrak{B}}(v_1) \quad \cdots \quad \psi_{\mathfrak{B}}(v_n))\big) = \big(e_1 \quad \cdots \quad e_n\big) = I_n. \end{aligned}$$

Corollario 36.3. *Sia $F : V \to W$ un isomorfismo tra due spazi finitamente generati aventi dimensione $n \geq 1$. Siano $\mathfrak{B}$ e $\mathfrak{C}$ basi di V e di W rispettivamente. Allora*

$$M^{\mathfrak{B}}_{\mathfrak{C}}(F^{-1}) = M^{\mathfrak{C}}_{\mathfrak{B}}(F)^{-1}. \tag{24}$$

Dimostrazione: Ricordiamo intanto che la matrice $M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ è invertibile stante il Teorema 32.3. Dalla relazione

$$F \circ F^{-1} = Id_W$$

segue, in forza del Teorema precedente:

$$M^{\mathfrak{C}}_{\mathfrak{B}}(F) M^{\mathfrak{B}}_{\mathfrak{C}}(F^{-1}) = M^{\mathfrak{C}}_{\mathfrak{C}}(Id_W) = I_n$$

da cui si ricava immediatamente la (24) moltiplicando ambo i membri per $M^{\mathfrak{C}}_{\mathfrak{B}}(F)^{-1}$.
$\square$

37. Cambiamenti di base

Sia V uno spazio vettoriale di dimensione $n > 0$ sul campo $\mathbb{K}$ e siano assegnate due basi $\mathfrak{B} = \{v_1, \dots, v_n\}$ e $\mathfrak{B}' = \{v'_1, \dots, v'_n\}$ di V.

Definizione 37.1. Si chiama **matrice di passaggio** da $\mathfrak{B}$ a $\mathfrak{B}'$ e si denota col simbolo

$$M_{\mathfrak{B},\mathfrak{B}'}$$

la matrice associata ai vettori $v'_1, \dots, v'_n$ rispetto a $\mathfrak{B}$.

Esempio 37.2. Ad esempio, considerate la base canonica $\mathfrak{B}_0 = \{(1, 0), (0, 1)\}$ di $\mathbb{R}^2$ e la base $\mathfrak{B} = \{(1, 2), (1, 1)\}$ allora risulta che la matrice di passaggio da $\mathfrak{B}_0$ a $\mathfrak{B}$ è:

$$M_{\mathfrak{B}_0,\mathfrak{B}} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix},$$

che si ottiene semplicemente scrivendo in colonna i vettori di $\mathfrak{B}$, mentre la matrice di passaggio da $\mathfrak{B}$ a $\mathfrak{B}_0$ risulta la seguente:

$$M_{\mathfrak{B},\mathfrak{B}_0} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

In tal caso le due colonne sono i vettori delle componenti di $(1, 0)$ e $(0, 1)$ rispetto a $\mathfrak{B}$; si ha infatti:

$$(1, 0) = -(1, 2) + 2(1, 1), \ (0, 1) = (1, 2) - (1, 1).$$

Osserviamo che la matrice di passaggio da una base ad un'altra è sempre una matrice invertibile; ciò è garantito dal Cor. (30.3). Si può ricavare ciò anche dal risultato seguente, che ne fornisce un'altra interpretazione:

Proposizione 37.3. *Assegnate due basi* $\mathfrak{B}$ *e* $\mathfrak{B}'$ *dello spazio vettoriale* V *risulta che:*

$$M_{\mathfrak{B},\mathfrak{B}'} = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V).$$

In particolare

$$M_{\mathfrak{B},\mathfrak{B}'} \in GL(n,\mathbb{K}).$$

Inoltre:

$$M_{\mathfrak{B}',\mathfrak{B}} = (M_{\mathfrak{B},\mathfrak{B}'})^{-1}.$$

Dimostrazione: Abbiamo infatti

$$M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V) = (\psi_{\mathfrak{B}}(Id_V(v_1')) \ \cdots \ \ \psi_{\mathfrak{B}}(Id_V(v_n'))) = (\psi_{\mathfrak{B}}(v_1') \ \cdots \ \ \psi_{\mathfrak{B}}(v_n')) = M_{\mathfrak{B},\mathfrak{B}'}.$$

Poichè Id_V è un isomorfismo, segue dal Corollario 32.3 che $M_{\mathfrak{B},\mathfrak{B}'}$ è invertibile. L'ultima affermazione segue dal Corollario 36.3 applicato a Id_V:

$$M_{\mathfrak{B}',\mathfrak{B}} = M^{\mathfrak{B}'}_{\mathfrak{B}}(Id_V) = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V^{-1})^{-1} = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V)^{-1} = (M_{\mathfrak{B},\mathfrak{B}'})^{-1}.$$

□

Il risultato seguente illustra come cambiano le coordinate di un vettore effettuando un cambiamento di base:

Teorema 37.4. *Dato un vettore* $v \in V$*, siano* x *e* x' *i vettori delle componenti di* v *rispetto alle basi* $\mathfrak{B}$ *e* $\mathfrak{B}'$ *rispettivamente. Denotata con* A *la matrice di passaggio da* $\mathfrak{B}$ *a* $\mathfrak{B}'$*, sussiste la relazione:*

$$x = Ax'. \tag{25}$$

Dimostrazione: Siccome $A = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V)$, basta applicare il Corollario 33.1 all'applicazione identica $Id_V : V \to V$, dove lo spazio di partenza è considerato munito della base $\mathfrak{B}'$, mentre quello di arrivo si considera munito della base $\mathfrak{B}$. Sappiamo infatti che per ogni $v \in V$ il legame tra le coordinate di v e di $Id_V(v) = v$ rispetto alle due basi (dove va considerata prima $\mathfrak{B}'$ e poi $\mathfrak{B}$) è proprio dato dalla (25). □

Studiamo infine come cambia la matrice associata ad un'applicazione lineare facendo scelte diverse delle basi negli spazi di partenza e di arrivo:

Teorema 37.5. *(del cambiamento di base)*

Siano V, W *spazi vettoriali non banali sullo stesso campo* $\mathbb{K}$*, entrambi finitamente generati. Sia* $F : V \to W$ *un'applicazione lineare. Si fissino due basi* $\mathfrak{B}$*,* $\mathfrak{B}'$ *di* V *e due basi* $\mathfrak{C}$*,* $\mathfrak{C}'$ *di* W*. Allora risulta*

$$M^{\mathfrak{C}'}_{\mathfrak{B}'}(F) = C^{-1} M^{\mathfrak{C}}_{\mathfrak{B}}(F)\, B \tag{26}$$

dove B *è la matrice di passaggio da* $\mathfrak{B}$ *a* $\mathfrak{B}'$*, mentre* C *è la matrice di passaggio da* $\mathfrak{C}$ *a* $\mathfrak{C}'$*.*

Dimostrazione: L'applicazione F può riguardarsi come composizione delle applicazioni seguenti:

$$\begin{array}{ccccccc} V & \xrightarrow{Id_V} & V & \xrightarrow{F} & W & \xrightarrow{Id_W} & W \\ \mathfrak{B}' & & \mathfrak{B} & & \mathfrak{C} & & \mathfrak{C}' \end{array}$$

dove abbiamo indicato la scelta di una base per ciascuno spazio coinvolto. Applicando il Teorema 36.1 si ottiene quindi:

$$\begin{aligned} M^{\mathfrak{C}'}_{\mathfrak{B}'}(F) &= M^{\mathfrak{C}'}_{\mathfrak{B}'}(Id_W \circ F \circ Id_V) = M^{\mathfrak{C}'}_{\mathfrak{B}'}((Id_W \circ F) \circ Id_V) = \\ &= M^{\mathfrak{C}'}_{\mathfrak{B}}(Id_W \circ F) M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V) = M^{\mathfrak{C}'}_{\mathfrak{B}}(Id_W \circ F) B = \\ &= M^{\mathfrak{C}'}_{\mathfrak{C}}(Id_W) M^{\mathfrak{C}}_{\mathfrak{B}}(F) = C^{-1} M^{\mathfrak{C}}_{\mathfrak{B}}(F) B \end{aligned}$$

dove si è tenuto conto della Prop. 37.3. □

38. Endomorfismi di uno spazio vettoriale

Definizione 38.1. Si chiama *endomorfismo* di un $\mathbb{K}$-spazio vettoriale V ogni applicazione lineare $F : V \to V$.

Sono esempi di endomorfismi l'endomorfismo nullo $0 : V \to V$ e quello identico $Id_V : V \to V$. Lo spazio vettoriale $Hom(V, V)$ costituito da tutti gli endomorfismi dello spazio V verrà denotato con $End(V)$.

Per gli endomorfismi si particolarizza la nozione di matrice associata, facendo uso di una singola di base di V, come segue:

Definizione 38.2. Sia V uno spazio vettoriale di dimensione finita $n \geq 1$, $F \in End(V)$ e $\mathfrak{B}$ una base di V. Si pone

$$M_{\mathfrak{B}}(F) := M^{\mathfrak{B}}_{\mathfrak{B}}(F).$$

La matrice quadrata $M_{\mathfrak{B}}(F) \in M_n(\mathbb{K})$ verrà detta la *matrice associata a F rispetto alla base* $\mathfrak{B}$.

Sappiamo che, fissata la base $\mathfrak{B}$, l'applicazione

$$M_{\mathfrak{B}} : End(V) \to M_n(\mathbb{K})$$

che associa ad ogni endomorfismo $F : V \to V$ la corrispondente matrice associata (rispetto alla base assegnata), è un isomorfismo. In particolare $End(V)$ ha dimensone n^2.

Applicando il Teorema del cambiamento di base in questo caso, otteniamo la seguente relazione tra le metrici associate ad un endomorfismo rispetto a due basi diverse:

Proposizione 38.3. *Sia $F \in End(V)$ e siano $\mathfrak{B}$ e $\mathfrak{B}'$ due basi di V. Allora, posto $A = M_{\mathfrak{B}}(F)$ e $B = M_{\mathfrak{B}'}(F)$, si ha:*

$$B = M^{-1}AM \tag{27}$$

dove M è la matrice di passaggio da $\mathfrak{B}$ a $\mathfrak{B}'$.

Conviene introdurre a questo proposito la seguente nozione di carattere generale:

Definizione 38.4. Siano $A, B \in M_n(\mathbb{K})$ matrici quadrate dello stesso ordine. Si dice che A e B sono **simili** se esiste una matrice invertibile $M \in GL(n, \mathbb{K})$ tale che:

$$B = M^{-1}AM. \tag{28}$$

La relazione di similitudine sull'insieme $M_n(\mathbb{K})$ è simmetrica, perchè se vale (28), allora si ha anche:

$$A = MBM^{-1},$$

con M^{-1} invertibile. Si tratta in realtà di una relazione di equivalenza. Notiamo intanto che ogni matrice A è simile ad A, avendosi

$$A = I_n^{-1}AI_n.$$

Lasciamo al lettore la semplice verifica che, date le matrici $A, B, C \in M_n(\mathbb{K})$, allora se A è simile a B e B è simile a C, segue che A è simile a C.

Dunque con questa terminologia possiamo riformulare la Proposizione precedente come segue:

Corollario 38.5. *Sia $F \in End(V)$. Le matrici associate a F rispetto a due basi diverse $\mathfrak{B}$ e $\mathfrak{B}'$ di V sono sempre simili tra loro.*

Proposizione 38.6. *Siano $A, B \in M_n(\mathbb{K})$, $n \geq 1$, due matrici simili. Allora*

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

Dimostrazione: Infatti, per ipotesi

$$B = M^{-1}AM$$

per un'opportuna matrice invertibile M, da cui, applicando il Teorema di Binet:

$$\det(B) = \det(M^{-1})\det(B)\det(M) = \frac{1}{\det(M)}\det(A)\det(M) = \det(A).$$

□

Il fatto che matrici simili hanno lo stesso determinante giustifica la definizione seguente, in forza del Corollario precedente:

Definizione 38.7. Siano V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $n \geq 1$ e $F : V \to V$ un endomorfismo. Si chiama **determinante** di F lo scalare

$$\det(F) := \det(M_{\mathfrak{B}}(F)),$$

dove $\mathfrak{B}$ è una base di V scelta in modo arbitrario.

Chiameremo *automorfismo* dello spazio V ogni endomorfismo che è un isomorfismo. L'insieme $Aut(V)$ di tutti gli automorfismi di V forma un gruppo rispetto all'operazione di composizione di funzioni. Il seguente risultato è conseguenza immediata del Teorema 32.3:

Proposizione 38.8. *Siano V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $n \geq 1$ e sia $F : V \to V$ un endomorfismo. Allora*

$$F \in Aut(V) \iff \det(F) \neq 0.$$

39. Endomorfismi diagonalizzabili

In questo paragrafo studiamo un problema importante in Algebra Lineare e di interesse in diversi campi della matematica. Si tratta di stabilire sotto quali condizioni un endomorfismo di un fissato spazio vettoriale sia rappresentabile con una matrice associata che sia di una tipologia più semplice possibile. Introduciamo a tal scopo la nozione di matrice diagonale:

Definizione 39.1. Una matrice quadrata $A = (a^i_j) \in M_n(\mathbb{K})$ si dice **diagonale** se

$$a^i_j = 0 \qquad \forall i \neq j.$$

In tal caso A è della forma:

$$A = \begin{pmatrix} a^1_1 & & & \\ & a^2_2 & & \\ & & \ddots & \\ & & & a^n_n \end{pmatrix},$$

dove si sottindende gli elementi di A non scritti sono tutti uguali a 0.

In generale, gli elementi a^i_i di una matrice quadrata A si dicono *diagonali* e formano la cosiddetta *diagonale principale* di A. Quindi in una matrice diagonale essi sono gli unici che possono essere non nulli.

Useremo anche la notazione

$$A = diag(\lambda_1, \ldots, \lambda_n)$$

per indicare la matrice diagonale di ordine n tale che

$$a^1_1 = \lambda_1, \ldots, a^n_n = \lambda_n.$$

Notiamo che, rappresentando tale matrice per colonne, abbiamo:

$$A = \begin{pmatrix} \lambda_1 e_1 & \cdots & \lambda_n e_n \end{pmatrix}.$$

Definizione 39.2. Siano V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $n \geq 1$ ed $F : V \to V$ un endomorfismo. Si dirà che F è **diagonalizzabile** se esiste una base $\mathfrak{B}$ di V tale che la matrice associata $M_{\mathfrak{B}}(F)$ a F rispetto a $\mathfrak{B}$ è diagonale. Una base di V con questa proprietà è detta *base diagonalizzante per* F.

Nel seguito discuteremo dei criteri utili per stabilire se un endomorfismo è diagonalizzabile. Discutiamo ora delle immediate conseguenze di questa condizione. Si supponga che la matrice $A = M_{\mathfrak{B}}(F)$ rispetto a un'opportuna base $\mathfrak{B}$ sia diagonale:

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Allora, per ogni vettore $v \in V$ il cui vettore delle componenti rispetto a $\mathfrak{B}$ è $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, abbiamo che il vettore $F(v)$ ha componenti date dal vettore:

$$y = \begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix}.$$

Infatti, sappiamo che y si ricava subito mediante la relazione $y = Ax$. Dunque, per ogni $v \in V$, il vettore $F(v)$ si ottiene da v semplicemente "riscalandone" le coordinate mediante i fattori costanti $\lambda_1, \dots, \lambda_n$ (indipendenti da v).

In particolare, posto $\mathfrak{B} = \{v_1, \dots, v_n\}$ possiamo ricavare che le immagini dei vettori della base diagonalizzante sono dati da:

$$F(v_i) = \lambda_i v_i, \qquad i = 1, \dots, n. \tag{29}$$

Ad esempio, $F(v_1) = \lambda_1 v_1$, perchè le componenti di v_1 sono $(1, 0, \dots, 0)$ e quindi $F(v_1)$ ha componenti $(\lambda_1, 0, \dots, 0)$. Analogamente per gli altri vettori v_i, usando il fatto che tale vettore ha componenti date da e_i.

Le (29) si possono anche giustificare usando l'isomorfismo $\psi_{\mathfrak{B}} : V \to \mathbb{K}^n$; infatti, tenendo conto della definizione di A:

$$A = (\psi_{\mathfrak{B}}(F(v_1)) \cdots \psi_{\mathfrak{B}}(F(v_n))) = diag(\lambda_1, \dots, \lambda_n) = (\lambda_1 e_1 \cdots \lambda_n e_n),$$

da cui

$$\psi_{\mathfrak{B}}(F(v_i)) = \lambda_i e_i$$

ovvero

$$\psi_{\mathfrak{B}}(F(v_i)) = \lambda_i \psi_{\mathfrak{B}}(v_i) = \psi_{\mathfrak{B}}(\lambda_i v_i)$$

e si ricava che $F(v_i) = \lambda_i v_i$.

Le relazioni (29) suggeriscono di introdurre la seguente definizione.

Definizione 39.3. Sia $F : V \to V$ un endomorfismo. Un vettore $v \in V$ si dice **autovettore** di F se $v \neq 0$ ed inoltre esiste $\lambda \in \mathbb{K}$ tale che

$$F(v) = \lambda v. \tag{30}$$

Si noti che lo scalare λ è *univocamente determinato* dalla condizione precedente: infatti se $\mu \in \mathbb{K}$ è tale che

$$F(v) = \mu v,$$

allora

$$\mu v = \lambda v$$

da cui $\mu = \lambda$ perchè $v \neq 0$.

Se v è un autovettore di F, l'unico scalare λ per cui è soddisfatta la (30) si chiamerà l'**autovalore** relativo a v.

