Appunti per gli studenti del corso di Geometria
Corso di Laurea in Fisica
A.A. 2023/24

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 universale di "piano euclideo". In ultima analisi, ciò che tale identificazione permette è la manipolazione degli oggetti geometrici mediante le operazioni fondamentali tra numeri reali: la somma e il prodotto. In tale procedimento sono di basilare importanza le proprietà formali che tali operazioni hanno. Su di esse si basa la nozione generale di *campo*, che sarà discussa in questo paragrafo.

Sia X un qualsiasi insieme non vuoto.

Definizione 1.1. Si chiama **operazione interna** (o semplicemente **operazione**) su X ogni applicazione $* : X \times X \to X$.
Si utilizza anche il termine *legge di composizione interna* per lo stesso concetto.

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

Dati $a, b \in X$, piuttosto che scrivere

$$*((a, b))$$

si scrive

$$a * b.$$

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

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

Esempio 1.3. Dato un insieme A, consideriamo l'insieme A^A di tutte le applicazioni da A in A. Allora l'operazione di composizione

$$*((f,g)) := f \circ g$$

è un'operazione su A^A.

Esempio 1.4. Dato un insieme A, abbiamo sull'insieme delle parti $\mathcal{P}(A)$ le operazioni fondamentiali di intersezione e unione:

$$\cap : \mathcal{P}(A) \times \mathcal{P}(A) \to \mathcal{P}(A), \ \cup : \mathcal{P}(A) \times \mathcal{P}(A) \to \mathcal{P}(A).$$

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

1) L'operazione $* : G \times G \to G$ è **associativa**, nel senso che: per ogni $x, y, z \in G$ risulta

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

2) Esiste un **elemento neutro**, ovvero un elemento $e \in G$ tale che:

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

per ogni $x \in G$.

3) Ogni elemento $x \in G$ è **invertibile**, nel senso che, in corrispondenza di esso esiste un altro $x' \in G$, detto *inverso di* x, per cui si abbia:

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

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.

  In relazione a questi esempi, si osservi che, denotato con X l'insieme numerico in esame, l'operazione di prodotto di X induce effettivamente un'operazione interna anche su X^*, ovvero una funzione:

  $$\cdot : X^* \times X^* \to X^*.$$

  Essa è ben definita, perchè se $a \neq 0$ e $b \neq 0$, la *legge di annullamento del prodotto* garantisce che $a \cdot b \neq 0$, ovvero il risultato dell'operazione è ancora un elemento *dello stesso insieme* X^*.

Invece *non* sono gruppi:

- $(\mathbb{N},+)$, $(\mathbb{N},\cdot)$ e $(\mathbb{Z},\cdot)$.

A corredo della definizione precedente, proviamo il seguente risultato:

Teorema 1.6. *Ogni gruppo $(G,*)$ ammette un unico elemento neutro ed inoltre per ogni $x \in G$, l'inverso x' di x è univocamente determinato.*

Dimostrazione: Infatti, ammettendo che e ed e' siano entrambi elementi neutri, risulta
$$e = e * e' = e'.$$
La prima uguaglianza è giustificata dal fatto che e' è elemento neutro, la seconda dal fatto che e gode della stessa proprietà.

Supponiamo poi che x' e x'' siano entrambi inversi dello stesso elemento x. Allora in particolare abbiamo
$$x * x' = e.$$
Applicando l'operazione $*$, "componendo" ambo i membri 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 è
$$e * x' = x''$$
ovvero
$$x' = x''.$$
□

Esempio 1.7. Dato un insieme qualsiasi non vuoto X, l'insieme G di tutte le *bigezioni* $f : X \to X$ 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 X. Ad esempio, se $X = \{1, \ldots, n\}$ è un insieme finito, si tratta del gruppo delle permutazioni su n oggetti, studiato nell'ambito del calcolo combinatorio.

Il fatto che l'operazione di composizione dia luogo effettivamente ad una operazione
$$\circ : G \times G \to G$$

è garantito dal fatto che la composizione di due applicazioni ingettive è anch'essa ingettiva, e analogamente per le applicazioni surgettive; dunque se f, g sono bigettive, tale è $f \circ g$.

L'elemento neutro di tale gruppo è l'applicazione identica $Id_X : X \to X$, mentre per ogni bigezione $f : X \to X$, l'inverso è semplicemente la funzione inversa $f^{-1} : X \to X$; ricordiamo che tale funzione associa ad ogni $y \in X$ l'unico elemento $x \in X$ tale che:

$$f(x) = y,$$

e che la relazione che intercorre tra f e f^{-1} è proprio:

$$f \circ f^{-1} = Id_X,\ f^{-1} \circ f = Id_X.$$

Il lettore verifichi infine la validità della proprietà associativa:

$$(f \circ g) \circ h = f \circ (g \circ h)$$

che in realtà sussiste qualsiasi siano le funzioni da X in X coinvolte (non necessariamente bigettive).

Una generalizzazione molto importante dell'esempio precedente è fornita dalla nozione seguente:

Definizione 1.8. Si chiama *gruppo di trasformazioni* di un insieme X ogni insieme non vuoto Φ di bigezioni $f : X \to X$, verificante le condizioni seguenti:

a) Per ogni $f, g \in \Phi$, anche $f \circ g$ appartiene a Φ:

b) Per ogni $f \in \Phi$, anche la funzione inversa f^{-1} appartiene a Φ.

Sotto queste condizioni, si può considerare la struttura algebrica $(\Phi, \circ)$, in quanto la a) garantisce che l'operazione di composizione induce un'operazione interna a Φ (che ovviamente soddisfa ancora la proprietà associativa).

Risulta che $(\Phi, \circ)$ è un gruppo. Infatti, abbiamo che $Id_X \in \Phi$, in quanto, fissata una bigezione $f \in \Phi$, allora $Id_X = f \circ f^{-1}$ appartiene a Φ in forza di b) e di a). Ora, Id_X opera come elemento neutro anche in $(\Phi, \circ)$ e ogni $f \in \Phi$ risulta invertibile grazie alla b).

Esempio 1.9. L'insieme costituito da tutte e sole le rotazioni intorno ad un punto fissato O è un gruppo di trasformazioni del piano.