Proposizione 39.4. *Siano $F \in End(V)$ un endomorfismo e sia $\mathfrak{B} = \{v_1, \dots, v_n\}$ una base di V. Allora si ha:*

$$\mathfrak{B} \text{ è una base diagonalizzante per } F \iff v_1, \dots, v_n \text{ sono autovettori di } F.$$

Dimostrazione: L'implicazione $\Rightarrow$ è stata già provata, tenendo conto della (29). Riguardo l'altra, supponiamo che ogni v_i sia un autovettore, con autovalore associato λ_i. Allora

$$\begin{aligned} M_{\mathfrak{B}}(F) &= (\psi_{\mathfrak{B}}(F(v_1)) \cdots \psi_{\mathfrak{B}}(F(v_n))) = (\psi_{\mathfrak{B}}(\lambda_1 v_1) \cdots \psi_{\mathfrak{B}}(\lambda_n v_n)) = \\ &= (\lambda_1 \psi_{\mathfrak{B}}(v_1) \cdots \lambda_n \psi_{\mathfrak{B}}(v_n)) = (\lambda_1 e_1 \cdots \lambda_n e_n) = diag(\lambda_1, \cdots, \lambda_n) \end{aligned}$$

e quindi $\mathfrak{B}$ è una base diagonalizzante per F. $\square$

Si ottiene quindi il seguente criterio di diagonalizzabilità:

Corollario 39.5. *Un endomorfismo $F : V \to V$ è diagonalizzabile se e solo se lo spazio vettoriale V ammette una base $\mathfrak{B}$ di V costituita da autovettori di F.*

40. Autovalori e autospazi di un endomorfismo

Definizione 40.1. Sia $F \in End(V)$ un endomorfismo di un $\mathbb{K}$-spazio vettoriale e sia $\lambda \in \mathbb{K}$. Si dirà che λ è un **autovalore di** F se esiste un vettore $v \in V$, $v \neq 0$, tale che

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

In altri termini, ciò significa che λ è l'autovolare associato ad un opportuno autovettore di F.

Definizione 40.2. L'insieme di tutti gli autovalori di F di un endomorfismo è detto **spettro** di F e si denota con $Sp(F)$.

Come vedremo, può accadere che lo spettro di un endomorfismo sia vuoto; ciò dipende dalle proprietà del campo $\mathbb{K}$.

Definizione 40.3. Sia $\lambda \in \mathbb{K}$ un autovalore dell'endomorfismo. Si chiama **autospazio** di F relativo a λ l'insieme

$$V_\lambda(F) := \{v \in V |\, F(v) = \lambda v\}.$$

Risulta che:

Proposizione 40.4. *Per ogni autovalore λ, $V_\lambda(F)$ è un sottospazio vettoriale di V.*

Dimostrazione: Basta osservare che

$$\begin{aligned} V_\lambda(F) &= \{v \in V |\, F(v) = \lambda v\} = \{v \in V |\, F(v) - \lambda v = 0_V\} \\ &= \{v \in V |\, F(v) - \lambda Id_V(v) = 0_V\} \\ &= \{v \in V |\, (F(v) - \lambda Id_V)(v) = 0_V\} \\ &= Ker(F - \lambda Id_V). \end{aligned}$$

□

Osserviamo esplicitamente che, per definizione di autovalore, se $\lambda \in Sp(F)$, allora $V_\lambda(F) \neq \{0\}$; inoltre

$$V_\lambda(F) = \{0_V\} \cup \{v \in V : v \text{ è autovettore di } F \text{ con autovalore } \lambda\}.$$

Ai fini di determinare lo spettro di un endomorfismo , si introduce un oggetto cruciale, di carattere algebrico. Come nei casi $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$, un polinomio $p(t)$ a coefficienti nel campo $\mathbb{K}$ è un'espressione formale del tipo

$$p(t) = a_0 + a_1 t + \cdots + a_k t^k.$$

Il grado d di un poliomio $p(t)$ non nullo è il più grande intero $d \geq 0$ tale che $a_d \neq 0$. Il coefficiente a_d si chiama coefficiente direttivo di $p(t)$.

La funzione polinomiale associata a $p(t)$ è la funzione:

$$p : x \in \mathbb{K} \mapsto a_0 + a_1 x + \cdots a_k x^k,$$

e si chiama radice di $p(t)$ ogni scalare $\lambda \in \mathbb{K}$ per cui $p(\lambda) = 0$.

Anche nel contesto dei campi qualsiasi vale il Teorema di Ruffini, che caratterizza le radici λ di $p(t)$ mediante la condizione che il polinomio di primo grado $t - \lambda$ divida $p(t)$, nel senso che:

$$p(t) = (t - \lambda) q(t),$$

con $q(t)$ un altro polinomio opportuno.

Definizione 40.5. Siano V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $n \geq 1$ ed $F : V \to V$ un endomorfismo. Si chiama **polinomio caratteristico** di F il polinomio $p_F(t)$ definito da:

$$p_F(t) = \det(F - t Id_V). \tag{31}$$

Si può provare che il polinomio caratteristico di un endomorfismo $F : V \to V$ è effettivamente un polinomio di grado $n = \mathtt{dim}(V)$, avente coefficiente direttivo $(-1)^n$. Non discuteremo una dimostrazione formale di ciò, limitandoci a osservare quanto segue. Ai fini di esplicitare l'espressione di $p_F(t)$, fissiamo una base qualsiasi $\mathfrak{B}$ di V. Sia $A = (a^i_j)$ la matrice associata ad F rispetto a tale base. Allora la matrice associata all'endomorfismo $F - tId_V$ è data da

$$M_{\mathfrak{B}}(F - tId_V) = M_{\mathfrak{B}}(F) - tM_{\mathfrak{B}}(Id_V) = A - tI_n$$

in virtù del fatto che l'applicazione

$$g \in End(V) \mapsto M_{\mathfrak{B}}(g) \in M_n(\mathbb{K})$$

è lineare (Teorema 35.1) e che, rispetto a qualsiasi base $\mathfrak{B} = \{v_1, \ldots, v_n\}$, si ha sempre

$$M_{\mathfrak{B}}(Id_V) = I_n,$$

infatti:

$$M_{\mathfrak{B}}(Id_V) = (\psi_{\mathfrak{B}}(Id(v_1)) \ \cdots \ \psi_{\mathfrak{B}}(Id(v_n)) = (\psi_{\mathfrak{B}}(v_1) \ \cdots \ \psi_{\mathfrak{B}}(v_n)) = (e_1 \ \cdots \ e_n) = I_n.$$

In conclusione, per definizione di determinante di un endomorfismo, abbiamo:

$$p_F(t) = |A - tI_n| = \begin{vmatrix} a^1_1 - t & a^1_2 & \cdots & a^1_n \\ a^2_1 & a^2_2 - t & \cdots & a^2_n \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a^n_1 & a^n_2 & \cdots & a^n_n - t \end{vmatrix}. \tag{32}$$

Ora, sviluppando il determinante mediante la regola di Laplace, ad es. rispetto alla prima riga, si vede che tale epressione è somma di prodotti di n scalari che sono elementi a^i_j della matrice con $i \neq j$ o della forma $a^i_i - t$, e da ciò si ricava che si tratta di un polinomio in t. Tra questi prodotti vi è $(a^1_1 - t) \cdots (a^n_n - t)$, e ciò giustifica che il grado è n e che il coefficiente di t^n di tale polinomo è $(-1)^n$.

Osservazione 40.6. La (32) viene utilizzata per calcolare esplicitamente il polinomio caratteristico. Mettiamo in evidenza che A è la matrice associata ad F rispetto ad una base che può essere scelta in modo arbitrario.

L'importanza del polinomio caratteristico di un endomorfismo risiede nel fatto che è coinvolto nella seguente caratterizzazione degli autovalori:

Teorema 40.7. *(Cayley-Hamilton) Sia V un $\mathbb{K}$-spazio vettoriale di dimensione $n \geq 1$ e sia assegnato un endomorfismo $F : V \to V$. Sia $\lambda \in \mathbb{K}$. Allora*

$$\lambda \text{ è autovalore di } F \iff \lambda \text{ è una radice del polinomio caratteristico di } F.$$

In altri termini, lo spettro di un endomorfismo F è costituito da tutte sole le radici del suo polinomio caratteristico $p_F(t)$.

DIMOSTRAZIONE: Infatti, in base alla definizione di autovalore:

$$\begin{aligned}
\lambda \in Sp(F) \iff & \exists v \in V \ t.c. \ v \neq 0 \ \text{ e } F(v) = \lambda v \\
\iff & \exists v \in V \, t.c. \, v \neq 0 \ \text{ e } \ (F - \lambda Id_V)(v) = 0 \\
\iff & \exists v \in \ t.c. \, v \neq 0 \ \text{ e } v \in Ker(F - \lambda Id_V) \\
\iff & Ker(F - \lambda Id_V) \neq \{0_V\} \\
\iff & F - \lambda Id_V \text{ non è un monomorfismo} \\
\iff & F - \lambda Id_V \text{ non è un isomorfismo} \\
\iff & \det(F - \lambda Id_V) = 0.
\end{aligned}$$

□

Osservazione 40.8. In base al teorema di Ruffini, $p_F(t)$ ha al più n radici, perchè ha grado n. Si ottiene quindi che ogni endomorfismo di V ammette al più n autovalori distinti, dove n è la dimensione di V.

Vogliamo ora discutere una caratterizzazione degli endomorfismi diagonalizzabili.

Premettiamo che, come nel caso reale, dato un polinomio $p(t)$ a coefficienti in $\mathbb{K}$ ed una sua radice λ, dire che λ ha *molteplicità* $m \in \mathbb{N}$, $m \geq 1$, significa che $(t-\lambda)^m$ divide p, mentre $(t-\lambda)^{m+1}$ non divide $p(t)$. Equivalentemente, $p(t)$ si fattorizza come

$$p(t) = (t-\lambda)^m q(t)$$

essendo $q(t)$ un polinomio tale che $q(\lambda) \neq 0$. Ciò premesso, per ogni autovalore si introducono le nozioni seguenti:

Definizione 40.9. Sia $F \in End(V)$ e sia $\lambda \in Sp(F)$. Si chiamerà **molteplicità algebrica** di λ e si denoterà con $m_a(\lambda)$, la molteplicità di λ come radice del polinomio caratteristico di F.

Si chiama **molteplicità geometrica** dell'autovalore λ, e si denota con $m_g(\lambda)$, la dimensione $\mathtt{dim}(V_\lambda(F))$ dell'autospazio a esso relativo.

Sussiste un il seguente risultato importante che mette in relazione le due molteplicità di un autovalore:

Teorema 40.10. *Sia assegnato un endomorfismo $F : V \to V$ e sia $\lambda \in Sp(F)$. Allora*

$$m_g(\lambda) \leq m_a(\lambda).$$

DIMOSTRAZIONE: Sappiamo che l'autospazio $V_\lambda(F)$ non è banale. Posto $k := m_g(\lambda)$, si fissi quindi una base $\mathfrak{B}' = \{v_1, \ldots, v_k\}$ di $V_\lambda(F)$ e si consideri una base $\mathfrak{B} = \{v_1, \ldots, v_n\}$ di V ottenuta da $\mathfrak{B}'$ per completamento. Calcoliamo quindi $p_F(t)$

utilizzando la (32) in corrispondenza di tale base, dove $A = M_{\mathfrak{B}}(F)$. Per ogni $i = 1, \dots, k$ abbiamo:

$$F(v_i) = \lambda v_i,$$

per cui le prime k colonne di A sono i vettori $\lambda e_1, \dots, \lambda e_k$; infatti, per definizione di matrice associata, per ogni $i = 1, \dots, k$ abbiamo:

$$A_{(i)} = \psi_{\mathfrak{B}}(F(v_i)) = \psi_{\mathfrak{B}}(\lambda v_i) = \lambda e_i.$$

Quindi possiamo scrivere la matrice A in questo modo:

$$A = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & B \\ & & \lambda & \\ \hline & & & \\ & 0 & & C \\ & & & \end{array}\right)$$

dove $B \in M_{k,n-k}(\mathbb{K})$ e $C \in M_{n-k}(\mathbb{K})$, dove la sottomatrice diagonale $diag(\lambda, \dots, \lambda)$ è quadrata di ordine k. Allora la matrice $A - tI_n$ è data da

$$A - tI_n = \left(\begin{array}{ccc|c} \lambda - t & & & \\ & \ddots & & B \\ & & \lambda - t & \\ \hline & & & \\ & 0 & & C - tI_{n-k} \\ & & & \end{array}\right).$$

Il determinante di questa matrice si può calcolare utilizzando lo sviluppo di Laplace rispetto alla prima colonna, ottenendo:

$$|A - tI_n| = (\lambda - t) \left|\begin{array}{ccc|c} \lambda - t & & & \\ & \ddots & & B' \\ & & \lambda - t & \\ \hline & & & \\ & 0 & & C - tI_{n-k} \\ & & & \end{array}\right|,$$

dove B' si ottiene da B cancellandone la prima riga; continuando allo stesso modo con lo sviluppo di Laplace per un totale di k volte, sempre rispetto alla prima colonna, si perviene a:

$$|A - tI_n| = (\lambda - t)^k |C - tI_{n-k}|.$$

Vediamo quindi che $p_F(t)$ si fattorizza come:

$$p_F(t) = (\lambda - t)^k q(t),$$

dove $q(t) = |C - tI_{n-k}|$, il che mostra che la molteplicità della radice λ è almeno k (supera strettamente k se $q(\lambda) = 0$), e ciò significa che $m_a(\lambda) \geq m_g(\lambda)$. □

Osservazione 40.11. Conviene osservare che, stante il Teorema del rango, la molteplicità geometrica di un autovalore λ si può calcolare come segue:

$$m_g(\lambda) = n - rg(A - \lambda I_n),$$

dove A denota ancora la matrice associata all'endomorfismo in questione, rispetto ad una base scelta a piacimento.

Prima di enunciare il criterio principale di diagonalizzabilità, richiamiamo un'altra nozione della teoria dei polinomi. Un polinomio $p(t)$ non nullo a coefficienti in $\mathbb{K}$ è detto *completamente riducibile* se la somma $m_1 + \cdots + m_k$ delle molteplicità delle sue radici distinte $\lambda_1, \ldots, \lambda_k$ coincide con il suo grado. Per il Teorema di Ruffini, ciò equivale alla richiesta che, a meno di una costante non nulla, il polinomio sia prodotto di fattori tutti di primo grado del tipo $t - \lambda_i$, cioè alla condizione che $p(t)$ possa fattorizzarsi come segue:

$$p(t) = \alpha(t - \lambda_1)^{m_1} \cdots (t - \lambda_k)^{m_k} \tag{33}$$

dove $\alpha \in \mathbb{K}$, $\alpha \neq 0$, e per ogni $i \neq j$, $\lambda_i \neq \lambda_j$, e ciascun m_j è un intero positivo.

Naturalmente, in tal caso le radici di $p(t)$ sono esattamente i k scalari λ_i, ciascuna di molteplicità m_i.

Teorema 40.12. *(Criterio di diagonalizzabilità) Sia $F : V \to V$ un endomorfismo del $\mathbb{K}$-spazio vettoriale V di dimensione $n > 0$. Le seguenti proprietà sono equivalenti:*

a) *F è diagonalizzabile.*

b) i) *Il polinomio caratteristico di F è completamente riducibile;*
   ii) *per ogni autovalore $\lambda \in Sp(F)$ si ha $m_g(\lambda) = m_a(\lambda)$.*

La dimostrazione che *a)* $\Rightarrow$ *b)* è sostanzialmente un esercizio: si assuma che F sia diagonalizzabile. Allora esiste una base $\mathfrak{B}$ di V rispetto alla quale la matrice associata a F è diagonale del tipo

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

dove gli scalari λ_i, $i = 1, \dots, n$ sulla diagonale principali non sono necessariamente distinti. Utilizzando tale matrice per calcolare il polinomio caratteristico mediante la (32), si perviene immediatamente a:

$$p_F(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n),$$

il che prova che $p_F(t)$ è completamente riducibile. Inoltre, per il Teorema 40.7, lo spettro di F contiene tutti e soli i λ_i distinti tra loro, e la molteplicità algebrica di ogni autovalore è determinata da quante volte ciascun λ_i figura ripetuto sulla diagonale principale della matrice diagonale. Per controllare ii), si consideri l'autovalore λ_j e sia m la sua molteplicità; la sua molteplicità geometrica si ottiene calcolando $n - rg(A - \lambda_j I_n)$; abbiamo:

$$A - \lambda_j I_n = diag(\lambda_1 - \lambda_j, \dots, \lambda_n - \lambda_j),$$

e il rango di questa matrice coincide col numero dei suoi elementi non nulli, ovvero il numero delle differenze $\lambda_i - \lambda_j$ non nulle; siccome m conta le ripetizioni di λ_j, questo numero è esattamente $n - m$, per cui in definitiva $m_g(\lambda_j) = n - (n - m) = m$.