Data una retta r del piano euclideo X, detta $\sigma : X \to X$ la simmetria assiale rispetto a r, allora $\Phi = \{Id_X, \sigma\}$ costituisce un gruppo di trasformazioni di X. 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

$$\sigma \circ \sigma = Id_X$$

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 in Φ 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 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.

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.

Dato un gruppo $(G,\cdot)$ la cui operazione è denotata con notazione **moltiplicativa**, 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}.

Definizione 1.10. Un gruppo $(G,*)$ si dice **abeliano** se l'operazione $*$ gode della proprietà **commutativa**:

$$\forall a,b \in G \quad a*b=b*a.$$

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

Teorema 1.11. *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 c^{-1} 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

$$a = b.$$

□

Come applicazione, abbiamo la seguente proprietà: se $a \in G$ è tale che:

$$a * a = a,$$

allora $a = e$.

Infatti l'uguaglianza precedente si può interpretare come

$$a * a = a * e$$

e quindi, semplificando:

$$a = e.$$

Un'altra importante applicazione della legge di cancellazione è la seguente: per ogni elemento a di un gruppo $(G, *)$ denotiamo con $t_a : G \to G$ la funzione definita da:

$$t_a(x) := a * x.$$

Essa si chiama *traslazione* (a sinistra) relativa ad a. Il lettore verifichi che si tratta di una bigezione e che la sua inversa è la traslazione relativa all'inverso di a^{-1} di a. In simboli:

$$(t_a)^{-1} = t_{a^{-1}}.$$

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

Definizione 1.12. Si chiama *campo* ogni struttura algebrica $(K, +, \cdot)$ le cui operazioni soddisfano i seguenti requisiti:

a) $(K, +)$ è un gruppo abeliano.

b) Posto $K^* = K - \{0\}$, anche $(K^*, \cdot)$ è un gruppo abeliano.

c) Sussistono le proprietà distributive del prodotto rispetto alla somma, ovvero:

$$x \cdot (y + z) = x \cdot y + x \cdot z, \quad (y + z) \cdot x = y \cdot x + z \cdot x$$

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

Osserviamo esplicitamente che 0 denota qui l'elemento neutro del gruppo $(K, +)$; inoltre la b) include la condizione che l'operazione di prodotto si restringa ad un'operazione *interna* su K^*; ciò significa che si postula quanto segue: se $a \neq 0$ e $b \neq 0$, allora si ha sempre $a \cdot b \neq 0$. Questa è la legge di annullamento del prodotto, ed è implicita nella definizione di campo data sopra.

Denoteremo di norma con 1 l'elemento neutro di $(K^*, \cdot)$. Notiamo che in base alla definizione, $1 \neq 0$.

Proposizione 1.13. *Sia $(K, +, \cdot)$ un campo. Allora per ogni $x \in K$ si ha:*

$$0 \cdot x = 0 = x \cdot 0.$$

Dimostrazione: Abbiamo, utilizzando la prima delle proprietà distributive nella c) nella 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 $(K, +)$. Analogamente, sfruttando la seconda proprietà distributiva, possiamo scrivere:

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

da cui deduciamo che $0 \cdot x = 0$. $\square$

Questo risultato garantisce che l'operazione di moltiplicazione di un campo è sempre *commutativa e associativa* (sebbene tali proprietà siano postulate esplicitamente solo per K^*). Inoltre, $(K, \cdot)$ ammette 1 come elemento neutro. Però $(K, \cdot)$ *non è un gruppo*. Infatti, 0 non è invertibile in $(K, \cdot)$: nessun x può essere suo inverso, avendosi $0 \cdot x = 0 \neq 1$.

Un'altra proprietà naturale che si deduce dagli assiomi di campo è la seguente. Coerentemente con la convenzione generale, l'opposto di un elemento x, rispetto alla struttura di gruppo $(K, +)$, si denoterà con $-x$.

Proposizione 1.14. *Se $(K, +, \cdot)$ allora per ogni $x \in K$ risulta:*

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

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

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

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

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

□

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. Possiamo introdurre una struttura di $\mathbb{K}$-spazio vettoriale sul prodotto cartesiano $\mathbb{K}^2 := \mathbb{K} \times \mathbb{K}$ introducente 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 come segue:

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

In particolare $\mathbb{R}^2$ (il familiare piano cartesiano) è un fondamentale esempio di spazio vettoriale reale.

La verifica degli assiomi di spazio vettoriale per $(\mathbb{K}^2, +, \cdot)$ è un semplice esercizio. 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 ammettte 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 abliano. 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:

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

mentre

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

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

Esempio 2.5. 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

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

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

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

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

Il vettore nullo è:

$$(0, 0, \ldots, 0),$$

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

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

In simboli:

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

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

Esempio 2.6. *(Prodotti diretti di spazi vettoriali)* Possiamo sfruttare la stessa idea per costruire in modo naturale un nuovo spazio a partire da due spazi vettoriali assegnati.

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

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

nel modo seguente

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

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

Definizione 2.7. Lo spazio vettoriale $V_1 \times V_2$ dotato delle operazioni (1) si chiama *prodotto diretto* di V_1 e V_2.

La nozione di prodotto diretto si estende in modo ovvio al caso di più spazi vettoriali $V_1, \dots, V_n$ con $n \geq 2$. Sul prodotto cartesiano $V_1 \times V_2 \cdots \times V_n$, costituito per definizione da tutte le n-ple ordinate $(v_1, \dots, v_n)$ con $v_i \in V_i$ per ogni $i = 1, \dots, n$, si introduce una struttura di spazio vettoriale le cui operazioni sono definite in modo analogo al caso $n = 2$:

$$(v_1, v_2, \dots, v_n)+(w_1, w_2, \dots, w_n) := (v_1+w_1, \dots, v_n+w_n), \ \lambda(v_1, \dots v_n) := (\lambda v_1, \dots, \lambda v_n).$$

Il vettore nullo è $(0_{V_1}, \dots, 0_{V_n})$ e per ogni $v = (v_1, \dots, v_n) \in V_1 \times \cdots \times V_n$, il vettore $-v$ coincide con $(-v_1, \dots, -v_n)$.