Alla dimostrazione dell'altra implicazione premettiamo il seguente

Lemma 40.13. *Si considerino $\lambda_1, \dots, \lambda_k$ autovalori distinti di F, con $k \geq 1$, e una sequenza di k vettori*

$$v_1, \dots, v_k$$

dove ciascun v_i appertiene all'autospazio $V_{\lambda_i}(F)$. Allora

$$v_1 + \cdots + v_k = 0 \Rightarrow v_1 = \cdots = v_k = 0.$$

Dimostrazione: Ragioniamo per induzione su k. Nel caso $k = 1$ non cè nulla da dimostrare. Supponiamo $k > 1$ e l'asserto vero per $k - 1$. Se

$$v_1 + \cdots + v_k = 0,$$

allora

$$v_1 + \cdots + v_{k-1} = -v_k,$$

per cui $v_1 + \cdots + v_{k-1}$ appartiene all'autospazio $V_{\lambda_k}(V)$. Quindi

$$F(v_1 + \cdots + v_{k-1}) = \lambda_k(v_1 + \cdots + v_{k-1}).$$

D'altra parte

$$F(v_1 + \cdots + v_{k-1}) = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}$$

e quindi per confronto si ottiene che:

$$\lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} = \lambda_k(v_1 + \cdots + v_{k-1}).$$

Ciò si riscrive

$$(\lambda_1 - \lambda_k)v_1 + \cdots + (\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Applicando l'ipotesi induttiva agli autovalori $\lambda_1, \dots, \lambda_{k-1}$, segue che:

$$(\lambda_i - \lambda_k)v_i = 0$$

per ogni $i \leq k-1$. Siccome λ_k è diverso dagli altri λ_i, consegue $v_i = 0$ per ogni $i \leq k-1$. Infine, dall'ipotesi $v_1 + \cdots + v_k = 0$ si ottiene anche $v_k = 0$. $\square$

Siamo ora in grado di dimostrare che le i) e ii) nel Teorema 40.12 garantiscono che F è diagonalizzabile. Assumiamo quindi i) e ii) e siano $\lambda_1, \ldots, \lambda_k$ gli autovalori distinti di F, ovvero le radici di $p_F(t)$. Per la i), posto $m_i := m_a(\lambda_i)$, abbiamo

$$m_1 + \cdots + m_k = n. \tag{34}$$

Inoltre, per la ii), per ogni $i = 1, \ldots, k$, ciascun autospazio $V_{\lambda_i}(F)$ ha dimensione m_i. Fissiamo una base $\mathfrak{B}_i$ di $V_{\lambda_i}(F)$, costituita dai vettori:

$$\mathfrak{B}_i = \{v_1^i, \ldots, v_{m_i}^i\}.$$

Affermiamo ora che la sequenza

$$\mathfrak{B} = \{v_1^1, \ldots, v_{m_1}^1, \ldots, v_1^k, \ldots, v_{m_k}^k\}$$

costituita da tutti i vettori delle basi $\mathfrak{B}_1, \ldots, \mathfrak{B}_k$ costituisce una base di V. Intanto $\mathfrak{B}$ è una sequenza di n vettori, stante la (34). Dunque basta provare che tali vettori sono linearmente indipendenti. Consideriamo a tal scopo una loro combinazione lineare nulla:

$$\alpha_1^1 v_1^1 + \cdots + \alpha_{m_1}^1 v_{m_1}^1 + \cdots + \alpha_1^k v_1^k + \cdots + \alpha_{m_k}^k v_{m_k}^k = 0.$$

Possiamo scrivere questa uguaglianza come:

$$v_1 + \cdots + v_k = 0,$$

dove abbiamo denotato con v_i il vettore $\alpha_1^i v_1^i + \cdots + \alpha_{m_i}^i v_{m_i}^i$, che appartiene all'autospazio $V_{\lambda_i}(F)$. Ora, siccome gli autovalori $\lambda_1, \ldots, \lambda_k$ sono tutti distinti, applicando il Lemma precedente otteniamo:

$$v_1 = 0, \ldots, v_k = 0.$$

Dunque, per ogni $i = 1, \ldots, k$:

$$\alpha_1^i v_1^i + \cdots + \alpha_{m_i}^i v_{m_i}^i = 0.$$

Infine, poichè i vettori $v_1^i, \ldots, v_{m_i}^i$, costituendo la base $\mathfrak{B}_i$, sono linearmente indipendenti, segue che per ogni fissato i, tutti gli scalari α_j^i sono nulli, e con ciò la dimostrazione è conclusa $\square$.

Osservazione 40.14. Conviene osservare che, dato $\lambda \in Sp(F)$, la condizione

$$m_g(\lambda) = m_a(\lambda)$$

è automaticamente soddisfatta nel caso in cui $m_a(\lambda) = 1$. Infatti abbiamo

$$1 \leq \mathtt{dim}(V_\lambda(F)) \leq m_a(\lambda) = 1,$$

perchè $V_\lambda(F) \neq \{0\}$. Diremo in tal caso che l'autovalore λ è *semplice*.

Ne consegue:

Corollario 40.15. *Un endomorfismo $F : V \to V$ che ammetta n autovalori distinti, $n = \mathtt{dim}(V)$, è sempre diagonalizzabile.*

41. Matrici diagonalizzabili

La nozione di diagonalizzabilità si applica anche a matrici, secondo la definizione seguente:

Definizione 41.1. Sia $\mathbb{K}$ un campo. Una matrice quadrata $A \in M_n(\mathbb{K})$, $n \geq 1$, si dice **diagonalizzabile** se è simile a una matrice diagonale.

Pertanto una matrice A è diagonalizzabile se e solo se esiste una matrice invertibile $M \in GL(n, \mathbb{K})$ tale che
$$M^{-1}AM = D,$$
con D diagonale.

Questa nozione risulta di fatto equivalente a quella introdotta per gli endomorfismi, in virtù del risultato che segue:

Proposizione 41.2. *Data una matrice $A \in M_n(\mathbb{K})$, si ha che:*
$$A \text{ è diagonalizzabile} \iff L_A : \mathbb{K}^n \to \mathbb{K}^n \text{ è diagonalizzabile.}$$

Dimostrazione:

Supponiamo che A sia diagonalizzabile; fissiamo $M \in GL(n, \mathbb{K})$ tale che $M^{-1}AM = D$ con D diagonale. Posto $M = (v_1 \cdots v_n)$, i vettori $v_1, \ldots, v_n$ di $\mathbb{K}^n$ costituiscono una base $\mathfrak{B}$ di $\mathbb{K}^n$ in quanto M ha rango massimo. Notiamo che, per costruzione, M coincide con la matrice di passaggio $M_{\mathfrak{B}_o, \mathfrak{B}}$. Allora abbiamo:
$$M_{\mathfrak{B}}(L_A) = M^{-1}M_{\mathfrak{B}_o}(L_A)M = M^{-1}AM = D,$$
e ciò mostra che L_A è diagonalizzabile. Il viceversa si prova in modo simile; omettiamo i dettagli. $\square$

Le nozioni di autovettore e autovalore si possono definire anche per matrici quadrate, in modo diretto, come segue:

Definizione 41.3. 1) Data una matrice $A \in M_n(\mathbb{K})$, un vettore $v \in \mathbb{K}^n$ si dice **autovettore** di A se $v \neq 0$ e
$$Av = \lambda v,$$
dove λ è uno scalare opportuno, detto autovalore associato a v.

2) Si chiama **autovalore** di A ogni scalare $\lambda \in \mathbb{K}$ per cui esiste un vettore non nullo $v \in \mathbb{K}^n$ tale che $Av = \lambda v$. L'autospazio corrispondente è $V_\lambda(A) = \{v \in \mathbb{K}^n : Av = \lambda v\}$.

3) Si chiama **polinomio caratteristico** di A, il polinomio $p_A(t) := det(A - tI_n)$.

Si osservi che richiedere che un vettore v di $\mathbb{K}^n$ sia autovettore di A equivale a richiedere che sia autovettore di L_A. Analogamente, gli autovalori di A sono esattamente gli autovalori di L_A, e così via.

In base alla discussione precedente, risulta in particolare che il problema di determinare una matrice invertibile M per cui risulti che $M^{-1}AM$ sia una matrice diagonale equivale a determinare una matrice $M = (v_1 \ \dots \ v_n)$ le cui colonne $v_1, \dots, v_n$ sono n autovettori linearmente indipendenti di A.

Evidentemente, tutti i risultati provati sopra riguardanti gli endomorfismi lineari, e il criterio di diagonalizzabilità, si applicano in modo ovvio e naturale anche alle matrici.

42. Spazi vettoriali Euclidei

In questa sezione iniziamo lo studio degli spazi vettoriali Euclidei: sono spazi vettoriali reali dotati di uno "strumento" aggiuntivo, detto prodotto scalare, adatto per effettuare misure, quali lunghezze e angoli, e mediante il quale è possibile trattare la nozione di "perpendicolarità".

D'ora in poi consideremo, salvo avviso contrario, solo spazi vettoriali reali.

Definizione 42.1. Si chiama *prodotto scalare* su uno spazio vettoriale reale V ogni applicazione

$$g : V \times V \to \mathbb{R}$$

avente i seguenti requisiti:

(1) g **è simmetrica**: $g(u, v) = g(v, u)$ per ogni $u, v \in V$;

(2) g **è bilineare**: $g(\lambda u + \mu v, w) = \lambda g(u, w) + \mu g(v, w)$
per ogni $u, v, w \in V$ e $\lambda, \mu \in \mathbb{R}$.

(3) g **è definita positiva**: per ogni $v \in V$, $v \neq 0$ si ha $g(v, v) > 0$.

Dato un prodotto scalare g, e due vettori u, v, il numero $g(u, v)$ prende il nome di *prodotto scalare di u e v* rispetto a g.

N. B. In letteratura spesso un prodotto scalare si denota anche con altri simboli, ad esempio con un punto $\cdot$ o mediante parentesi angolari $<\, , >$. Anche qui adottereremo talvolta una di queste varianti.

Notiamo che, stante la proprietà 1) di simmetria, ogni prodotto scalare soddisfa anche la proprietà seguente:

$$g(u, \lambda v + \mu w) = \lambda g(u, v) + \mu g(v, w)$$

che si ottiene direttamente dalla 2). Ciò giustifica la terminologia "bilineare": g è lineare in entrambi gli argomenti, nel senso che, fissato un vettore $u \in V$, entrambe le funzioni

$$v \in V \mapsto g(u, v) \in \mathbb{R}, \quad v \in V \mapsto g(v, u) \in \mathbb{R},$$

sono lineari.

Osservazione 42.2. La proprietà (3) si può riformulare come segue: per ogni v si ha $g(v,v) \geq 0$ e
$$g(v,v) = 0 \Rightarrow v = 0.$$

Osserviamo anche quanto segue riguardo il prodotto scalare del vettore nullo con gli altri vettori:

Proposizione 42.3. *Sia $g : V \times V \to \mathbb{R}$ un prodotto scalare. Allora per ogni $v \in V$ si ha:*
$$g(v, 0_V) = 0.$$

Dimostrazione: Infatti, appicando la bilinearità di g abbiamo:
$$g(v, 0_V) = g(v, 0_V + 0_V) = g(v, 0_V) + g(v, 0_V).$$
□

Definizione 42.4. Si chiama *spazio vettoriale Euclideo* ogni coppia (V, g) costituita da uno spazio vettoriale reale ed un fissato prodotto scalare su V.

Discutiamo alcuni esempi.

Definizione 42.5. Si chiama *prodotto scalare standard* su $\mathbb{R}^n$ $(n \geq 1)$, l'applicazione
$$g_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
definita come segue
$$g_0(x, y) := x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$
per ogni coppia di vettori $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$.

La definizione di g_0 può anche riscriversi facendo uso del prodotto matriciale righe per colonne:
$$g_0(x, y) = (x_1 \ x_2 \ \cdots \ x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = {}^t x \, y,$$
dove si conviene sempre di identificare ogni vettore di $\mathbb{K}^n$ con un vettore colonna (per cui ${}^t x$ è un vettore riga).

La verifica che g_0 è un prodotto scalare è piuttosto semplice. Notiamo che la simmetria deriva direttamente dalla definizione. Anche il fatto che g_0 è definita positiva, in quanto
$$g(x, x) = x_1^2 + \cdots + x_n^2$$
e tale quantità è sempre ≥ 0, e si annulla solo se $x_1 = \cdots = x_n = 0$, ovvero solo se $x = 0$. Riguardo la bilinearità, possiamo notare che dove conveniamo che ogni vettore di $\mathbb{R}^n$ è un vettore colonna Ciò premesso, la bilinearità segue direttamente

dalla caratterizzazione esplicita delle funzioni lineari $F : \mathbb{K}^n \to \mathbb{K}$; fissato un vettore y si ha infatti che

$$x \mapsto g_0(x, y) = y_1 x_1 + \cdots + y_n x_n$$

è una funzione lineare.

Il prodotto scalare standard $g_0(u, v)$ di due vettori di $\mathbb{R}^n$ si denoterà anche mediante il simbolo $\cdot$.

Esempio 42.6. È possibile munire lo spazio $M_n(\mathbb{R})$ delle matrici quadrate di ordine n di un prodotto scalare ponendo

$$g(A, B) := tr({}^tAB).$$

Qui $tr(A)$ denota la *traccia* di una matrice A quadrata, definita come la somma degli elementi sulla diagonale principale di A:

$$tr(A) := a_1^1 + \cdots + a_n^n.$$

Si osserva che

$$tr(A + B) = tr(A) + tr(B),\ tr(\lambda A) = \lambda tr(A),$$

per cui la funzione $tr : M_n(\mathbb{R}) \to \mathbb{R}$ è lineare.

Verifichiamo che tale applicazione $g : M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to \mathbb{R}$ soddisfa le (1)-(3).

(1) La verifica della simmetria è basata sul fatto che la traccia di una matrice coincide con la traccia della sua trasposta:

$$g(B, A) = tr({}^tBA) = tr({}^t({}^tBA)) = tr({}^tAB) = g(A, B).$$

(2) La bilinearità è conseguenza della linearità della funzione traccia e dell'operazione di trasposizione:

$$g(\lambda A + \mu B, C) = tr((\lambda {}^tA + \mu {}^tB)C) = \lambda\, tr({}^tAC) + \mu\, tr({}^tBC) = \lambda g(A, C) + \mu g(B, C).$$

(3) Per provare che g è definita positiva, possiamo osservare che, data una matrice $A \in M_n(\mathbb{R})$ e posto $C = {}^tAA = (c_j^i)$, allora l'elemento generico c_i^i sulla diagonale di C è dato dal prodotto righe per colonne tra la i-ma riga di ${}^tA$ e la i-ma colonna di A:

$$c_i^i = ({}^tA)^{(i)} A_{(i)}.$$

D'altra parte, rappresentando la trasposta di A per righe abbiamo:

$${}^tA = \begin{pmatrix} {}^tA_{(1)} \\ \vdots \\ {}^tA_{(n)} \end{pmatrix},$$

ovvero la i-ma riga della trasposta di A coincide con la i-ma colonna di A, ma riscritta come riga. Segue che c_i^i si può esprimere mediante il prodotto scalare standard:

$$c_i^i = {}^tA_{(i)} A_{(i)} = A_{(i)} \cdot A_{(i)}.$$

In conclusione

$$g(A, A) = tr(C) = \sum_{i=1}^{n} A_{(i)} \cdot A_{(i)}.$$

Dunque $g(A, A) \geq 0$, e se $g(A, A) = 0$, allora per ogni colonna di A risulta $A_{(i)} \cdot A_{(i)} = 0$ e quindi $A_{(i)} = 0$; pertanto $A = 0$.

ESEMPIO 42.7. Ogni spazio vettoriale reale di dimensione finita V si può munire di un prodotto scalare, che si può costruire a partire da una base assegnata $\mathfrak{B}$. Si pone

$$g(u, v) := {}^t x\, y$$

dove x e y sono i vettori delle componenti di u e v rispetto alla base in questione. In altri termini:

$$g(u, v) = g_o(\psi_{\mathfrak{B}}(u), \psi_{\mathfrak{B}}(v)).$$

Si lascia al lettore la verifica della bilinearità di g (si sfrutti la linearità di $\psi_{\mathfrak{B}}$). La verifica che g è definita positiva poggia sul fatto che $\psi_{\mathfrak{B}}$ è un isomorfismo, in particolare sulla sua ingettività: infatti, per ogni $v \in V$ se $g(v, v) = 0$, allora

$$g_o(\psi_{\mathfrak{B}}(v), \psi_{\mathfrak{B}}(v)) = 0,$$

da cui $\psi_{\mathfrak{B}}(v) = 0$ e quindi $v = 0$.

ESEMPIO 42.8. L'esempio precedente consiste nel "trasportare" il prodotto scalare standard g_0 di $\mathbb{R}^n$ su V. Questo procedimento può generalizzarsi come segue. Sia (W, g) uno spazio vettoriale Euclideo e sia

$$F : V \to W$$

un monomorfismo dove V è un qualsiasi spazio vettoriale reale. Allora g induce un prodotto scalare g su V, denotato con F^*g e definito come segue:

$$(F^*g)(u, v) := g(F(u), F(v)), \quad \forall u, v \in V.$$

Tale prodotto scalare si chiama il *pull-back* di g mediante F. Si lasciano le verifiche delle proprietà di F^*g come esercizio. Anche in questo caso il fatto che F è un monomorfismo è essenziale per controllare che g è definita positiva:

$$(F^*g)(v, v) = 0 \Rightarrow g(F(v), F(v)) = 0 \Rightarrow F(v) = 0 \Rightarrow v = 0.$$

L'esempio precedente corrisponde al pull-back ψ^*g_o del prodotto scalare standard su $\mathbb{R}^n$ mediante l'isomorfismo $\psi_{\mathfrak{B}} : V \to \mathbb{R}^n$.

ESEMPIO 42.9. Se (V, g) è uno spazio vettoriale Euclideo, ogni sottospazio vettoriale $W \subset V$ è esso stesso uno spazio vettoriale Euclideo in modo canonico: la restrizione

$$g|_{W \times W} : W \times W \to \mathbb{R}$$

a W del prodotto scalare di V è manifestamente un prodotto scalare anche su W.

43. Basi ortonormali

Definizione 43.1. Sia (V, g) uno spazio vettoriale Euclideo. Si chiama *norma* o *lunghezza* di un vettore v il numero reale non negativo:

$$||v||_g := \sqrt{g(v,v)}.$$

La norma di un vettore si denota anche più semplicemente con $||v||$ se è chiaro dal contesto rispetto a quale prodotto scalare g viene calcolata.

Si noti che, essendo g definita positiva, si ha:

$$||v||_g = 0 \iff v = 0.$$

Inoltre, per ogni scalare λ risulta:

$$||\lambda v||_g = |\lambda| ||v||_g. \tag{35}$$

Infatti: $|\lambda v||_g = \sqrt{g(\lambda v, \lambda v)} = \sqrt{\lambda^2 g(v,v)} = |\lambda| ||v||_g$.

Definizione 43.2. Si chiama *versore* o *vettore unitario* ogni vettore v tale che $||v||_g = 1$.

Notiamo che ogni vettore non nullo v può essere riscalato in modo da ottenere un versore: infatti è sufficiente considerare il vettore:

$$\frac{v}{||v||_g}.$$

Definizione 43.3. Sia (V, g) uno spazio vettoriale Euclideo di dimensione $n \geq 1$. Una sequenza di n vettori

$$v_1, \ldots, v_n$$

si chiama *base ortonormale* se ogni v_i è un versore e

$$g(v_i, v_j) = 0 \text{ per ogni } i, j \text{ tali che } i \neq j,$$

ovvero i vettori della sequenza sono versori mutuamente ortogonali.

Ogni base ortonormale è effettivamente una base di V: ciò è conseguenza del risultato seguente, che ne garantisce la lineare indipendenza (pertanto trattandosi di una sequenza di n vettori è una base):

Proposizione 43.4. *Sia*

$$v_1, \ldots, v_k, \ k \geq 1$$

una sequenza di vettori tutti non nulli e mutuamente ortogonali. Allora $v_1, \ldots, v_k$ sono linearmente indipendenti.

Dimostrazione: Si consideri una combinazione lineare banale dei vettori in questione:

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0 \tag{$*$}$$

dove i λ_i sono scalari. Fissiamo j, $1 \leq j \leq k$; considerando il prodotto scalare di ambo i membri nella (*) con v_j segue

$$g(\lambda_1 v_1 + \cdots + \lambda_k v_k, v_j) = 0.$$

Stante la bilinearità del prodotto scalare, il primo membro si sviluppa come segue:

$$\lambda_1 g(v_1, v_j) + \cdots + \lambda_k g(v_k, v_j) = 0,$$

ma per l'ipotesi sui vettori in questione tutti i prodotti scalari $g(v_i, v_j)$ per $i \neq j$ sono nulli, sicchè l'uguaglianza precedente si riduce a

$$\lambda_j g(v_j, v_j) = 0$$

da cui, essendo $v_j \neq 0$, che garantisce che $g(v_j, v_j) \neq 0$, si conclude

$$\lambda_j = 0.$$

□

Le condizioni che definiscono una base ortonormale si può formulare in modo sintetico come segue:

$$g(v_i, v_j) = \delta^i_j,$$

dove δ^i_j è il cosiddetto *simbolo di Kronecker*, che per definizione è uno scalare che vale 1 quando $i = j$ e 0 altrimenti. Proveremo che ogni spazio vettoriale Euclideo è dotato di basi ortonormali.

Notiamo che la base canonica $\{e_1, \dots, e_n\}$ è un esempio di base ortonormale, rispetto al prodotto scalare standard di $\mathbb{R}^n$.

Discutiamo prima alcune proprietà rilevanti di tali basi.

44. Componenti di un vettore rispetto a basi ortonormali

Le basi ortonormali costituiscono un ottimo strumento di calcolo, in quanto permettono di determinare in modo immediato le componenti di un vettore: si ha infatti

Proposizione 44.1. *Sia $\mathfrak{B} = \{v_1, \dots, v_n\}$ una base ortonormale di uno spazio vettoriale Euclideo (V, g) di dimensione n. Allora per ogni vettore v si ha:*

$$v = g(v, v_1)v_1 + \cdots + g(v, v_n)v_n.$$

Pertanto le componenti di un vettore v sono semplicemente i prodotti scalari tra esso ed i vettori della base. Questo fatto generalizza la formula (5) già nota nel caso della base canonica di $\mathbb{R}^n$.

Dimostrazione: Dette $(x_1, \dots, x_n)$ le componenti di v rispetto alla base in questione abbiamo che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Si tratta di mostrare che ogni x_i coincide con $g(v, v_i)$. Ciò segue calcolando direttamente il prodotto scalare:

$$g(v, v_i) = g(\sum_{k=1}^{n} x_k v_k, v_i) = \sum_{k=1}^{n} x_k g(v_k, v_i) = x_i g(v_i, v_i) = x_i.$$

□

Anche il prodotto scalare tra due vettori si calcola in modo semplice disponendo di una base ortonormale; il risultato seguente mostra che effettuare un prodotto scalare $g(u, v)$ consiste nell'effettuare il prodotto scalare standard tra i vettori $x, y \in \mathbb{R}^n$ delle componenti di u e v.

Proposizione 44.2. *Sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base ortonormale di uno spazio vettoriale Euclideo (V, g) di dimensione n e siano u e v due vettori. Siano $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ i vettori delle componenti di u e v rispettivamente. Allora si ha:*

$$g(u, v) = \sum_{i=1}^{n} x_i y_i = x \cdot y, \tag{36}$$

dove $\cdot$ è il prodotto scalare standard su $\mathbb{R}^n$.

Dimostrazione: Infatti, è sufficiente espandere il calcolo di $g(u, v)$, in modo simile a quanto visto nelle argomentazioni fatte in precedenza, tenendo conto della bilinearità di g:

$$g(u, v) = g(\sum_{i=1}^{n} x_i v_i, v) = \sum_{i=1}^{n} x_i g(v_i, v) = \sum_{i=1}^{n} x_i y_i.$$

L'ultimo passaggio è giustificato dalla proposizione precedente applicata al vettore v, che garantisce che $g(v_i, v) = y_i$. □

Successivamente interpreteremo questo fatto mediante il concetto di isometria tra spazi vettoriali Euclidei.

45. Esistenza di basi ortonormali

Proviamo ora che ogni spazio vettoriale Euclideo (di dimensione finita) ammette effettivamente basi ortonormali.

Teorema 45.1. *Sia (V, g) uno spazio vettoriale Euclideo di dimensione $n \geq 1$. Si fissi una base*

$$\mathfrak{B} = \{v_1, \ldots, v_n\}.$$

Allora esiste una base ortonormale

$$\mathfrak{B}' = \{v'_1, \ldots, v'_n\}$$

tale che, per ogni k, $1 \leq k \leq n$ si abbia:

$$L(v_1, \ldots, v_k) = L(v'_1, \ldots, v'_k).$$

Dimostrazione: Procederemo per induzione su n. Nel caso $n = 1$, per costruire la base richiesta è sufficiente porre

$$v'_1 := \frac{v_1}{||v_1||_g}.$$

Supponiamo quindi $n > 1$ e l'asserto vero per spazi vettoriali Euclidei di dimensione $n-1$. Consideriamo quindi il sottospazio $W := L(v_1, \ldots, v_{n-1})$ di V; esso ha dimensione $n-1$ e $\{v_1, \ldots, v_{n-1}\}$ è una sua base. Per l'ipotesi induttiva, W ammette una base ortonormale

$$v'_1, \ldots, v'_{n-1}$$

tale che

$$L(v_1, \ldots, v_k) = L(v'_1, \ldots, v'_k)$$

per ogni $k \leq n-1$. Si tratta quindi di completare tale insieme di vettori ad una base ortonormale di V. A questo scopo, consideriamo il seguente vettore:

$$u_n := v_n - g(v_n, v'_1)v'_1 - \cdots - g(v_n, v'_{n-1})v'_{n-1} = v_n - \sum_{i=1}^{n-1} g(v_n, v'_i)v'_i.$$

Risulta che $u_n \neq 0$: infatti se fosse $u_n = 0$, allora avremmo

$$v_n = g(v_n, v'_1)v'_1 + \cdots + g(v_n, v'_{n-1})v'_{n-1}$$

da cui

$$v_n \in L(v'_1, \ldots, v'_{n-1}) = L(v_1, \ldots, v_{n-1})$$

contro la lineare indipendenza dei vettori della base $\mathfrak{B}$.

Inoltre risulta che u_n è ortogonale a tutti vettori $v'_1, \ldots, v'_{n-1}$: fissato infatti uno di questi, v'_j, risulta:

$$\begin{aligned} g(u_n, v'_j) &= g(v_n - \sum_{i=1}^{n-1} g(v_n, v'_i)v'_i, v'_j) = \\ &= g(v_n, v'_j) - g(v_n, v'_j)g(v'_j, v'_j) = \\ &= g(v_n, v'_j) - g(v_n, v'_j) = 0, \end{aligned}$$

dove si è ancora sfruttata la bilinearità del prodotto scalare e il fatto che l'unico prodotto scalare $g(v'_i, v'_j)$ non nullo è quello corrispondente a $i = j$.

A questo punto è sufficiente porre

$$v'_n := \frac{u_n}{||u_n||_g}$$

ottenendo così una base ortonormale $v'_1, \ldots, v'_n$ che ha il requisito richiesto. $\square$

La dimostrazione appena discussa suggerisce un algoritmo per costruire una base ortonormale $v'_1, \dots, v'_n$ a partire da una base assegnata $v_1, \dots, v_n$, detto **algoritmo di Gram-Schmidt**. La sequenza di vettori $v'_1, \dots, v'_n$ si può definire in modo ricorsivo come segue:

Per $k = 1$ si pone $v'_1 := \frac{v_1}{||v_1||_g}$.

Per $k > 1$ si pone $v'_k := \frac{u_k}{||u_k||_g}$, dove $u_k = v_k - \sum_{i=1}^{k-1} g(v_k, v'_i)\, v'_i$.

Esempio 45.2. Determiniamo una base ortonormale $\{v'_1, v'_2, v'_3\}$ di $\mathbb{R}^3$ partendo dalla base $\mathfrak{B} = \{v_1 = (1,0,1), v_2 = (0,1,-1), v_3 = (0,3,0)\}$. Applicando passo dopo passo l'algoritmo di Gram-Schmidt possiamo costruire i tre vettori v'_i in successione come segue:

$$v'_1 = \frac{v_1}{||v_1||} = \frac{1}{\sqrt{2}}(1,0,1) = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}).$$

$$u_2 = v_2 - (v_2 \cdot v'_1)v'_1 = (0,1,-1) + \frac{\sqrt{2}}{2}v'_1 = (\frac{1}{2}, 1, -\frac{1}{2})$$

$$v'_2 = \frac{u_2}{||u_2||} = (\frac{1}{2}\sqrt{\frac{2}{3}}, \sqrt{\frac{2}{3}}, -\frac{1}{2}\sqrt{\frac{2}{3}})$$

$$u_3 = v_3 - (v_3 \cdot v'_1)v'_1 - (v_3 \cdot v'_2)v'_2 = (0,3,0) - 0 - (1,2,-1) = (-1,1,1)$$

$$v'_3 = \frac{u_3}{||u_3||} = (-\frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}).$$

46. Matrici ortogonali

Definizione 46.1. Una matrice quadrata $A \in M_n(\mathbb{R})$ si dice *ortogonale* se

$$ {}^tAA = I_n. \tag{37}$$

Per definizione quindi I_n è ortogonale.

Proposizione 46.2. *Ogni matrice ortogonale A è invertibile e si ha*

$$A^{-1} = {}^tA.$$

Inoltre $\det(A) = 1$ *oppure* $\det(A) = -1$*. Infine se A è ortogonale, tale è la trasposta di A.*

Infatti dalla relazione (37) segue

$$\det({}^tA)\det(A) = 1$$

per il Teorema di Binet e quindi

$$\det(A)^2 = 1.$$

Da qui l'affermazione sul determinante di A, e quindi l'invertibilità di A. Moltiplicando ambo i membri di (37) a destra per la matrice inversa di A si ottiene che

${}^tA = A^{-1}$. In particolare, ciò garantisce anche che vale l'altra relazione $A\,{}^tA = I_n$ e ciò, in base alla definizione, dice che anche ${}^tA$ è una matrice ortogonale, perchè $A = {}^t({}^tA)$.

Proposizione 46.3. *L'insieme $O(n)$ di tutte le matrici ortogonali di ordine n è un gruppo rispetto all'operazione di moltiplicazione righe per colonne.*

La dimostrazione è semplice: abbiamo $I_n \in O(n)$, per cui $O(n)$ ammette l'elemento neutro rispetto alla moltiplicazione tra matrici. Abbiamo già stabilito che l'inversa di una matrice ortogonale è anch'essa ortogonale. Si lascia al lettore la verifica del fatto che il prodotto di due matrici ortogonali è dello stesso tipo.

Sussiste la seguente utile interpretazione geometrica della condizione (37):

Proposizione 46.4. *Per ogni matrice $A \in M_n(\mathbb{R})$, le proprietà seguenti sono equivalenti:*
a) A è ortogonale.
b) Le colonne $A_{(1)}, \ldots, A_{(n)}$ costituiscono una base ortonormale di $(\mathbb{R}^n, \cdot)$
c) Le righe $A^{(1)}, \ldots, A^{(n)}$ costituiscono una base ortonormale di $(\mathbb{R}^n, \cdot)$.

Dimostrazione: Basta provare che $a)$ e $b)$ sono equivalenti, perchè A è ortogonale se e solo se lo è la trasposta ${}^tA$. L'elemento c^i_j di posto (i, j) della matrice prodotto $C = {}^tAA$ concide con il prodotto tra la i-ma riga di ${}^tA$ e la j-ma colonna di A. Quindi:

$$c^i_j = ({}^tA)^{(i)}\, A_{(j)} = {}^t(A_{(i)})A_{(j)} = A_{(i)} \cdot A_{(j)}.$$

Osservando che $I_n = (\delta^i_j)$, consegue che ${}^tAA = I_n$ se e solo se $c^i_j = \delta^j$ ovvero se e solo se

$$A_{(i)} \cdot A_{(j)} = \delta^i_j,$$

ma questa è esattamente la condizione che $A_{(1)}, \ldots, A_{(n)}$ sia una base ortonormale di $\mathbb{R}^n$. $\square$

Esempio 46.5. Usando la caratterizzazione di cui sopra, non è difficile determinare tutte le matrici ortogonali di ordine 2. Risulta:

$$O(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\} \cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}.$$

Studiamo ora le matrice di passaggio tra basi ortormali in uno spazio vettoriale Euclideo qualsiasi.

Proposizione 46.6. *Sia (V, g) uno spazio vettoriale Euclideo di dimensione $n \geq 1$ e sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una fissata base ortonormale. Sia $\mathfrak{B}' = \{v'_1, \ldots, v'_n\}$ un'altra base di V. Risulta allora:*

$$\mathfrak{B}' \text{ è una base ortonormale} \iff M_{\mathfrak{B},\mathfrak{B}'} \in O(n).$$

Dimostrazione: Denotata con A la matrice di passaggio $M_{\mathfrak{B},\mathfrak{B}'}$, abbiamo che le colonne di A sono:

$$A = (\psi_{\mathfrak{B}}(v'_1) \cdots \psi_{\mathfrak{B}}(v'_n)).$$

Ora, $\mathfrak{B}'$ è base ortonormale se e solo se

$$g(v'_i, v'_j) = \delta^i_j,$$

che si può riscrivere, utilizzando ancora la (39):

$$\psi_{\mathfrak{B}}(v'_i) \cdot \psi_{\mathfrak{B}}(v'_j) = \delta^i_j$$

e ciò equivale alla condizione che le colonne di A formino una base ortonormale di $\mathbb{R}^n$, ovvero che $A \in O(n)$. $\square$

Osservazione 46.7. Date due basi ortonormali $\mathfrak{B}$ e $\mathfrak{B}'$, e detta A la matrice di passaggio da $\mathfrak{B}$ a $\mathfrak{B}'$, ricordiamo la relazione

$$x = Ax'$$

che lega i vettori x e x' delle coordinate di uno stesso vettore v rispetto alle due basi. Qui $x \in \mathbb{R}^n$ è il vettore delle coordinate di v nella base "vecchia", mentre x' si riferisce alla base "nuova" $\mathfrak{B}'$. Nel caso generale, quando sono coinvolte basi qualsiasi, per ricavare le coordinate "nuove" x' note le x, occorre determinare l'inversa A^{-1} di A, in quanto:

$$x' = A^{-1}x.$$

Nel caso specifico, tale operazione è immediata, perchè, essendo A ortogonale, sappiamo che $A^{-1} = {}^tA$ e si ottiene quindi la relazione più conveniente:

$$x' = {}^tAx.$$

47. Isometrie lineari

In questo paragrafo denotiamo con (V, g) e (W, g') due spazi vettoriali Euclidei della stessa dimensione $n \geq 1$.

Definizione 47.1. Diremo che un'applicazione lineare $F : (V, g) \to (W, g')$ è un'*isometria lineare*, o *trasformazione ortogonale*, oppure che è un *operatore unitario* se conserva i prodotti scalari, nel senso che:

$$g(u, v) = g'(F(u), F(v)) \tag{38}$$

per ogni $u, v \in V$.

Esempio 47.2. Fissata una base ortonormale $\mathfrak{B} = \{v_1, \dots, v_n\}$ di uno spazio vettoriale Euclideo (V, g), l'isomorfismo ad essa associato:

$$\psi_{\mathfrak{B}} : V \to \mathbb{R}^n$$

è un'isometria lineare tra (V, g) e $(\mathbb{R}^n, \cdot)$, dove $\cdot$ è il prodotto scalare standard.

Si tratta infatti di una rilettura della (36):

$$g(u, v) = x \cdot y = \psi_{\mathfrak{B}}(u) \cdot \psi_{\mathfrak{B}}(v). \tag{39}$$

Utilizzando questo esempio otteniamo il seguente risultato di classificazione:

Teorema 47.3. *Ogni spazio vettoriale Euclideo (V, g) n-dimensionale è isometrico a $(\mathbb{R}^n, g_0)$.*

Discutiamo ora un'importante caratterizzazione delle isometrie lineari:

Teorema 47.4. *Sia $F : (V, g) \to (W, g')$ un'applicazione lineare tra spazi vettoriali della stessa dimensione $n \geq 1$. Sia $\mathfrak{B} = \{v_1, \dots, v_n\}$ una base ortonormale di V. Sono equivalenti:*

a) F è un'isometria.

b) $\{F(v_1), \dots, F(v_n)\}$ è base ortonormale di W.