Chiaramente questa tipologia di esempi include $\mathbb{K}^n$ che può essere pensato come prodotto di n spazi vettoriali tutti uguali a $\mathbb{K}$.

3. Significato geometrico delle operazioni di spazio vettoriale di $\mathbb{R}^2$

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.

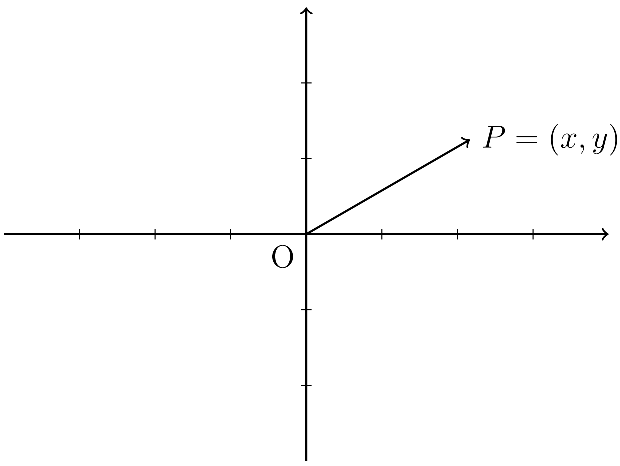
Dato un vettore $\overrightarrow{AB}$ e un punto X, esiste un unico vettore $\overrightarrow{XY}$ che ha lo stesse caratteristiche di $\overrightarrow{AB}$: esso si dice *equipollente ad* $\overrightarrow{AB}$. I vettori equipollenti $\overrightarrow{AB}$ e $\overrightarrow{XY}$ sono due rappresentanti 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 alla nozione di vettore di $\mathbb{R}^2$ e alle operazioni che ne realizzano la struttura di spazio vettoriale. Pensiamo innanzitutto a $\mathbb{R}^2$ come un piano cartesiano, il cui riferimento ha origine in un dato punto O. Sappiamo che ogni $P = (x, y) \in \mathbb{R}^2$ si identifica (e si può rappresentare graficamente come) un punto del piano, avente ascissa x e ordinata y nel nostro riferimento. In particolare $(0, 0)$ è l'origine O.

Per interpretare invece P come vettore, lo rappresenteremo mediante il vettore applicato $\vec{OP}$. 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 $\mathbb{R}^2$:



Per rendersi conto dell'utilità e dell'interpretazione geometrica della somma tra vettori di $\mathbb{R}^2$ e della moltiplicazione per gli scalari, cominciamo discutendo il seguente fatto:

Proposizione 3.1. *Siano $A \neq B$ e $C \neq D$ punti di $\mathbb{R}^2$. Allora*
La retta per A e B è parallela alla retta per C e $D \Longleftrightarrow \exists t \in \mathbb{R} : B - A = t(D - C)$.

Questo risultato esprime il concetto puramente geometrico di *direzione* nel piano mediante l'operazione puramente algebrica di differenza di vettori di $\mathbb{R}^2$, e la moltiplicazione per scalari.

Dimostrazione: Esamineremo il caso generico in cui le rette nell'enunciato siano entrambe non parallele all'asse delle ordinate. Si invita il lettore a completare la verifica nel caso in cui una della rette sia parallela allo stesso asse.

Dunque, assunto che $A = (a_1, a_2)$, $B = (b_1, b_2)$, $C = (c_1, c_2)$ e $D = (d_1, d_2)$, stiamo assumendo $a_1 \neq b_1$ e $c_1 \neq d_1$. Siano r la retta per A e B e r' la retta per B e C. Allora r e r' hanno coefficienti angolari m e m' dati dai rapporti:

$$m = \frac{b_2 - a_2}{b_1 - a_1},\ m' = \frac{d_2 - c_2}{d_1 - c_1}.$$

Pertanto r e r' sono parallele se e solo se

$$\frac{b_2 - a_2}{b_1 - a_1} = \frac{d_2 - c_2}{d_1 - c_1}.$$

D'altra parte, operando nello spazio vettoriale $\mathbb{R}^2$ abbiamo:

$$B - A = (b_1 - a_1, b_2 - a_2),\ D - C = (d_1 - c_1, d_2 - c_2).$$

Ora possiamo fare la seguente osservazione di carattere generale: siano $u = (x_1, y_1)$ e $v = (x_2, y_2)$ vettori di $\mathbb{R}^2$, tali che $x_1 \neq 0$ e $x_2 \neq 0$. Allora si ha:

$$\frac{y_1}{x_1} = \frac{y_2}{x_2} \iff \exists t \in \mathbb{R} :\ u = tv. \tag{$*$}$$

Applicando questo principio ai vettori $u = B - A$ e $v = D - C$ consegue quindi la tesi. Resta da spiegare l'equivalenza in (*). Supponiamo dapprima che $u = tv$ per un opportuno scalare t. Allora $t \neq 0$, perchè altrimenti avremmo $u = 0(x_2, y_2) = (0, 0)$ contro il fatto che $x_1 \neq 0$. Esplicitando la relazione $u = tv$ abbiamo:

$$\begin{cases} x_1 = tx_2 \\ y_1 = ty_2 \end{cases}.$$

Allora uguagliando i rapporti tra primi membri e i secondi membri segue che:

$$\frac{y_1}{x_1} = \frac{ty_2}{tx_2} = \frac{y_2}{x_2}$$

come volevasi. Viceversa, si assuma che valga $\frac{y_1}{x_1} = \frac{y_2}{x_2}$ e denotiamo tali rapporti con m. Allora $y_1 = mx_1$ e possiamo riscrivere il vettore u come:

$$u = (x_1, y_1) = (x_1, mx_1) = x_1(1, m)$$

e analogamente

$$v = x_2(1, m).$$

Possiamo allora ricavare u in funzione di v come segue:

$$u = \frac{x_1}{x_2} x_2(1, m) = tv$$

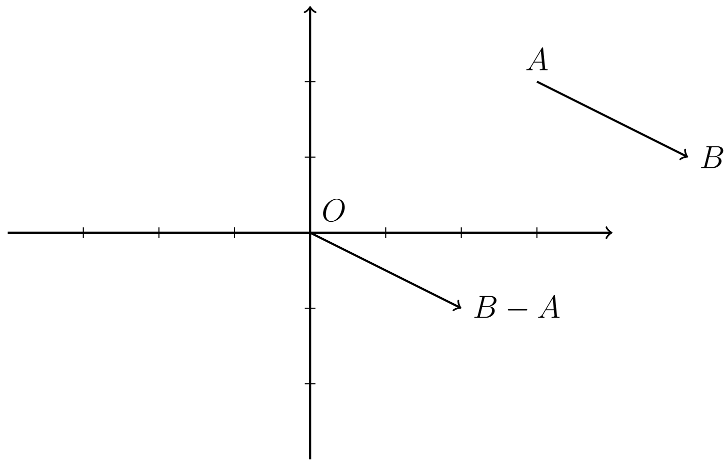
dove abbiamo posto $t := \frac{x_1}{x_2}$. $\square$

In modo simile, anche l'equipollenza tra vettori geometrici si può interpretare usando semplicemente le differenze tra vettori:

Proposizione 3.2. *Dati i punti $A, B, C, D \in \mathbb{R}^2$, allora*

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

In particolare, ciò significa che, a meno di equipollenza la differenza $B - A$ rappresenta il vettore $\overrightarrow{AB}$: precisamente tale vettore è equipollente a $\overrightarrow{O(B-A)}$.



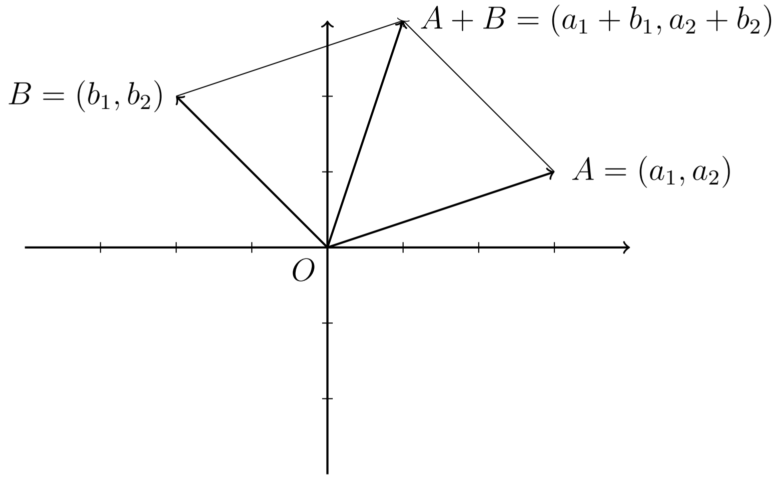
Tralasciamo la verifica dell'enunciato della Proposizione precedente, che è un altro esercizio di geometria analitica nel piano cartesiano. Notiamo solo che il fatto che i segmenti AB e CD siano congruenti si esprime mediante l'uguaglianza delle distanze $d(A, B) = d(C, D)$; in forza della relazione pitagorica:

$$d(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2},$$

e l'analoga per la distanza tra C e D, vediamo subito come siano naturalmente coinvolti i vettori $B - A$ e $D - C$.

A questo punto possiamo dare anche un'interpretazione geometrica della somma di vettori di $\mathbb{R}^2$:

Proposizione 3.3. *Siano A, B vettori di $\mathbb{R}^2$. Allora la somma $A + B \in \mathbb{R}^2$ coincide con il vettore somma $\overrightarrow{OA} + \overrightarrow{OB}$ secondo la regola del parallelogramma.*



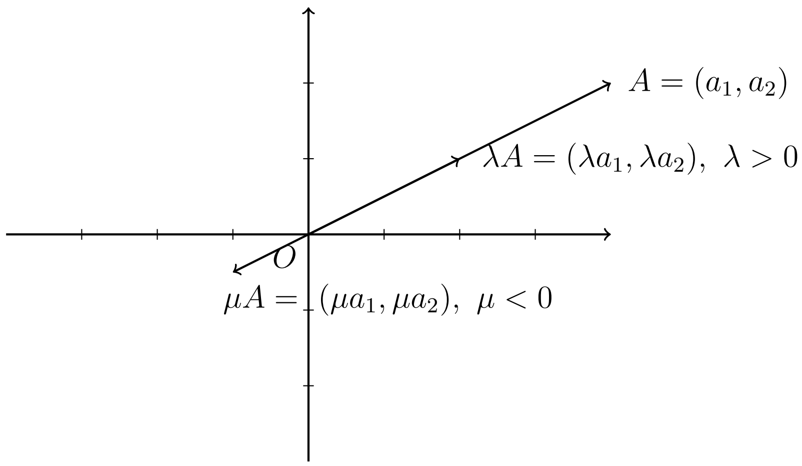
Dimostrazione: Considerando il caso generico in cui O, A, B non sono allineati, si tratta di verificare che $O, B, A + B, A$ sono vertici di un parallelogramma, ovvero che i vettori $\overrightarrow{OA}$ e $\overrightarrow{B(A + B)}$ sono equipollenti. Possiamo fare ciò usando il criterio enunciato nella Proposizione 3.2: infatti abbiamo

$$A - O = (A + B) - B.$$

Alternativamente, il lettore può procedere con una verifica diretta che $O, B, A, A+B$ è un parallelogramma utilizzando le coordinate e controllando il parallelismo tra la retta per O, A e la retta per B e $A + B$ ed il parallelismo tra la retta per O, B e la retta per A e $A + B$. $\square$

Infine, l'operazione di moltiplicazione per uno scalare ha la consueta interpretazione di "riscalamento" di un segmento orientato:

Proposizione 3.4. *Dato un vettore $A \in \mathbb{R}^2$, e uno scalare $\lambda \in \mathbb{R}$ non nullo, allora λA 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 $\lambda < 0$.*



4. Spazi vettoriali di funzioni

Introduciamo 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 le operazioni di somma $+$ e di prodotto per scalari $\cdot$: date due funzioni $f : X \to V$ e $g : X \to V$, allora la funzione somma $f+g$ si definisce come la funzione $f + g : X \to V$ 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 λf si definisce come la funzione $\lambda f : X \to V$ tale che:

$$(\lambda f)(x) := \lambda f(x), \quad \forall x \in X.$$

Risulta allora che $(V^X, +, \cdot)$ è anch'esso uno spazio vettoriale sul campo $\mathbb{K}$.