Dimostrazione:

a) $\Rightarrow$ *b)* È immediata perchè F conserva il prodotto scalare:

$$g'(F(v_i), F(v_j)) = g(v_i, v_j) = \delta^i_j.$$

b) $\Rightarrow$ *a)* Supponiamo che $\{F(v_1), \dots, F(v_n)\}$ sia una base ortonormale, che denotiamo con $\mathfrak{C}$. Osserviamo ora che per ogni $v \in V$, risulta che v e $F(v)$ hanno le stesse componenti rispetto alle due basi $\mathfrak{B}$ e $\mathfrak{C}$: ciò si può verificare direttamente usando la linearità di F: se v ha componenti $(x_1, \dots, x_n)$ allora

$$F(v) = F(x_1 v_1 + \cdots + x_n v_n) = x_1 F(v_1) + \cdots + x_n F(v_n).$$

Ciò premesso, la tesi segue subito dalla (36) applicata sia in (V, g) che in (W, g'): entrambi i prodotti scalari $g(u, v)$ e $g'(F(u), F(v))$ sono dati dal prodotto scalare standard $x \cdot y$ dei vettori delle componenti di u e v. $\square$

Corollario 47.5. *Ogni isometria lineare $F : (V, g) \to (W, g')$ è un isomorfismo.*

Ciò è conseguenza immediata del Teorema 32.2 di caratterizzazione degli isomorfismi. Notiamo anche che, in base alla stessa caratterizzazione, l'inversa F^{-1} di un'isometria lineare F è anch'essa un'isometria lineare; infatti, se F trasforma la base ortonormale $\{v_1, \dots, v_n\}$ nella base ortonormale $\{F(v_1), \dots, F(v_n)\}$, allora F^{-1} trasforma quest'ultima nella prima.

48. Relazione tra isometrie lineari e matrici ortogonali

Teorema 48.1. *Sia $F : (V, g) \to (W, g')$ un'applicazione lineare tra spazi vettoriali Euclidei della stessa dimensione $n \geq 1$. Siano $\mathfrak{B}$ e $\mathfrak{C}$ due basi ortonormali rispettivamente di V e di W. Allora:*

$$F \text{ è un'isometria lineare} \iff M^{\mathfrak{C}}_{\mathfrak{B}}(F) \in O(n).$$

Dimostrazione: Supposto che $\mathfrak{B} = \{v_1, \ldots, v_n\}$, abbiamo che le colonne della matrice $A := M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ sono gli n vettori

$$\psi_{\mathfrak{C}}(F(v_1)), \ldots, \psi_{\mathfrak{C}}(F(v_n))$$

di $\mathbb{R}^n$. Ora, in base al Teorema 47.4, F è un'isometria lineare se e solo se se i vettori $\{F(v_1), \ldots, F(v_n)\}$ costituiscono una base ortonormale di W, e ciò è vero se e solo se i vettori $\{\psi_{\mathfrak{C}}(F(v_1)), \ldots, \psi_{\mathfrak{C}}(F(v_n))\}$ costituiscono una base ortonormale di $\mathbb{R}^n$, perchè sia $\psi_{\mathfrak{C}} : W \to \mathbb{R}^n$ che la sua inversa sono isometrie lineari, e questo accade se e solo se $A \in O(n)$. $\square$

Corollario 48.2. *Sia $A \in M_n(\mathbb{R})$. Allora $A \in O(n)$ se e solo se:*

$$(Ax) \cdot (Ay) = x \cdot y \quad \forall x, y \in \mathbb{R}^n. \tag{40}$$

Infatti, poichè A coincide con la matrice $M_{\mathfrak{B}_o}(L_A)$ associata a $L_A : (\mathbb{R}^n, \cdot) \to (\mathbb{R}^n, \cdot)$ rispetto alla base canonica, che è ortonormale, abbiamo che $A \in O(n)$ se e solo se L_A è un'isometria lineare, il che accade se e solo se sussiste la (40).

49. Ortogonalità in uno spazio vettoriale Euclideo

Il concetto di perpendicolarità tra vettori si estende a sottospazi in modo naturale, come segue:

Definizione 49.1. Sia (V, g) uno spazio vettoriale Euclideo e sia $W \subset V$ un sottospazio vettoriale. Si chiama *complemento ortogonale* di W (o semplicemente *ortogonale di W*) l'insieme

$$W^\perp := \{v \in V \,|\, g(v, w) = 0 \text{ per ogni } w \in W\},$$

ovvero l'insieme di tutti i vettori che sono perprendicolari a *tutti* i vettori del sottospazio W.

Osservazione 49.2. Se $W = L(w_1, \ldots, w_k)$ allora risulta:

$$W^\perp = \{v \in V | g(v, w_i) = 0 \;\; \forall i = 1, \ldots, k\}.$$

In altri termini, per controllare che un vettore v appartenga a $W^\perp$ è sufficiente controllare che esso sia perpendicolare a tutti i vettori di un fissato sistema di generatori. Questo fatto è utile perchè riconduce una verifica che a priori coinvolge infiniti vettori ad un numero finito di casi.

La giustificazione di ciò coinvolge come al solito la bilinearità di g: se un vettore v è ortogonale a tutti i w_i, allora per ogni $w \in W$ si ha che v è ortogonale a w perchè w è una combinazione lineare $w = x_1 w_1 + \cdots + x_k w_k$ e quindi

$$g(v, w) = \sum_{i=1}^{k} x_i g(v, w_i) = 0.$$

Proposizione 49.3. *Per ogni sottospazio W di uno spazio vettoriale Euclideo (V, g) si ha che $W^\perp$ è un sottospazio vettoriale di V.*

Dimostrazione: Notiamo che evidentemente $0 \in W^\perp$, per cui tale insieme non è vuoto. Considerati due vettori u e v di $W^\perp$ allora, considerato un vettore arbitrario $w \in W$, si ha:

$$g(u + v, w) = g(u, w) + g(v, w) = 0 + 0 = 0$$

il che mostra che anche $u + v \in W^\perp$. Similmente, per ogni scalare λ e per ogni $u \in W^\perp$ risulta:

$$g(\lambda u, w) = \lambda g(u, w) = \lambda 0 = 0,$$

per cui anche λu appartiene a $W^\perp$. □

Osservazione 49.4. Notiamo i casi estremi: $\{0_V\}^\perp = V$ e $V^\perp = \{0_V\}$ corrispondenti rispettivamente a $W = \{0_V\}$ e $W = V$.

Nel secondo caso si utilizza il fatto che se se un vettore $v \in V$ è ortogonale a tutti i vettori di V, allora in particolare $g(v, v) = 0$, e quindi l'unica possibilità è che $v = 0$.

Teorema 49.5. *Sia W un sottospazio vettoriale dello spazio vettoriale Euclideo (V, g). Si assuma V di dimensione finita. Allora si ha:*

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

Dimostrazione: Risulta innanzitutto che:

$$W \cap W^\perp = \{0\},$$

in quanto se w è un vettore che appartiene sia a W che a $W^\perp$, allora necessariamente $g(w, w) = 0$ e quindi $w = 0$. Resta quindi da provare che $V = W + W^\perp$. Osserviamo che possiamo supporre $W \neq \{0\}$, perchè $\{0\}^\perp = V$ e quindi se il sottospazio è banale, l'asserto è senz'altro vero.

Fissiamo allora una base ortonormale $\mathfrak{B} = \{w_1, \ldots, w_k\}$ di W (rispetto alla restrizione del prodotto scalare g a W), dove $k = \mathtt{dim}(W) \geq 1$. Dato un vettore qualsiasi $v \in V$, vogliamo provare che v si può decomporre come:

$$v = w + u \tag{41}$$

con $w \in W$ e $u \in W^{\perp}$ opportuni vettori. Costruiamo dapprima il vettore w. Consideriamo a tal fine il seguente vettore di W:

$$w := g(v, w_1)w_1 + \cdots + g(v, w_k)w_k.$$

Questa scelta è di fatto obbligata, ed è suggerita dal fatto che, se vale la (41), allora

$$g(v, w_i) = g(w + u, w_i) = g(w, w_i),$$

e quindi le componenti del vettore w rispetto alla base assegnata devono necessariamente coincidere con i prodotti scalari $g(v, w_i)$. Poniamo poi

$$u := v - w.$$

Per concludere la dimostrazione, occorre provare che u appartiene a $W^{\perp}$. Per verificare questo, è sufficiente controllare che w è ortogonale a ciascuno dei vettori della base $\mathfrak{B}$. In effetti, per ogni $j \in \{1, \ldots, k\}$ abbiamo:

$$\begin{aligned} g(u, w_j) &= g(v, w_j) - g(\sum_{i=1}^{k} g(v, w_i)w_i, w_j) = g(v, w_j) - \sum_{i=1}^{k} g(v, w_i)g(w_i, w_j) \\ &= g(v, w_j) - g(v, w_j)g(w_j, w_j) \\ &= g(v, w_j) - g(v, w_j) = 0. \end{aligned}$$

□

Come conseguenza del Teorema precedente, risulta che fissato un sottospazio W di uno spazio vettoriale Euclideo (V, g), ogni vettore v si decompone in modo *unico* come

$$v = w + u,$$

dove $w \in W$ e $u \in W^{\perp}$. Il vettore w prende il nome di **proiezione ortogonale** di v su W.

Applicando la formula di Grassmann consegue:

Corollario 49.6. *Per ogni sottospazio W di uno spazio vettoriale Euclideo (V, g) si ha:*

$$\mathtt{dim}(W^{\perp}) = \mathtt{dim}(V) - \mathtt{dim}(W).$$

Osservazione 49.7. Valgono inoltre i seguenti fatti:

a) Per ogni sottospazio W si ha $(W^{\perp})^{\perp} = W$.

b) Dati due sottospazi U e W si ha: $U \subset W \Rightarrow W^{\perp} \subset U^{\perp}$.

Per giustificare a), basta notare che l'inclusione $W \subset (W^{\perp})^{\perp}$ è di semplice verifica in quanto se $w \in W$ e $u \in W^{\perp}$, allora certamente $g(w, u) = 0$ (ogni vettore di W è automaticamente ortogonale a ogni vettore di $W^{\perp}$). Ma ciò è sufficiente per concludere che i due spazi coincidono, in quanto essi hanno la stessa dimensione; infatti:

$$\mathtt{dim}((W^{\perp})^{\perp}) = \mathtt{dim}(V) - \mathtt{dim}(W^{\perp}) = \mathtt{dim}(V) - (\mathtt{dim}(V) - \mathtt{dim}(W)) = \mathtt{dim}(W).$$

Riguardo b), assunto che $U \subset W$, considerati $v \in W^{\perp}$ e $u \in U$, si ha che $g(u,v) = 0$ perchè u è anche un vettore di W.

50. La disuguaglianza di Cauchy-Schwarz

Proposizione 50.1. *(Formula di polarizzazione) Siano u, v vettori arbitrari di uno spazio vettoriale Euclideo (V, g). Si ha:*

$$||u+v||^2 = ||u||^2 + 2g(u,v) + ||v||^2. \tag{42}$$

La dimostrazione si ottiene direttamente sviluppando $g(u+v, u+v)$, tenendo conto della bilinearità e simmetria di g:

$$||u+v||^2 = g(u+v, u+v) = g(u,u)+g(u,v)+g(v,u)+g(v,v) = ||u||^2+2g(u,v)+||v||^2.$$

Teorema 50.2. *(disuguaglianza di Cauchy-Schwarz). Siano u, v vettori arbitrari di uno spazio vettoriale Euclideo (V, g). Allora si ha:*

$$|g(u,v)| \leq ||u||||v||. \tag{43}$$

Dimostrazione: La disuguaglianza da provare si può riformulare come:

$$g(u,v)^2 \leq ||u||^2||v||^2.$$

Essa è banalmente un'uguaglianza se uno dei due vettori è il vettore nullo; supponiamo quindi entrambi i vettori diversi da zero. Studiamo dapprima il caso in cui $v = tu$, per un opportuno scalare t. Anche in questo caso la (43) è un'uguaglianza, infatti:

$$g(u,v)^2 = t^2 g(u,u)^2 = t^2||u||^4 = ||u||^2||tu||^2 = ||u||^2||v||^2.$$

Supponiamo ora che $v \notin L(u)$. Allora, per ogni $t \in \mathbb{R}$ si ha

$$||tu+v|| > 0$$

in quanto $tu + v \neq 0$. Quindi

$$t^2||u||^2 + 2tg(u,v) + ||v||^2 > 0$$

per ogni $t \in \mathbb{R}$. Dunque l'equazione di secondo grado in t:

$$||u||^2t^2 + 2g(u,v)t + ||v||^2 = 0$$

non ammette radici reali; considerandone il discriminante dev'essere $\frac{\Delta}{4} = g(u,v)^2 - ||u||^2||v||^2 < 0$ e quindi la tesi. $\square$

La disuguaglianza di Cauchy-Schwarz permette di introdurre in modo analitico la nozione di angolo tra vettori:

Definizione 50.3. Siano u e v vettori non nulli di uno spazio vettoriale Euclideo (V, g). Si chiama **angolo convesso** tra u e v l'unico scalare $\theta \in [0, \pi]$ tale che:

$$\cos\theta = \frac{g(u,v)}{||u||||v||}.$$

Tale definizione ha senso perchè, in virtù di (43), abbiamo $-1 \leq \frac{g(u,v)}{||u||||v||} \leq 1$, e la funzione $\cos : [0, \pi] \to [-1, 1]$ è bigettiva.

Notiamo ad esempio che u e v sono ortogonali tra loro se e solo se $\pi = \frac{\pi}{2}$; inoltre, per ogni scalare non nullo $\lambda \in \mathbb{R}$, l'angolo θ tra v e λv è $\theta = 0$ se e solo se $\lambda > 0$, mentre $\theta = \pi$ se e solo se $\lambda < 0$.

Corollario 50.4. *(Disuguaglianza triangolare) Per ogni u, v vettori di uno spazio vettoriale Euclideo (V, g) si ha:*

$$||u + v|| \leq ||u|| + ||v||. \tag{44}$$

È sufficiente provare che:

$$||u + v||^2 \leq (||u|| + ||v||)^2.$$

Infatti:

$$||u + v||^2 = ||u||^2 + 2g(u, v) + ||v||^2 \leq |u||^2 + 2||u||||v|| + ||v||^2 = (||u|| + ||v||)^2.$$

51. Sottospazi affini di $\mathbb{R}^n$

In questo paragrafo iniziamo una trattazione della geometria analitica nello spazio Euclideo $\mathbb{R}^n$ $(n \geq 1)$, basata sulle nozioni di algebra lineare fin qui esposte. Gli oggetti di studio comprendono i punti, le rette ed i piani della geometria elementare, ma la loro definizione sarà fondata sul concetto di spazio vettoriale.

Nel seguito considereremo lo spazio vettoriale $\mathbb{R}^n$ munito (salvo avviso contrario) del prodotto scalare standard, che verrà denotato con il punto $\cdot$. Ogni elemento di $\mathbb{R}^n$ verrà chiamato *punto* o *vettore* a seconda del contesto e dell'interpretazione geometrica che si vuole attribuire. Di norma, nel primo caso esso sarà denotato con una lettera maiuscola, nel secondo caso con una lettera minuscola.

Definizione 51.1. Sia chiama **sottopazio affine** di $\mathbb{R}^n$ ogni sottoinsieme del tipo

$$P_o + W := \{P_o + w | \, w \in W\}$$

dove $P_0 \in \mathbb{R}^n$ è un punto fissato e $W \subset \mathbb{R}^n$ è un sottospazio vettoriale assegnato.

Diremo che $S = P_o + W$ è il sottospazio affine *passante per* P_0 e avente *giacitura* W (oppure sottospazio affine *passante per* P_0 e *parallelo a* W).

Si pone

$$\mathtt{dim}(P_o + W) := \mathtt{dim}(W).$$

Osserviamo che il punto P_0 appartiene effettivamente al sottospazio affine $P_o + W$, in quanto

$$P_0 = P_0 + 0$$

dove denotiamo con 0 il vettore nullo di $\mathbb{R}^n$. Inoltre risulta che $P_o + W$ è completamente determinato da un suo punto qualsiasi e dalla sua giacitura:

Proposizione 51.2. *Sia Q_o un punto del sottospazio affine $P_o + W$. Allora si ha*

$$P_o + W = Q_o + W.$$

Dimostrazione: Infatti, per ipotesi il punto Q_o si può scrivere

$$Q_o = P_0 + w_0$$

dove $w_o \in W$ è un vettore opportuno. Si consideri quindi un punto $P = P_o + w$ di $P_o + W$, dove $w \in W$. Allora

$$P = P_0 + w = Q_0 - w_0 + w = Q_o + (w - w_0)$$

dove il vettore $w - w_0$ appartiene a W in quanto W è un sottospazio vettoriale. Ciò mostra che P appartiene anche al sottospazio affine $Q_o + W$ e, stante l'arbitrarietà di P, l'inclusione $\subset$ è provata. In modo del tutto simile si verifica l'altra inclusione. $\square$

Esempio 51.3. Per ogni punto P, l'insieme $\{P\}$ è un sottospazio affine: infatti

$$\{P\} = P + \{0\};$$

in altri termini, la giacitura in tal caso è il sottospazio vettoriale banale di $\mathbb{R}^n$. Il sottospazio $\{P\}$ è ancora denominato *punto* ed ha dimensione 0. Tutti i sottospazi affini 0-dimensionali sono di questo tipo.

Definizione 51.4. I sottospazi affini di dimensione 1 di $\mathbb{R}^n$ si dicono *rette*, mentre quelli di dimensione 2 si dicono *piani*. Inoltre si chiamano *iperpiani* i sottospazi affini di dimensione $n - 1$.

Si introduce anche la seguente terminologia:

Definizione 51.5. Dato un sottospazio affine $S = P_0 + W$, un vettore $v \in \mathbb{R}^n$ si dice *parallelo* a S se $v \in W$; si dice *perpendicolare* o *ortogonale* a S se $v \in W^{\perp}$.

Esempio 51.6. Consideriamo un sistema lineare di $m \geq 1$ equazioni nelle incognite $(x_1, \ldots, x_n)$:

$$Ax = b.$$

Se tale sistema è compatibile, allora l'insieme S delle soluzioni del sistema:

$$S = \left\{ P = (x_1, \ldots, x_n) \in \mathbb{R}^n : \; A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\} \tag{45}$$

è un sottospazio affine di $\mathbb{R}^n$ di dimensione $\mathtt{dim}(S) = n - rg(A)$.

Infatti, dalla teoria dei sistemi lineari è noto che, fissata una soluzione particolare $P_0 = (x_0^1, \ldots, x_0^n)$ del sistema, si ha

$$S = P_0 + W$$

dove W denota lo spazio vettoriale delle soluzioni del sistema lineare omogeneo associato:

$$Ax = 0.$$

Pertanto S è il sottospazio affine passante per P_0 con giacitura W. Inoltre il Teorema di Rouchè-Capelli dice che il sistema ha ∞^{n-r} soluzioni, dove $r = rg(A)$; ma ciò significa in termini rigorosi che $\mathtt{dim}(W) = n - r$.

Introduciamo ora una notazione utile che permette di caratterizzare in modo conveniente l'appartenenza di un punto ad un dato sottospazio affine:

Definizione 51.7. Dati due punti A, B di $\mathbb{R}^n$, il vettore

$$\overrightarrow{AB} := B - A$$

si chiamerà vettore applicato di estremi A e B.

Si osservi che segue immediatamente dalla definizione che $\overrightarrow{BA} = -\overrightarrow{AB}$.

Proposizione 51.8. *Dati un punto P_0 ed un sottospazio vettoriale W si ha:*

$$P_0 + W = \{P \in \mathbb{R}^n | \; \overrightarrow{P_0P} \in W\}.$$

Dimostrazione: Anche questa è una semplice verifica: considerato un punto $P = P_0 + w$ del sottospazio affine $P_0 + W$, si ha

$$\overrightarrow{P_0P} = P - P_0 = w \in W.$$

Viceversa, se un punto P è tale che $\overrightarrow{P_0P} \in W$, denotato con w tale vettore abbiamo che $P = P_0 + w$ e quindi P appartiene al sottospazio affine in esame. $\square$

Utilizzeremo questo risultato per dare la seguente importante caratterizzazione dei sottospazi affini, che afferma che l'esempio discusso sopra ne costituisce il prototipo:

Teorema 51.9. *Sia $S \subset \mathbb{R}^n$ un sottoinsieme. Sono proprietà equivalenti:*

a) S è un sottospazio affine;

b) S è l'insieme delle soluzioni di un sistema lineare compatibile

$$Ax = b,$$

ovvero

$$S = \left\{ P = (x_1, \dots, x_n) : A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\}. \tag{46}$$

Inoltre, vera a) oppure b), la giacitura W di S è lo spazio delle soluzioni del sistema lineare omogeneo

$$Ax = 0$$

e si ha $\texttt{dim}(S) = n - rg(A)$.

Dimostrazione: Abbiamo già discusso che $b) \Rightarrow a)$ nell'Esempio 51.6, insieme alle ultime affermazioni. Resta da provare che $a) \Rightarrow b)$. Supponiamo che S sia il sottospazio affine $S = P_0 + W$. Osserviamo che, se $W = \mathbb{R}^n$, allora anche $S = \mathbb{R}^n$ e S si può considerare come l'insieme delle soluzioni del sistema (di rango 0) costituito dall'unica equazione:

$$0x_1 + \cdots + 0x_n = 0.$$

Supponiamo quindi $W \neq \mathbb{R}^n$; segue che $\texttt{dim}(W^\perp) > 0$. Fissiamo una base $u_1, \dots, u_m$ di $W^\perp$. Abbiamo che un punto $P = (x_1, \dots, x_n)$ di $\mathbb{R}^n$ appartiene a S se solo se

$$\overrightarrow{P_0P} \in W,$$

ma essendo $W = (W^\perp)^\perp$ ciò accade se e solo se il vettore $\overrightarrow{P_0P}$ è ortogonale a tutti i vettori di $W^\perp$, il che avviene se e solo se

$$\overrightarrow{P_0P} \cdot u_i = 0, \quad i = 1, \dots, m, \tag{47}$$

(si ricordi l'Oss. 49.2). Ciò fornisce un sistema di equazioni lineari che descrive S: infatti le (47) si riscrivono

$$(P - P_o) \cdot u_i = 0, \quad i = 1, \dots, m,$$

ovvero:

$$P \cdot u_i = P_o \cdot u_i, \quad i = 1, \dots, m$$

o, in termini matriciali:

$${}^t u_i \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = b_i, \qquad i = 1, \dots, m$$

dove abbiamo posto $b_i := P_o \cdot u_i$. Pertanto S coincide con l'insieme delle soluzioni del sistema
$$Ax = b$$
dove A è la matrice con m righe e n colonne $A = \begin{pmatrix} {}^t u_1 \\ \vdots \\ {}^t u_m \end{pmatrix}$ e $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. $\square$

Per sintetizzare il fatto che un sottospazio affine S è l'insieme delle soluzioni di un sistema lineare, si scrive
$$S : Ax = b$$
e le equazioni del sistema si dicono *le equazioni di* S. Questa scrittura corrisponde esplicitamente alla (46).

Quindi, ad esempio, un piano di $\mathbb{R}^3$ è descritto da una singola equazione lineare non banale, mentre una retta da un sistema di due equazioni indipendenti, nel senso che la matrice A dei coefficienti ha rango 2. Torneremo sulla geometria di $\mathbb{R}^3$ più in dettaglio in seguito.

Più in generale, un sottospazio affine di *codimensione* m, cioè di dimensione $n-m$ può essere descritto mediante un sistema lineare di m equazioni, che si può scegliere di rango m.

52. Sistemi di riferimento affini

La descrizione dei sottospazi affini che abbiamo ottenuto nel paragrafo precedente fa uso delle coordinate naturali $(x_1, \ldots, x_n)$ di un punto generico $P = (x_1, \ldots, x_n)$ di $\mathbb{R}^n$; nella pratica è necessario poter disporre di sistemi di riferimento opportunamente scelti per descrivere gli oggetti geometrici (ad es. luoghi) mediante equazioni, come il lettore è abituato a fare nell'ordinaria geometria analitica piana. Si introduce a questo scopo questa definizione fondamentale:

Definizione 52.1. Si chiama *riferimento affine* di $\mathbb{R}^n$ ogni coppia $(O, \mathfrak{B})$ dove O è un punto fissato, detto *origine* del riferimento, mentre $\mathfrak{B}$ è una base fissata di $\mathbb{R}^n$.

Un riferimento Cartesiano (ortonormale) è un riferimento affine la cui base $\mathfrak{B}$ è una base ortonormale.

N.B. Non si confonda il punto O con il vettore nullo 0.

Le coordinate di un punto si attribuiscono, in presenza di un riferimento, nel modo specificato dalla definizione seguente:

Definizione 52.2. Sia $\mathfrak{R} = (O, \mathfrak{B})$ un riferimento affine. Sia $P \in \mathbb{R}^n$ un punto. Si chiamano *coordinate* di P rispetto a $\mathfrak{R}$ le componenti $(x_1, \ldots, x_n)$ del vettore $\overrightarrow{OP}$ rispetto alla base $\mathfrak{B}$.

Si userà la scrittura $P(x_1, \ldots, x_n)$ per indicare il fatto che P ha coordinate $(x_1, \ldots, x_n)$ rispetto ad un riferimento assegnato. Dunque, posto $\mathfrak{B} = \{u_1, \ldots, u_n\}$, allora $P(x_1, \ldots, x_n)$ vuol dire che

$$\overrightarrow{OP} = x_1 u_1 + \cdots + x_n u_n,$$

ovvero che

$$P = O + x_1 u_1 + \cdots + x_n u_n.$$

Giova notare esplicitamente che le scritture $P(x_1, \ldots, x_n)$ e $P = (x_1, \ldots, x_n)$ significano cose diverse.

Esempio 52.3. Il riferimento *standard* di $\mathbb{R}^n$ è $\mathfrak{R}_0 = (0_{\mathbb{R}^n}, \mathfrak{B}_0)$ dove $\mathfrak{B}_0 = \{e_1, \ldots, e_n\}$ è la base canonica. Si tratta naturalmente di un riferimento Cartesiano. Solo rispetto a questo riferimento abbiamo che ogni $P = (x_1, \ldots, x_n)$ ha coordinate $(x_1, \ldots, x_n)$.

Teorema 52.4. *Fissato un riferimento affine $\mathfrak{R} = (O, \mathfrak{B})$, la funzione*

$$\psi_{\mathfrak{R}} : \mathbb{R}^n \to \mathbb{R}^n$$

che associa ad ogni punto P il vettore $(x_1, \ldots, x_n)$ delle sue coordinate rispetto a $\mathfrak{R}$ è una bigezione.

Dimostrazione: Conosciamo la funzione $\psi_{\mathfrak{B}} : \mathbb{R}^n \to \mathbb{R}^n$ che associa ad ogni vettore v le sue componenti rispetto alla base, e sappiamo che si tratta di un isomorfismo lineare. Per definizione delle coordinate, per ogni punto P si ha

$$\psi_{\mathfrak{R}}(P) = \psi_{\mathfrak{B}}(\overrightarrow{OP}).$$

Ciò suggerisce che la funzione $\psi_{\mathfrak{R}}$ in questione può riguardarsi come la composta di due funzioni:

$$\mathbb{R}^n \xrightarrow{G} \mathbb{R}^n \xrightarrow{\psi_{\mathfrak{B}}} \mathbb{R}^n$$

dove $G : \mathbb{R}^n \to \mathbb{R}^n$ è definita ponendo

$$G(P) := \overrightarrow{OP}.$$

Ora, è di semplice verifica il fatto che G è bigettiva: dato infatti un vettore $X \in \mathbb{R}^n$ esiste uno ed un solo $P \in \mathbb{R}^n$ tale che

$$G(P) = X$$

ovvero tale che

$$\overrightarrow{OP} = P - O = X;$$

si tratta del punto $P = X + O$. Dunque $\psi_{\mathfrak{R}} = \psi_{\mathfrak{B}} \circ G$ è bigettiva in quanto composizione di bigezioni. $\square$

Osservazione 52.5. Giova osservare che questo risultato significa che, rispetto ad un riferimento affine, vi è una corrispondenza biunivoca tra punti e n-ple di coordinate, esattamente come nella geometria analitica nota allo studente; vale a dire: due punti $P(x_1, \dots, x_n)$ e $Q(x'_1, \dots, x'_n)$ coincidono se e solo se hanno le stesse coordinate, cioè $x_i = x'_i$ per ogni i, mentre fissata $(x^1_o, \dots, x^n_o)$ una n-pla di scalari, vi è esattamente un punto P di coordinate $(x^1_o, \dots, x^n_o)$.

Ad esempio, considerato il riferimento

$$\mathfrak{R} = (O = (1,2), \mathfrak{B} = \{u_1 = (3,1), u_2 = (\sqrt{2}, -1)\})$$

del piano Euclideo $\mathbb{R}^2$, allora il punto di coordinate $(1, -2)$ rispetto a tale riferimento è

$$P = O + u_1 - 2u_2 = (1,2) + (3,1) - 2(\sqrt{2}, -1) = (4 - 2\sqrt{2}, 5).$$

53. Cambiamenti di riferimento

Proposizione 53.1. *(Cambiamento di riferimento affine) Siano* $\mathfrak{R} = (O, \mathfrak{B})$ *e* $\mathfrak{R}' = (O', \mathfrak{B}')$ *due riferimenti affini di* $\mathbb{R}^n$. *Per ogni punto* P*, denotati con* x *e* x' *i vettori delle coordinate di* P *rispetto a* $\mathfrak{R}$ *e a* $\mathfrak{R}'$*, sussiste la relazione:*

$$x = Cx' + d \tag{48}$$

dove $C = M_{\mathfrak{B},\mathfrak{B}'}$ *è la matrice di passaggio dalla base* $\mathfrak{B}$ *alla base* $\mathfrak{B}'$ *e* d *è il vettore delle coordinate dell'origine* O' *del riferimento* $\mathfrak{R}'$ *rispetto al riferimento* $\mathfrak{R}$.

Se i riferimenti in questione sono entrambi Cartesiani, allora $C \in O(n)$.

Ai fini di provare la (48) è conveniente far uso della seguente identità, detta *relazione di Chasles*: per ogni $A, B, C \in \mathbb{R}^n$ si ha

$$\overrightarrow{AB} = \overrightarrow{AC} + \overrightarrow{CB}.$$

Infatti

$$\overrightarrow{AC} + \overrightarrow{CB} = C - A + B - C = B - A = \overrightarrow{AB}.$$

Ora, considerati i due riferimenti nell'enunciato della Proposizione precedente, e un punto P qualsiasi, abbiamo:

$$\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P}.$$

La (48) si ricava uguagliando i vettori delle componenti di ambo i membri, rispetto alla base $\mathfrak{B}$. Per quel che concerne il primo membro, tale vettore è il vettore delle coordinate affini x di P rispetto a $\mathfrak{R}$; quindi abbiamo:

$$x = \psi_{\mathfrak{B}}(\overrightarrow{OO'}) + \psi_{\mathfrak{B}}(\overrightarrow{O'P}).$$

Analogamente, si ha $\psi_{\mathfrak{B}}(\overrightarrow{OO'}) = d$, dove d è il vettore delle coordinate di O' nello stesso riferimento. D'altra parte $x' = \psi_{\mathfrak{B}'}(\overrightarrow{O'P})$; ma, applicando la (25), sappiamo che tra i vettori delle componenti di $\overrightarrow{O'P}$ rispetto alle due basi $\mathfrak{B}$ e $\mathfrak{B}'$ sussiste la relazione:

$$\psi_{\mathfrak{B}}(\overrightarrow{O'P}) = Cx'$$

e quindi si ottiene la (48). L'ultima affermazione è immediata applicazione della Prop. 46.6.

Torniamo ora a discutere i sottospazi affini; in presenza di un riferimento non standard, il Teorema 51.9 ammette la seguente generalizzazione:

Teorema 53.2. *Si fissi un riferimento affine $\mathfrak{R} = (O, \mathfrak{B})$ di $\mathbb{R}^n$. Sia $S \subset \mathbb{R}^n$ un sottoinsieme. Sono proprietà equivalenti:*

a) S è un sottospazio affine.

b) S è l'insieme dei punti le cui coordinate rispetto a $\mathfrak{R}$ sono soluzioni di un sistema lineare compatibile

$$Ax = b,$$

ovvero

$$S = \left\{ P(x_1, \ldots, x_n) : A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\} \tag{49}$$

Inoltre, vera a) oppure b), la giacitura W di S è il sottospazio vettoriale costituito dai vettori le cui componenti rispetto alla base $\mathfrak{B}$ sono soluzioni del sistema lineare omogeneo

$$Ax = 0,$$

ovvero, posto $\mathfrak{B} = \{v_1, \ldots, v_n\}$:

$$W = \{v = \sum_{i=1}^{n} x_i v_i |\, Ax = 0\}.$$

Si ha inoltre $\mathtt{dim}(S) = n - rg(A)$.

Dimostrazione: Fissiamo un riferimento affine $\mathfrak{R}$, rispetto al quale denotiamo con x' il vettore delle coordinate di un punto generico. Consideriamo la matrice di passaggio C dalla base canonica alla base $\mathfrak{B}$. Sappiamo che un sottoinsieme $S \subset \mathbb{R}^n$ è un sottospazio affine se e solo se è della forma:

$$S = \left\{ P = (x_1, \ldots, x_n) : A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\}$$

ovvero, considerando le coordinate dei punti rispetto al riferimento standard $\mathfrak{R}_0 = (0, \mathfrak{B}_0)$:

$$S = \{P(x_1, \ldots, x_n) :\ Ax = b\}.$$

Utilizzando la (48), rispetto al riferimento $\mathfrak{R}$ lo stesso insieme è descritto come segue:

$$S = \{P(x'_1, \dots, x'_n) : A(Cx' + d) = b\}$$

cioè

$$S = \{P(x'_1, \dots, x'_n) : A'x' = b'\},$$

dove abbiamo posto $A' := AC$ e $b' := b - Ad$. Riguardo la giacitura W di S, in modo simile, sfruttando ancora la (25) abbiamo che

$$W = \{v = (x_1, \dots, x_n) \in \mathbb{R}^n : Ax = 0\} = \{v = \sum_{i=1}^{n} x'_i v_i| : (AC)x' = 0\}.$$

□

54. Sottospazi affini di $\mathbb{R}^3$

Nel caso familiare dello spazio tridimensionale $\mathbb{R}^3$ abbiamo solo tre tipi di sottospazi: punti, rette e piani. Fissato un riferimento affine $\mathfrak{R} = (O, \{u_1, u_2, u_3\})$, ogni piano α è rappresentato da un'equazione non banale:

$$\alpha : ax + by + cz + d = 0, \quad (a, b, c) \neq 0.$$

Ogni retta r è l'insieme dei punti le cui coordinate soddisfano un sistema del tipo

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}, \quad rg \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

Geometricamente, ciò comporta che ogni retta è l'intersezione di due piani opportuni. Discuteremo in generale l'intersezione tra sottospazi affini nel prossimo paragrafo.