Il vettore nullo di tale spazio è la funzione costante di valore 0_V, che denoteremo semplicemente con il simbolo 0; dunque si tratta della funzione

$$0(x) := 0_V, \quad \forall x \in V,$$

che associa ad ogni x il vettore nullo di V.

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

$$f'(x) := -f(x)$$

che associa cioè ad ogni punto x il vettore opposto di $f(x)$ in V.

La verifica della validità di tutti degli assiomi di spazio vettoriale è un esercizio molto istruttivo per prendere confidenza con la nozione di spazio vettoriale.

5. Proprietà elementari degli spazi vettoriali

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

Proposizione 5.1. *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 $(V,+)$, 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 che $\lambda \neq 0$; 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 $v = 0_V$.

□

6. 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 6.1. 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:

$$\forall u, v \in V \qquad f(u+v) = f(u) + f(v). \tag{2}$$

$$\forall u \in V \; \forall \lambda \in \mathbb{K} \qquad f(\lambda u) = \lambda f(u). \tag{3}$$

Ogni applicazione lineare deve soddisfare il seguente vincolo:

Proposizione 6.2. *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 6.3. Dati due $\mathbb{K}$-spazi vettoriali V e V', l'applicazione nulla

$$0 : V \to V'$$

(si veda il paragrafo precedente) è lineare.

Esempio 6.4. 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. Anche la funzione $x \mapsto x^2$ non è lineare.

Si noti la seguente caratterizzazione:

Proposizione 6.5. *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: $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$.

Dimostrazione: La dimostrazione è un esercizio. □

Fissati i vettori u e v ogni vettore della forma $\lambda u + \mu v$ è detto una *combinazione lineare* di u e v. Questo concetto viene generalizzato nel paragrafo seguente.

7. Combinazioni lineari di vettori

Definizione 7.1. Siano V un $\mathbb{K}$-spazio vettoriale e

$$v_1, \dots, v_k$$

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

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

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

Esempio 7.2. 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.

Proposizione 7.3. *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 6.5. Una dimostrazione più rigorosa si può fare agevolmente per induzione sul numero k dei vettori coinvolti.

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

Lo spazio $\mathbb{K}^n$ 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 8.1. Sia $n \geq 1$. Si chiama *base canonica* di $\mathbb{K}^n$ l'insieme $\{e_1, \ldots, e_n\}$ costituito dai vettori così definiti:

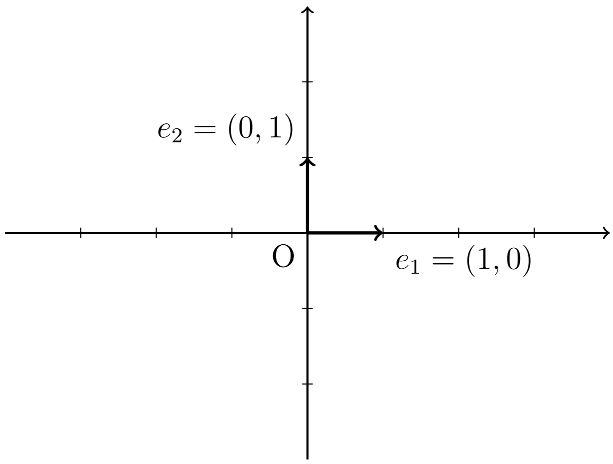
$$e_1 := (1, 0, \ldots, 0),\ e_2 := (0, 1, 0, \ldots, 0), \ldots, e_n := (0, 0, \ldots, 0, 1). \tag{4}$$

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 8.2. *Ogni vettore* $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ *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

$$\begin{aligned}
&x_1 e_1 + \cdots + x_n e_n = \\
&x_1(1, 0, \dots, 0) + x_2(0, 1, \dots, 0) + \cdots + x_n(0, 0, \dots, 0, 1) = \\
&(x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \cdots + (0, 0, \dots, 0, x_n) = \\
&(x_1 + 0 + \cdots + 0, 0 + x_2 + 0 + \cdots + 0, \cdots, 0 + \cdots + 0 + x_n) = \\
&(x_1, x_2, \dots, x_n).
\end{aligned}$$

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

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n$$

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

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

e pertanto necessariamente $\lambda_i = x_i$ per ogni $i = 1, \dots, n$. $\square$

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 8.3. *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*

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

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 del tipo (6) è effettivamente lineare è un esercizio. Inoltre è immediato dalla definizione di una tale f che $f(e_i) = a_i$: l'unica coordinata non nulla di e_i è la i-ma e quindi $f(e_i) = a_i \cdot 1 = a_i$.

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 7.3:

$$f(x) = x_1f(e_1) + \cdots + x_nf(e_n)$$

da cui si ottiene che f è della forma richiesta, ponendo $a_i := f(e_i)$ per ogni i. $\square$

Esempio 8.4. (Proiezioni) In particolare, le n funzioni

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

sono tutte lineari. Esse si chiamano *proiezioni naturali* o semplicemente *proiezioni*.

9. Isomorfismi

La seguente proprietà è fondamentale nella teoria degli spazi vettoriali.

Teorema 9.1. *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 $u, v \in V$; allora

$$(g\circ f)(u+v) = g(f(u+v)) = g(f(u)+f(v)) = g(f(u))+g(f(v)) = (g\circ f)(u)+(g\circ f)(v).$$

In modo del tutto analogo si prova che $(g \circ f)(\lambda u) = \lambda(g \circ f)(u)$ per ogni $\lambda \in \mathbb{K}$. $\square$

Teorema 9.2. *Siano V, V' spazi vettoriali sullo stesso campo $\mathbb{K}$.*

Se $f : V \to V'$ è bigettiva, anche l'applicazione inversa $f^{-1} : V' \to V$ è lineare.