I punti $\{P\}$ invece corrispondono alle soluzioni di sistemi di Cramer nelle incognite (x, y, z):

$$P : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \\ a''x + b''y + c''z + d'' = 0 \end{cases}, \quad \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \neq 0.$$

Le rette sono descrivibili in modo conveniente anche mediante *equazioni parametriche*: si fissi un punto $P_o(x_0, y_0, z_0)$ su una retta r di giacitura W, di modo che $r = P_0 + W$. Dato un vettore non nullo v parallelo a r, abbiamo $W = L(v)$ in quanto il sottospazio W è 1-dimensionale. Dunque

$$r = \{P \in \mathbb{R}^3 : \overrightarrow{P_0P} \in W\} = \{P \in \mathbb{R}^3 : \exists t \in \mathbb{R} : \overrightarrow{P_0P} = tv\} = \{P_0 + tv : \ t \in \mathbb{R}\}.$$

Denotate con (l, m, n) le componenti di v rispetto alla base del riferimento, abbiamo dunque che il generico punto $P \in r$ è:

$$P = P_0 + tv = O + x_0u_1 + y_0u_2 + z_0u_3 + t(lu_1 + mu_2 + nu_3) = (x_0 + tl)u_1 + (y_0 + tm)u_2 + (z_0 + tn)u_3,$$

cioè è il punto $P(x_0 + tl, y_0 + tm, z_0 + tn)$, al variare di $t \in \mathbb{R}$. Ciò mostra che le retta r può descriversi mediante le equazioni

$$\begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad t \in \mathbb{R}.$$

che sono dette le sue *equazioni parametriche.*

Di solito i numeri l, m, n si chiamano *parametri direttori* della retta nel riferimento fissato. Si noti che il punto P_0 si riottiene in corrispondenza del valore $t = 0$ del parametro.

55. Posizioni reciproche tra sottospazi affini

In questo parafrafo iniziamo lo studio della geometria affine in $\mathbb{R}^n$, il cui oggetto di studio sono le relazioni tra sottospazi affini (geometria di posizione); essa è incentrata sui concetti definiti di seguito:

Definizione 55.1. Due sottospazi affini S e T di $\mathbb{R}^n$ aventi giaciture rispettivamente U e W si dicono:

- **incidenti** se $S \cap T \neq \emptyset$;
- **paralleli** se $U \subset W$ oppure $W \subset U$;
- **sghembi** se non sono paralleli e $S \cap T = \emptyset$.

Prima di proseguire è opportuno notare che:

Proposizione 55.2. *Se S e T sono sottospazi affini incidenti, con giaciture U e W, allora:*

$$S \subset T \iff U \subset W.$$

Dimostrazione: Infatti, fissato un punto $P_0 \in S \cap T$, possiamo rappresentare i due sottospazi nella forma

$$S = P_0 + U, \ T = P_0 + W \tag{50}$$

e quindi è evidente che, se $U \subset W$, allora $S \subset T$. Riguardo l'altra implicazione, se $S \subset T$, allora dato $u \in U$ si ha che $P_0 + u$ appartiene a S e quindi anche a T; pertanto esiste un vettore $w \in W$ tale che

$$P_0 + u = P_0 + w$$

e quindi necessariamente $u = w \in W$. $\square$

Riguardo il parallelismo consegue:

Proposizione 55.3. *Dati due sottospazi affini S e T risulta:*

$$S \text{ e } T \text{ paralleli} \Rightarrow S \subset T \text{ oppure } T \subset S \text{ oppure } S \cap T = \emptyset.$$

Dimostrazione: Supponiamo che S e T siano paralleli: se $S \cap T = \emptyset$ si ha la tesi. Altrimenti S e T sono incidenti e, in base alla definizione di parallelismo, $U \subset W$ oppure $W \subset U$: in virtù di quanto stabilito sopra, uno dei due sottospazi è contenuto nell'altro. $\square$

Nel seguito scriveremo $S//T$ per indicare il fatto che S e T sono sottospazi paralleli.

Si noti anche la seguente caratterizzazione della condizione di incidenza:

Proposizione 55.4. *Dati i sottospazi affini $S = A + U$ e $T = B + W$ di $\mathbb{R}^n$ si ha:*

$$S \text{ e } T \text{ sono incidenti} \Longleftrightarrow \overrightarrow{AB} \in U + W.$$

Dimostrazione: Se S e T sono incidenti, si fissi $P_o \in S \cap T$. Allora

$$\overrightarrow{AB} = \overrightarrow{AP_o} + \overrightarrow{P_oB} = \overrightarrow{AP_o} - \overrightarrow{BP_o} \in U + W.$$

Viceversa, supponiamo che si abbia $\overrightarrow{AB} \in U+W$. Allora esistono due vettori $u \in U$ e $w \in W$ tali che:

$$\overrightarrow{AB} = u + w.$$

Dunque

$$B - w = A + u \in S \cap T$$

per cui S e T sono incidenti. $\square$

La nostra intuizione e conoscenza pregressa suggeriscono che nello spazio $\mathbb{R}^3$ due piani distinti α e β sono sempre o incidenti o paralleli, avendosi nei due casi, rispettivamente:

$$\alpha \cap \beta = r \text{ retta}, \ \alpha \cap \beta = \emptyset.$$

Mentre per una retta r ed un piano α, deve verificarsi una delle seguenti:

$$r \subset \alpha, \ r \cap \alpha = \{P\}, \ r \cap \alpha = \emptyset.$$

Invece, nel caso di due rette distinte r e s le possibili posizioni reciproche sono:

$$r \text{ e } s \text{ incidenti}, \ r//s, \ r \text{ e } s \text{ sghembe}.$$

Nel primo caso $r \cap s$ è un punto, mentre negli ultimi due casi, mutuamente esclusivi, $r \cap s = \emptyset$.

Queste affermazioni si possono (e si devono) tutte provare in modo rigoroso in base alle definizioni di retta e piano che abbiamo introdotto.

Prima di fare questo, discuteremo alcuni risultati di natura generale su cui si basa lo studio della geometria di posizione in $\mathbb{R}^n$.

Proposizione 55.5. *Siano S e T sottospazi affini incidenti di $\mathbb{R}^n$. Allora $S \cap T$ è un sottospazio affine, la cui giacitura è l'intersezione delle giaciture dei due sottospazi.*

Dimostrazione: Fissato un punto $P_o \in S \cap T$ rappresentiamo ancora S e T mediante la (50). Allora risulta

$$S \cap T = P_o + (U \cap W)$$

e poichè $U \cap W$ è un sottospazio vettoriale di $\mathbb{R}^n$, segue l'asserto. Questa uguaglianza si giustifica subito tenendo conto della condizione (51.8) di appartenza di un punto ad un dato sottospazio affine. Infatti, un punto P appartiene a $S \cap T$ se e solo se si ha simultaneamente sia $\overrightarrow{P_oP} \in U$ che $\overrightarrow{P_0P} \in W$, ovvero se e solo se $\overrightarrow{P_oP} \in U \cap W$ e ciò significa che P appartiene al sottospazio $P_o + (U \cap W)$. $\square$

Al fine di analizzare la posizione reciproca di due sottospazi, un'informazione utile è sapere se essi sono contenuti entrambi in un altro sottospazio (ad es. nel caso di due rette sapere se esse sono o no complanari); il seguente risultato è basilare:

Teorema 55.6. *Siano $S = A + U$ e $T = B + W$ due sottospazi affini di $\mathbb{R}^n$. Allora esiste il più piccolo (per inclusione) sottospazio affine contenente sia S che T. Esso si chiama* **sottospazio congiungente** *S e T, si denota con $S \vee T$, ed è dato da:*

$$S \vee T = A + L(\overrightarrow{AB}) + (U + W). \tag{51}$$

Dimostrazione: Consideriamo il sottospazio affine $E = A + L(\overrightarrow{AB}) + (U + W)$. Osserviamo che E passa ovviamente per A, ma anche per B, in quanto per definizione, $\overrightarrow{AB}$ appartiene alla sua giacitura. Siccome quest'ultima contiene sia U che W, abbiamo che E contiene sia S che T. Resta da verificare che E è il più piccolo sottospazio cha ha questa proprietà. Sia S' un altro sottospazio affine contenente S e T; allora S' passa per A e per B e possiamo rappresentare S' come $S' = A + U'$ dove U' è la giacitura di S'. In particolare $\overrightarrow{AB} \in U'$. Segue che $U \subset U'$, $W \subset U'$ e pertanto $L(\overrightarrow{AB}) + (U + W) \subset U'$ da cui $E \subset S'$. $\square$

Vale la seguente formula per calcolare la dimensione di $S \vee T$, che è la versione affine della formula di Grassmann (16):

Teorema 55.7. *Siano S e T due sottospazi affini di $\mathbb{R}^n$, aventi giaciture rispettivamente U e W. Allora la dimensione di $S \vee T$ si calcola come segue:*

1) Nel caso $S \cap T \neq \emptyset$: $\mathtt{dim}(S \vee T) = \mathtt{dim}(S) + \mathtt{dim}(T) - \mathtt{dim}(S \cap T)$.

2) Nel caso $S \cap T = \emptyset$: $\mathtt{dim}(S \vee T) = \mathtt{dim}(S) + \mathtt{dim}(T) + 1 - \mathtt{dim}(U \cap W)$.

Dimostrazione:

1) In questo caso sappiamo che $\overrightarrow{AB} \in U + W$ (Prop. 55.4) per cui, tenendo conto della (51) abbiamo che la giacitura di $S \vee T$ è $U + W$. Inoltre la giacitura di $S \cap T$ è $U \cap W$. Pertanto, applicando la formula di Grassmann abbiamo:

$$\mathtt{dim}(S \vee T) = \mathtt{dim}(U + W) = \mathtt{dim}(U) + \mathtt{dim}(W) - \mathtt{dim}(U \cap W) = \mathtt{dim}(S) + \mathtt{dim}(T) - \mathtt{dim}(S \cap T).$$

2) In tal caso $\overrightarrow{AB} \notin (U+W)$, per cui

$$\mathtt{dim}((U+W)+L(\overrightarrow{AB})) = \mathtt{dim}(U+W)+1,$$

in quanto la somma di $L(\overrightarrow{AB})$ e di $U+W$ è diretta. Allora

$$\mathtt{dim}(S \vee T) = \mathtt{dim}(U+W)+1 = \mathtt{dim}(U)+\mathtt{dim}(W)-\mathtt{dim}(U \cap W)+1$$

da cui la tesi. □

Come applicazione, discutiamo la seguente caratterizzazione dei sottospazi sghembi:

Proposizione 55.8. *Siano S e T sottospazi affini con $\mathtt{dim}(S) \leq \mathtt{dim}(T)$. Allora:*

$$S \text{ e } T \text{ sono sghembi} \iff S \cap T = \emptyset \text{ e } \mathtt{dim}(S \vee T) \geq \mathtt{dim}(T)+2.$$

Dimostrazione: Denotiamo al solito con U e W le giaciture dei due sottospazi. Notiamo che, essendo per ipotesi $\mathtt{dim}(S) \leq \mathtt{dim}(T)$, tenendo conto della definizione di parallelismo, abbiamo $S//T$ se e solo se $U \subset W$. Infatti, nel caso in cui $W \subset U$, si ha necessariamente $W = U$.

$\Rightarrow$) Supponiamo che S e T siano sghembi; dalla formula di Grassmann, caso 2), otteniamo che

$$\mathtt{dim}(S \vee T) = \mathtt{dim}(T)+1+(\mathtt{dim}(S)-\mathtt{dim}(U \cap W)). \tag{52}$$

Ora notiamo che $\mathtt{dim}(S)-\mathtt{dim}(U \cap W) > 0$, perchè altrimenti si avrebbe $\mathtt{dim}(U) = \mathtt{dim}(U \cap W)$, da cui $U \cap W = U$ e quindi $U \subset W$; ciò significherebbe $S//T$, contro l'ipotesi. Pertanto $\mathtt{dim}(S \vee T) \geq \mathtt{dim}(T)+2$.

$\Leftarrow$) Supponiamo che $\mathtt{dim}(S \vee T) \geq \mathtt{dim}(T)+2$ e S e T non incidenti. Vale ancora la (52), per cui otteniamo, stante l'ipotesi:

$$\mathtt{dim}(T)+1+(\mathtt{dim}(S)-\mathtt{dim}(U \cap W)) \geq \mathtt{dim}(T)+2$$

e quindi

$$(\mathtt{dim}(S)-\mathtt{dim}(U \cap W)) \geq 1.$$

Pertanto $\mathtt{dim}(U \cap W) \neq \mathtt{dim}(U)$ e quindi $U \not\subset W$ e questo implica che S e T non possono essere paralleli. □

Soffermiamoci ora sulla geometria di posizione in $\mathbb{R}^3$.

Proposizione 55.9. *Due piani di $\mathbb{R}^3$ non sono mai sghembi. Quindi se α e β sono piani distinti, essi sono sempre o incidenti o paralleli. Si ha:*

$$\alpha//\beta \iff \alpha \cap \beta = \emptyset.$$

Inoltre, se α e β sono incidenti, allora $\alpha \cap \beta$ è una retta.

Dimostrazione: La prima affermazione segue direttamente dalla Prop. precedente, avendosi $\mathtt{dim}(\beta) + 2 = 4 > 3$. Supponendo ora $\alpha \neq \beta$, è chiaro che se α e β sono paralleli, allora sono disgiunti (cfr. Prop. 55.3), e viceversa, se sono disgiunti, allora devono necessariamente essere paralleli, non potendo essere sghembi. Infine, consideriamo il caso in cui α e β sono incidenti; lo spazio congiungente $\alpha \vee \beta$ non è un piano perchè altrimenti coinciderebbe con entrambi i piani in quanto li contiene, ma ciò contraddice il fatto che $\alpha \neq \beta$. Quindi $\alpha \vee \beta = \mathbb{R}^3$ e la formula di Grassmann nel caso 1) implica che $\mathtt{dim}(\alpha \cap \beta) = 2 + 2 - 3 = 1$. $\square$

Discutiamo ora una caratterizzazione rilevante della circostanza che due rette di $\mathbb{R}^3$ siano sghembe. Definiamo due rette **complanari** se esiste un piano che contiene entrambe.

Proposizione 55.10. *Siano r e s rette distinte di $\mathbb{R}^3$. Allora:*

$$r \text{ e } s \text{ sono sghembe} \iff r \text{ e } s \text{ non sono complanari.}$$

Quindi:

$$r \text{ e } s \text{ sono complanari} \iff r \text{ e } s \text{ sono incidenti oppure sono parallele.}$$

Dimostrazione: In base alla Prop. 55.8, r e s sono sghembe se e solo se $r \cap s = \emptyset$ e $\mathtt{dim}(r \vee s) \geq 3$, cioè se e solo se $r \cap s = \emptyset$ e $r \vee s = \mathbb{R}^3$. Ora, se queste condizioni sono vere, certamente r e s non sono complanari, perchè $r \vee s$ è il più piccolo sottospazio contenente entrambe le rette. Viceversa, se r e s non sono complanari, allora chiaramente $r \vee s = \mathbb{R}^3$ e si ha anche $r \cap s = \emptyset$, perchè altrimenti applicando la formula di Grassmann, nel caso 1), avremmo

$$3 = \mathtt{dim}(r \vee s) = 2 - \mathtt{dim}(r \cap s) \leq 2,$$

il che è una contraddizione. Quindi r e s sono sghembe. $\square$

Riassumendo, per le rette abbiamo le seguenti possibilità:

Proposizione 55.11. *Due rette distinte r e s di $\mathbb{R}^3$ possono essere incidenti, parallele o sghembe. Nel primo caso, la loro intersezione è un punto, nel secondo caso r e s sono disgiunte; in entrambi questi casi, esse sono complanari e contenute in un unico piano. Se r e s sono sghembe, allora sono disgiunte e non complanari.*

Lasciamo infine al lettore il compito di esaminare in dettaglio il caso di una retta r ed un piano α, provando i fatti seguenti:

Proposizione 55.12. *Siano r una retta e α un piano di $\mathbb{R}^3$. Allora può accadere che che r e α siano incidenti oppure paralleli. Se r e α sono incidenti, allora o $r \subset \alpha$ oppure $r \cap \alpha$ è un punto. Inoltre, se $r \not\subset \alpha$, si ha:*

$$r // \alpha \iff r \cap \alpha = \emptyset.$$

Giova osservare che anche in questo caso, è escluso a priori che una retta ed un piano siano sghembi, sempre perchè $\mathtt{dim}(\alpha) + 2 = 4 > 3$.

56. Perpendicolarità e condizioni analitiche

Definizione 56.1. Due sottospazi affini S e T di $\mathbb{R}^n$ aventi giaciture rispettivamente U e W, si dicono **perpendicolari** se risulta

$$u \cdot w = 0$$

per ogni $u \in U$ e $w \in W$, oppure se

$$u' \cdot w' = 0$$

per ogni $u' \in U^\perp$ e $w' \in W^\perp$.

In tal caso scriveremo $S \perp T$.

Si osservi che le condizioni in questione si possono riformulare come segue: S e T sono perpendicolari se e solo se

$$U \subset W^\perp \text{ oppure } U^\perp \subset W.$$

Infatti richiedere che $u \cdot w = 0$ per ogni $u \in U$ e $w \in W$, vuol dire che $U \subset W^\perp$, mentre richiedere che $u' \cdot w' = 0$ per ogni $u' \in U^\perp$ e $w' \in W^\perp$, significa imporre che $U^\perp \subset \left(W^\perp\right)^\perp$, ma $\left(W^\perp\right)^\perp = W$.