Dimostrazione: Siano $u', v' \in V'$; allora, stante la bigettività di f tali vettori sono della forma

$$u' = f(u), \quad v' = f(v)$$

per opportuni u, v vettori di V, univocamente determinati. Allora abbiamo, sfruttando la linearità di f:

$$f^{-1}(u' + v') = f^{-1}(f(u) + f(v)) = f^{-1}(f(u + v)) = u + v = f^{-1}(u') + f^{-1}(v').$$

Nell'ultimo passaggio si è tenuto conto del fatto che $u = f^{-1}(u')$ e analogamente per v. In modo simile si prova che $f^{-1}(\lambda u') = \lambda f^{-1}(u')$ per ogni $\lambda \in \mathbb{K}$. $\square$

Definizione 9.3. Un'applicazione lineare bigettiva $f : V \to V'$ tra due $\mathbb{K}$-spazi vettoriali prende il nome di *isomorfismo*.

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:

$$f : V \xrightarrow{\cong} V'.$$

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.

Il fatto che l'inverso di un isomorfismo è anch'esso un isomorfismo garantisce che, se $V \cong V'$, allora è vero anche che $V' \cong V$ (proprietà *simmetrica* della relazione di isomorfismo).

Come conseguenza del Teorema 9.1 abbiamo anche che, se se V, V' e V'' sono spazi vettoriali sullo stesso campo $\mathbb{K}$, e si sa che $V \cong V'$, $V' \cong V''$, allora si può concludere che si ha anche $V \cong V''$ (proprietà *transitiva* della relazione di isomorfismo).

Assegnato dunque un campo $\mathbb{K}$, la relazione di isomorfismo permette di "classificare" tutti gli spazi vettoriali su $\mathbb{K}$. Nel seguito ci occuperemo di tale classificazione nel caso di spazi vettoriali di dimensione finita, dando un preciso significato matematico all'attributo "dimensione".

Un isomorfismo $f : V \to V$ prende il nome anche di *automorfismo* dello spazio vettoriale V. Un esempio basilare è l'applicazione identica

$$Id_V : V \to V.$$

Proposizione 9.4. *Sia V un $\mathbb{K}$-spazio vettoriale. L'insieme $GL(V)$ di tutti i suoi automorfismi è un gruppo di trasformazioni di V.*

Dimostrazione: Abbiamo appena osservato che $Id_V \in GL(V)$, per cui tale insieme è non vuoto. Inoltre la composizione $f \circ g$ di due automorfismi f e g è anch'esso un automorfismo, ovvero $f \circ g \in GL(V)$, stante ancora il Teorema 9.1. Analogamente, il Teorema 9.2 garantisce che per ogni $f \in GL(V)$ si ha che $f^{-1} \in GL(V)$. $\square$

Il gruppo $(GL(V), \circ)$ è noto come *gruppo lineare generale* su V.

10. Nucleo ed immagine di un'applicazione lineare

Definizione 10.1. Sia $f : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Si chiama *nucleo di* f il seguente sottoinsieme di V:

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

Si chiama *immagine di* f il seguente sottoinsieme di V':

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

L'importanza del nucleo è stabilita dal seguente risultato di caratterizzazione delle applicazioni lineari ingettive:

Teorema 10.2. *Sia $f : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Allora*

$$f \text{ è ingettiva} \iff Ker(f) = \{0_V\}.$$

Dimostrazione: Supponiamo f ingettiva. Sia $u \in Ker(f)$; allora $f(u) = 0_{V'}$, ovvero stante la Prop. 6.2, $f(u) = f(0_V)$, da cui $u = 0_V$.

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

$$f(u) = f(v).$$

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

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

ovvero per la linearità di f:

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

pertanto $u - v \in Ker(f)$ e quindi per l'ipotesi $u - v = 0_V$, ovvero $u = v$. □

Esempio 10.3. Il nucleo della funzione lineare $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x, y) := 2x + \sqrt{3}y$$

si può descrivere come segue:

$$Ker(f) = \{(t, -\frac{2}{\sqrt{3}}t) | \, t \in \mathbb{R}\}$$

oppure

$$Ker(f) = \{(-\frac{\sqrt{3}}{2}s, s) | \, s \in \mathbb{R}\}.$$

In entrambe le descrizioni, vediamo che tutti i vettori del nucleo dipendono da un solo parametro reale.

Definizione 10.4. Un'applicazione lineare ingettiva verrà detta *monomorfismo*. Un'applicazione lineare surgettiva verrà invece detta *epimorfismo*.

Osserviamo che, per definizione, $f : V \to V'$ è un epimorfismo se e solo se $Im(f) = V'$.

11. Sottospazi

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

Definizione 11.1. Un sottoinsieme non vuoto S di V si dice *sottospazio vettoriale* di V se sono soddisfatte le seguenti condizioni:

a) $\forall u, v \in S \quad u + v \in S$.
b) $\forall u \in S \, \forall \lambda \in \mathbb{K} \quad \lambda u \in S$.

La terminologia "sottospazio" è chiarita dalla seguente proposizione.

Proposizione 11.2. *Sia $S \subset V$ un sottospazio vettoriale. Allora si ha che $0_V \in S$. Inoltre S 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: Le condizioni a) e b) permettono di restringere le operazioni di somma e prodotto esterno su S, ottenendo nuove operazioni

$$S \times S \to S, \ \mathbb{K} \times S \to S.$$

È evidente che l'operazione di somma così ottenuta continua a essere associativa e commutativa. Inoltre l'operazione di prodotto esterno verifica le proprietà B) della definizione di spazio vettoriale, ed anche le C) sono immediatamente soddisfatte. Resta da chiarire che $(S, +)$ è un gruppo. Sia $u \in S$; allora, utilizzando la b) abbiamo che anche ilvettore opposto $-u$ appartiene ad S, in quanto $-u = (-1)u \in S$ (Prop. 5.1). In particolare, fissato un elemento $u \in S$ (è possibile perchè S non è vuoto), usando a) si ottiene che $0_V \in S$; infatti:

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

A questo punto è immediato constatare che $(S, +)$ è un gruppo con elemento neutro 0_V. $\square$

Esempio 11.3. 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, se r non è parallela all'asse delle ordinate, essa è un sottoinsieme del tipo

$$r = \{(x, mx) | \, x \in \mathbb{R}\}$$

(dove m è il coefficiente angolare della retta stessa). Dati quindi due vettori (x, mx) e (x', mx') e uno scalare λ, si ha:

$$(x, mx) + (x', mx') = (x + x', m(x + x')) \in r$$

e

$$\lambda(x, mx) = (\lambda x, m\lambda x) \in r.$$

Il lettore verifichi che anche l'asse delle ordinate $s = \{(0, y) | \, y \in \mathbb{R}\}$ è un sottospazio.

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.

Teorema 11.4. *Sia $f : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali. Allora $Ker(f)$ è sottospazio vettoriale di V, mentre $Im(f)$ è sottospazio di V'.*

DIMOSTRAZIONE: Infatti, sappiamo che $0_V \in Ker(f)$, per cui $Ker(f)$ è non vuoto. Siano $u, v \in \mathrm{Ker}(f)$; allora

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

per cui $u + v \in \mathrm{Ker}(f)$: inoltre per ogni scalare λ abbiamo:

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

In modo molto simile si procede per $Im(f)$. $\square$

ESEMPIO 11.5. Consideriamo nuovamente la funzione lineare $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x, y) := 2x + \sqrt{3}y.$$

Il suo nucleo

$$Ker(f) = \{(x, y) \in \mathbb{R}^2 | \, 2x + \sqrt{3}y = 0\}$$

è quindi un sottospazio vettoriale di $\mathbb{R}^2$. Questo fatto è in accordo anche con quanto discusso nell'Esempio 11.3, perchè sappiamo che tale insieme si interpreta come una retta per l'origine. In particolare, $Ker(f)$ è esso stesso uno spazio vettoriale reale, stante la Prop. 23.1. Mostriamo che tale spazio è isomorfo a $\mathbb{R}$. Possiamo definire in modo naturale due isomorfismi:

$$p : Ker(f) \to \mathbb{R}, \quad h : \mathbb{R} \to Ker(f)$$

come segue:

$$p(x, y) := x, \quad h(t) := (t, -\frac{2}{\sqrt{3}}t).$$

Il lettore verifichi che si tratta effettivamente di isomorfismi (uno inverso dell'altro), controllando per entrambi la linearità e la bigettività. Ad esempio, p risulta ingettivo perchè il suo nucleo è banale: infatti se $(x, y) \in Ker(f)$ e $p(x, y) = 0$, allora $x = 0$, ma ciò comporta anche che $y = -\frac{2}{\sqrt{3}}x = 0$, dunque $Ker(p) = \{(0, 0)\}$.

Il primo isomorfimo è la restrizione di una proiezione: i punti della retta vengono proiettati sull'asse delle ascisse. La funzione h è suggerita dalla rappresentazione parametrica di $Ker(f)$ discussa in precedenza. Ogni punto $t \in \mathbb{R}$ dell'asse delle ascisse determina un punto sulla retta, essendo questa il grafico di una funzione di variabile reale.

12. Somma e intersezione di sottospazi

Siano assegnati un $\mathbb{K}$-spazio vettoriale V e due suoi sottospazi vettoriali U_1 e U_2.

Definizione 12.1. Si chiama *somma* di U_1 ed U_2 il sottoinsieme di V:

$$U_1 + U_2 := \{u_1 + u_2 \,|\, u_1 \in U_1,\ u_2 \in U_2\}.$$

Proposizione 12.2. *La somma $U_1 + U_2$ è un sottospazio vettoriale di V.*

Dimostrazione: Naturalmente $U_1 + U_2 \neq \emptyset$ perchè entrambi gli spazi U_i, $i = 1, 2$ non sono vuoti. Siano $z, w \in U_1 + U_2$. Allora $z = u_1 + u_2$ per opportuni vettori $u_i \in U_i$ e $w = v_1 + v_2$ per certi $v_i \in U_i$. Segue, in forza delle proprietà associativa e commutativa della somma:

$$z + w = (u_1 + u_2) + (v_1 + v_2) = (u_1 + v_1) + (u_2 + v_2)$$

il che mostra che $z + w \in U_1 + U_2$ in quanto, essendo gli U_i sottspazi, $u_i + v_i \in U_i$. In modo analogo si verifica che per ogni $\lambda \in \mathbb{K}$ risulta $\lambda z \in U_1 + U_2$. $\square$

Si osservi che risulta:

$$U_1 \subset U_1 + U_2, \quad U_1 \subset U_1 + U_2,$$

perchè ogni vettore u di U_1 può essere scritto nella forma $u = u + 0_V$ e analogamente per U_2.

Definizione 12.3. Dati i sottospazi U_1 e U_2, la somma $U_1 + U_2$ verrà detta *diretta* se

$$U_1 \cap U_2 = \{0_V\}. \tag{7}$$

In tal caso, in luogo di $U_1 + U_2$ si scriverà $U_1 \oplus U_2$.

Se accade che

$$V = U_1 \oplus U_2, \tag{8}$$

diremo che lo spazio V è somma diretta dei suoi sottospazi U_1 e U_2.

Esempio 12.4. Pensando a $\mathbb{R}^2$ come piano cartesiano, due rette distinte r e s passanti per l'origine sono due sottospazi vettoriali la cui somma è diretta. Se r è l'asse delle ascisse e s è l'asse delle ordinate, allora abbiamo:

$$\mathbb{R}^2 = r \oplus s$$

perchè ogni vettore (x, y) si scrive nella forma:

$$(x, y) = (x, 0) + (0, y)$$

con $(x, 0) \in r$ e $(0, y) \in s$.

Sussiste la seguente caratterizzazione della condizione (8).

Proposizione 12.5. *Dati i sottospazi* U_1 *e* U_2 *di* V*, le seguenti proprietà sono equivalenti:*

a) $V = U_1 \oplus U_2$;

b) *Ogni vettore* $v \in V$ *si scrive in modo unico nella forma* $v = u_1 + u_2$ *con* $u_i \in U_i,\ i = 1, 2$.