Nel primo caso, necessariamente:

$$\mathtt{dim}(S) + \mathtt{dim}(T) \leq n$$

mentre nel secondo caso deve aversi:

$$\mathtt{dim}(S) + \mathtt{dim}(T) \geq n.$$

Ciò è immediata conseguenza del Cor. 49.6.

Esempio 56.2. Ad esempio, date due rette r e s, fissati due vettori non nulli v e w paralleli rispettivamente a r e s, allora $r \perp s$ se e solo se $u \cdot w = 0$.

Invece, dati due piani di $\mathbb{R}^3$, la prima delle due condizione nella Definizione 56.1 non può verificarsi, e risulta $\alpha \perp \beta$ se e solo se, fissati due vettori non nulli n e n' perpendicolari rispettivamente ad α e a β, si ha $n \cdot n' = 0$.

Ciò perchè le corrispondenti giaciture U e W hanno dimensione 2 e quindi $U^\perp = L(n)$ mentre $W^\perp = L(n')$.

Discutiamo ora le condizioni analitiche affinchè due sottospazi di $\mathbb{R}^3$ siano paralleli oppure perpendicolari. In tutta la discussione seguente è fissato un riferimento Cartesiano $\mathfrak{R} = (O, \{u_1, u_2, u_3\})$. Scriviamo al solito $P(x, y, z)$ per denotare un punto di coordinate (x, y, z) rispetto al riferimento; usiamo anche la notazione $v(a, b, c)$ per indicare il vettore $v = au_1 + bu_2 + cu_3$ di componenti (a, b, c) rispetto alla base $\mathfrak{B}$.

Osserviamo innanzitutto che dato un piano α di equazione:

$$\alpha : ax + by + cz + d = 0,$$

allora il vettore $n(a, b, c)$ è sempre **perpendicolare** al piano in questione.

Infatti, sappiamo che la giacitura W del piano è descritta dall'equazione

$$W : ax + by + cz = 0,$$

e pertanto utilizzando il prodotto scalare possiamo scrivere:

$$W = \{v \in \mathbb{R}^3 : v \cdot n = 0\} = L(n)^\perp.$$

Ciò è vero perchè la base $\mathfrak{B}$ è ortonormale: si ricordi la (36), che garantisce che $v \cdot n = ax + by + cz$.

Ciò premesso, ricaviamo i seguenti criteri:

Condizioni di parallelismo e perpendicolarità retta-piano

Consideriamo una retta r ed un piano α di equazioni:

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad \alpha : ax + by + cz + d = 0.$$

Allora risulta

$$r//\alpha \iff al + bm + cn = 0, \quad r \perp \alpha \iff \exists \rho \in \mathbb{R} - \{0\}\ t.c. \begin{cases} l &= \rho a \\ m &= \rho b \\ n &= \rho c \end{cases}.$$

La condizione di perpendicolarità si scrive anche sotto forma di rapporti uguali:

$$\frac{l}{a} = \frac{m}{b} = \frac{n}{c},$$

con la convenzione che se uno dei denominatori è zero, tale dev'essere il numeratore.

La giustificazione dei due criteri segue quasi direttamente dalle definizioni: la giacitura di r è infatti $L(v)$ dove v è il vettore di componenti (l, m, n), mentre per quanto premesso sopra $W = L(n)^\perp$ dove n è il vettore $n(a, b, c)$. Pertanto $r//\alpha$ se e solo se $U \subset W$, ma ciò accade se e solo se $v \in W$ ovvero se e solo se $v \cdot n = 0$, ovvero $al + bm + cn = 0$.

Invece $r \perp \alpha$ se e solo se $U \subset W^\perp$ ovvero se e solo se $U \subset L(n)$, il che accade se e solo se $v = \rho n$ per un opportuno scalare non nullo ρ.

Condizioni di parallelismo e perpendicolarità tra rette

Consideriamo due rette di equazioni

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad r' : \begin{cases} x = x'_0 + tl' \\ y = y'_0 + tm' \\ z = z'_0 + tn' \end{cases}.$$

Risulta:

$$r//s \iff \exists \rho \in \mathbb{R} - \{0\}\ t.c. \begin{cases} l' &= \rho l \\ m' &= \rho m \\ n' &= \rho n \end{cases}, \quad r \perp s \iff ll' + mm' + nn' = 0.$$

Infatti la condizione di parallelismo è che le giaciture $L(v)$ e $L(v')$ delle due rette coincidano, dove $v'(l', m', n')$, il che accade se e solo se $v' = \rho v$, mentre la condizione di perpendicolarità è $L(v) \subset L(v')^{\perp}$ che si riduce a $v \cdot v' = 0$.

Condizioni di parallelismo e perpendicolarità tra piani

Consideriamo infine due piani α e β di equazioni:

$$\alpha : ax + by + cz + d = 0, \quad \beta : a'x + b'y + c'z + d' = 0.$$

Si ha:

$$\alpha // \beta \iff \exists \rho \in \mathbb{R} - \{0\}\, t.c. \begin{cases} a' &= \rho a \\ b' &= \rho b \\ c' &= \rho c \end{cases}, \quad \alpha \perp \beta \iff aa' + bb' + cc' = 0.$$

Infatti le giaciture dei piani sono $L(n)^{\perp}$ e $L(n')^{\perp}$, dove $n(a, b, c)$ e $n'(a', b', c')$. Pertanto essi sono paralleli se e solo se $L(n)^{\perp} = L(n')^{\perp}$, ma ciò equivale a $L(n) = L(n')$ ovvero $n' = \rho n$. Invece, stante la definizione, la condizione di perpendicolarità equivale a richiedere che $n \cdot n' = 0$.

Esempio 56.3. Dato il piano $\alpha : 3x - \sqrt{2}z + 5 = 0$, allora i piani paralleli ad esso sono tutti e soli quelli aventi equazione del tipo

$$3x - \sqrt{2}z + k = 0$$

al variare di k in $\mathbb{R}$. Questa è una descrizione analitica del *fascio* di piani paralleli al piano dato.

57. La distanza euclidea

Mediante il prodotto scalare è possibile dare una definizione semplice di distanza tra punti:

Definizione 57.1. Si chiama **distanza Euclidea** tra i punti P e Q di $\mathbb{R}^n$ il numero

$$d(P, Q) := ||\overrightarrow{PQ}||.$$

Si osserva che $d(P, Q) \geq 0$ e si ha $d(P, Q) = 0$ se e solo se $P = Q$.

Un'altra proprietà basilare e familiare della distanza tra punti è la cosiddetta disuguaglianza triangolare.

Proposizione 57.2. *(Disuguaglianza triangolare) Dati i punti A, B, C di $\mathbb{R}^n$ si ha sempre*

$$d(A, B) \leq d(A, C) + d(C, B).$$

Infatti,

$$d(A, B) = ||\overrightarrow{AB}|| = ||\overrightarrow{AC} + \overrightarrow{CB}|| \leq ||\overrightarrow{AC}|| + ||\overrightarrow{CB}|| = d(A, C) + d(C, B).$$

Osservazione 57.3. Riassumendo, la funzione distanza $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ soddisfa le seguenti condizioni:

(1) $d(A, B) \geq 0$
(2) $d(A, B) = 0$ se e solo se $A = B$
(3) $d(A, B) = d(B, A)$
(4) $d(A, B) \leq d(A, C) + d(C, B)$.

valide per ogni A, B, C punti coinvolti. Si osservi che la 3) è conseguenza immediata della proprietà (35) della norma. Queste proprietà sono sufficientemente forti da poter sviluppare una teoria generale dei cosiddetti *spazi metrici*, insiemi su cui è assegnata una nozione di distanza che soddisfa le proprietà precedenti assunte come assiomi.

Tornando alle distanze, in questo contesto il Teorema di Pitagora assume la seguente formulazione:

Teorema 57.4. *(Teorema di Pitagora generalizzato). Siano A, B, C punti di $\mathbb{R}^n$. Allora:*

$$d(B, C)^2 = d(A, B)^2 + d(A, C)^2 - 2\,\overrightarrow{AB} \cdot \overrightarrow{AC}. \tag{53}$$

Dimostrazione: Si ha infatti, utilizzando la formula di polarizzazione (42):

$$\begin{aligned} d(B, C)^2 &= ||\overrightarrow{BC}||^2 = ||\overrightarrow{BA} + \overrightarrow{AC}||^2 = ||\overrightarrow{BA}||^2 + 2(\overrightarrow{BA} \cdot \overrightarrow{AC}) + ||\overrightarrow{AC}||^2 = \\ &= d(A, B)^2 - 2(\overrightarrow{AB} \cdot \overrightarrow{AC}) + d(A, C)^2. \end{aligned}$$

□

La distanza Euclidea si può calcolare utilizzando le coordinate dei punti coinvolti, rispetto a un fissato riferimento Cartesiano, utilizzando la formula fornita dalla Proposizione seguente, che generalizza quanto noto allo studente nel caso della geometria cartesiana del piano:

Proposizione 57.5. *Sia $\mathfrak{R} = (O, \mathfrak{B})$ un riferimento Cartesiano di $\mathbb{R}^n$. Dati i punti $P(x_1, \dots, x_n)$ e $Q(y_1, \dots, y_n)$ si ha:*

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Dimostrazione: Posto $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, dalla relazione

$$\overrightarrow{PQ} = \overrightarrow{PO} + \overrightarrow{OQ} = -\overrightarrow{OP} + \overrightarrow{OQ}$$

segue che il vettore delle componenti di $\overrightarrow{PQ}$ rispetto a $\mathfrak{B}$ è dato da:

$$\psi_{\mathfrak{B}}(\overrightarrow{PQ}) = -\psi_{\mathfrak{B}}(\overrightarrow{OP}) + \psi_{\mathfrak{B}}(\overrightarrow{OQ}) = y - x.$$

Quindi, essendo la base ortonormale:

$$d(P, Q) = ||\overrightarrow{PQ}|| = \sqrt{(y - x) \cdot (y - x)} = \sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2}.$$

□

58. Distanza tra punti e sottospazi

Teorema 58.1. *Siano S un sottospazio affine di $\mathbb{R}^n$ e A un punto. Allora esiste uno ed un solo punto $H \in S$ tale che il vettore $\overrightarrow{AH}$ è perpendicolare a S. Inoltre*

$$d(A,H) < d(A,P) \text{ per ogni punto } P \in S,\ P \neq H.$$

Tale punto prende il nome di **proiezione ortogonale** *di A sul sottospazio S.*

La proiezione ortogonale H del punto A realizza quindi la distanza *minima* tra A ed i punti del sottospazio S: in formule

$$d(A,H) = min\{d(A,P) :\ P \in S\},$$

ed è l'unico punto con questa proprietà. Ha senso quindi definire il numero $d(A,H)$ la **distanza del punto A dal sottospazio** S, che si denota con $d(A,S)$.

Nel caso in cui $A \in S$ si ha evidentemente $H = A$ perchè il vettore nullo è ortogonale a S. Pertanto in tal caso $d(A,S) = 0$.

Dimostrazione: Denotata con W la giacitura di S, sfrutteremo il fatto che

$$\mathbb{R}^n = W \oplus W^\perp.$$

Fissato un punto P_0 su S, decomponiamo il vettore $\overrightarrow{P_0A}$ come

$$\overrightarrow{P_0A} = w + u$$

dove $w \in W$ e $u \in W^\perp$. In corrispondenza del vettore w, che è la proiezione ortogonale di $\overrightarrow{P_0A}$ su W, possiamo considerare il punto

$$H := P_0 + w$$

che appartiene per costruzione a S in quanto $S = P_0 + W$. Mostriamo che tale punto ha la proprietà richiesta; abbiamo infatti:

$$\overrightarrow{AH} = H - A = P_0 + w - A = w + P_0 - A = w + (-w - u) = -u \in W^\perp.$$

Quindi $\overrightarrow{AH}$ è effettivamente perpendicolare al sottospazio. Riguardo l'unicità di H, supponiamo che H' sia un altro punto di S per cui $\overrightarrow{AH'}$ risulti perpendicolare a S; allora si avrebbe

$$\overrightarrow{HH'} = \overrightarrow{HA} + \overrightarrow{AH'} \in W^\perp.$$

Ma, essendo $H' \in S$, si avrebbe anche $\overrightarrow{HH'} \in W$. Dunque $\overrightarrow{HH'} = 0$ e quindi $H = H'$.

Infine, consideriamo un punto qualsiasi P di S diverso da H; per il Teorema di Pitagora (53), tenendo conto del fatto che i vettori $\overrightarrow{AH}$ e $\overrightarrow{HP}$ sono perpendicolari in quanto $\overrightarrow{HP} \in W$, si ricava:

$$d(A,P)^2 = d(A,H)^2 + d(H,P)^2 > d(A,H)^2$$

essendo $d(H,P) > 0$ perchè $P \neq H$. $\square$

Come caso rilevante, stabiliamo una formula per calcolare la distanza tra un punto ed un iperpiano:

Teorema 58.2. *Sia α un iperpiano di $\mathbb{R}^n$ avente equazione*

$$\alpha : a_1x_1 + \cdots + a_nx_n + d = 0$$

in un fissato riferimento Cartesiano. Allora per ogni punto $A(x_1^0, \ldots, x_n^0)$ si ha:

$$d(A, \alpha) = \frac{|a_1x_1^0 + \cdots + a_nx_n^0 + d|}{\sqrt{a_1^2 + \cdots + a_n^2}}.$$

Si osservi che la formula in questione dà correttamente che $d(A, \alpha) = 0$ esattamente nel caso in cui A sia un punto di α.

Dimostrazione: Si tratta di esplicitare la distanza $P(A, H)$ dove $H \in \alpha$ è la proiezione ortogonale di A sull'iperpiano. Cominciamo con l'osservare che, denotato con n il vettore di componenti $(a_1, \ldots, a_n)$ rispetto alla base del riferimento, in analogia a quanto si è visto nel caso dei piani di $\mathbb{R}^3$, risulta che n è un vettore non nullo perpendicolare ad α, e la giacitura W di α è

$$W = \{v \in \mathbb{R}^n | \, v \cdot n = 0\} = L(n)^\perp.$$

Inoltre, riscrivendo l'equazione di α usando il prodotto scalare, abbiamo:

$$\alpha = \{P \in \mathbb{R}^n : \ \overrightarrow{OP} \cdot n + d = 0\}$$

Dunque, essendo $\overrightarrow{AH}$ perpendicolare a α, abbiamo che $\overrightarrow{AH} \in L(n)$ e quindi il punto H è della forma

$$H = A + tn$$

per un opportuno scalare $t \in \mathbb{R}$, da determinarsi imponendo che tale punto appartenga all'iperpiano. Ciò accade se e solo se

$$\overrightarrow{OH} \cdot n + d = 0,$$

ovvero

$$((A - O) + tn) \cdot n + d = 0,$$

cioè

$$\overrightarrow{OA} \cdot n + t\, n \cdot n + d = 0,$$

da cui si ricava:

$$t = -\frac{\overrightarrow{OA} \cdot n + d}{||n||^2}.$$

Siamo ora in grado di calcolare la distanza tra A e H:

$$d(A, H) = ||\overrightarrow{AH}|| = \left\| -\frac{\overrightarrow{OA} \cdot n + d}{||n||^2} n \right\| = \frac{|\overrightarrow{OA} \cdot n + d|}{||n||},$$

ed esplicitando il rapporto ottenuto si ottiene subito la formula enunciata. $\square$

59. Cenno sulle isometrie

Definizione 59.1. Un'applicazione $f : \mathbb{R}^n \to \mathbb{R}^n$ si dice *isometria* (o movimento rigido) se conserva le distanze tra punti, cioè se

$$d(P, Q) = d(f(P), f(Q)), \quad \forall P, Q \in \mathbb{R}^n.$$

Nel caso $n = 2$ le isometrie sono i movimenti nel piano Euclideo che vengono utilizzati nelle argomentazioni di geometria sintetica concernenti la sovrapposizione di figure (preservandone la forma) e la nozione di congruenza (ad es. tra segmenti, angoli, triangoli, ecc).

L'importante risultato seguente caratterizza le isometrie dal punto di vista analitico:

Teorema 59.2. *Le isometrie $f : \mathbb{R}^n \to \mathbb{R}^n$ sono tutte e sole le funzioni del tipo*

$$f(X) = AX + b \quad \forall X \in \mathbb{R}^n$$

dove $A \in O(n)$ è una matrice ortogonale fissata e $b \in \mathbb{R}^n$ è un vettore fissato.

Come caso particolare abbiamo la traslazione $\tau_b : \mathbb{R}^n \to \mathbb{R}^n$ di vettore b definita da

$$\tau_b(X) := X + b,$$

corrispondente a $A = I_n$. Dunque ogni isometria è una funzione composta

$$f = \tau_b \circ L_A$$

di una traslazione e di un'isometria lineare $L_A : \mathbb{R}^n \to \mathbb{R}^n$.

Questo fatto implica che le isometrie sono tutte bigezioni, e che esse formano un gruppo di trasformazioni di $\mathbb{R}^n$.

Nel caso $n = 2$, un Teorema di Chasles afferma che vi sono solo quattro tipi di isometrie del piano: traslazioni (caratterizzate dal fatto che $det(A) = 1$ e di non ammettere punti fissi); rotazioni (caratterizzate dal fatto che $det(A) = 1$ e aventi punti fissi, di fatto un solo punto fisso, il centro della rotazione); riflessioni rispetto a una retta (caratterizzate da $det(A) = -1$ e aventi punti fissi, di fatto tutti e soli i punti della retta asse di riflessione); glissosimmetrie: sono isometrie composte da una riflessione seguita da una traslazione di vettore parallelo all'asse di simmetria (caratterizzate da $det(A) = -1$ e non aventi punti fissi).