Dimostrazione: *a)* $\Rightarrow$ *b)*. Stante la definizione di $U_1 + U_2$, assumendo a) è chiaro che ogni $v \in V$ si può scrivere come enunciato in b); supponiamo che ciò possa farsi in due modi diversi $v = u_1 + u_2 = v_1 + v_2$, con ovvio significato dei simboli. Allora $u_1 - v_1 = v_2 - u_2$ ed il vettore con cui coincidono ambo i membri appartiene sia ad U_1 che ad U_2 in virtù delle proprietà dei sottospazi. Pertanto, per la a), si ha $u_1 - v_1 = 0_V = v_2 - u_2$ e di qui la b).

b) $\Rightarrow$ *a)*. Sia $z \in U_1 \cap U_2$; poichè

$$z = z + 0_V = 0_V + z$$

applicando la *b)*, dal confronto di tali rappresentazioni di v si ottiene che necessariamente $z = 0_V$. $\square$

Proposizione 12.6. *Dati due sottospazi* U_1 *e* U_2 *di uno spazio vettoriale* V*, si ha che anche la loro intersezione:*

$$U_1 \cap U_2$$

è un sottospazio vettoriale.

Dimostrazione: L'intersezione in questione è non vuota, in quanto il vettore nullo appartiene a entrambi i sottospazi. Dati due vettori u e w in $U_1 \cap U_2$ e due scalari λ, μ, allora la combinazione lineare $\lambda u + \mu v$ appartiene ad U_1, in quanto U_1 è sottospazio, e i vettori in questione vi appartengono; per lo stesso motivo $\lambda u + \mu v$ appartiene anche a U_2, dunque appartiene a $U_1 \cap U_2$. $\square$

13. Spazi vettoriali di applicazioni lineari

In questa sezione ci occupiamo di ulteriori importanti proprietà delle applicazioni lineari tra due spazi vettoriali, che coinvolgono le operazioni di somma di funzioni e moltiplicazione di scalari per funzioni, già considerate nel paragrafo 4.

Proposizione 13.1. *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. $\square$

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

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

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

In particolare quindi 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 $0 : V \to W$.

Come esempio fondamentale, studiamo più in dettaglio $Hom(\mathbb{K}^n,\mathbb{K}^m)$, determinandone in modo esplicito gli elementi:

Teorema 13.3. *Dati due interi $n,m \geq 1$ le funzioni lineari $F : \mathbb{K}^n \to \mathbb{K}^m$ sono tutte e sole le funzioni del tipo*

$$F(x) := (F_1(x), \cdots, F_m(x)), \quad \forall x = (x_1,\ldots,x_n) \in \mathbb{K}^n, \tag{9}$$

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 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 8.3.

Dimostrazione: Ogni funzione del tipo (9) 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 $F : \mathbb{K}^n \to \mathbb{K}^m$: allora, per ogni $x \in \mathbb{K}^n$, le m coordinate $F_1(x), \ldots, F_m(x)$ del vettore $F(x) \in \mathbb{K}^m$ sono univocamente determinate; quindi restano definite m funzioni scalari $F_1, \ldots, F_m$: $\mathbb{K}^m \to \mathbb{K}$; per costruzione, abbiamo quindi:

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

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

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

Poichè p_i è lineare (cfr. Es. 8.4), per concludere basta utilizzare il fatto che la composizione di applicazioni lineari è lineare. □

14. Applicazioni lineari e sottospazi

Proposizione 14.1. *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

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

b) Ricordiamo che, per definizione:

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

e chiaramente tale insieme è non vuoto perchè tale è U (ad es. abbiamo certamente che $0'_V = f(0_V) \in f(U)$ avendosi $0_V \in U$).

Fissiamo due vettori arbitrari $f(u_1)$ e $f(u_2)$ di $f(U)$, dove u_1, u_2 sono vettori di U e consideriamo una loro combinazione lineare con coefficienti λ, μ. Tale combinazione si può scrivere:

$$\lambda f(u_1) + \mu f(u_2) = f(\lambda u_1 + \mu u_2),$$

ma tale vettore appartiene ancora a $f(U)$, in quanto, essendo U sottospazio, il vettore $\lambda u_1 + \mu u_2$ appartiene a U. $\square$

Notiamo che, considerando $U = V$ nell'enuniciato b) della Proposizione precedente, riotteniamo il fatto che $Im(f) = f(V)$ è sottospazio di V'.

15. 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 costruire una copia isomorfa del primo nel secondo.

Teorema 15.1. *Siano n, m interi con $n > m \geq 1$. Allora non esistono applicazioni lineari ingettive $f : \mathbb{K}^n \to \mathbb{K}^m$.*

Dimostrazione: Si tratta di provare che, per ogni $n \geq 2$, ogni applicazione lineare

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

con $m < n$ ha nucleo non banale, ovvero che esiste un vettore $x \neq 0$ tale che $F(x) = 0$. Sappiamo dal Teorema 9 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.

Quindi vogliamo provare che, per ogni $n \geq 2$ ogni il sistema di m equazioni

$$\begin{cases} F_1(x) = 0 \\ \quad\vdots \\ F_m(x) = 0 \end{cases} \tag{10}$$

con $m < n$ ha sempre una soluzione non banale $x = (x_1, \ldots, x_n)$ cioè una soluzione costituita da n scalari $x_1, \ldots, x_n$ non tutti nulli.

L'idea della dimostrazione è che un tale sistema si può risolvere con il metodo di sostituzione, partendo dalla prima equazione che non sia un'identità della forma $0 = 0$, ricavando una delle incognite in funzione di tutte le altre; sostituendo nelle altre equazioni, si ottiene un sistema che coinvolge solo tali incognite, e si può reiterare il procedimento al più m volte complessive, risolvendo, in sequenza, ciascuna equazione rispetto *ad una sola incognita.* Al termine di questo processo, resta un set di variabili indipendenti o libere, di modo che tutte le altre sono funzioni lineari di queste. Per costruzione, tutte le soluzioni del sistema si ottengono attribuendo valori *arbitrari* alle incognite *libere.* Il fatto che $m < n$ garantisce che vi è almeno un incognita libera, a cui si può attribuire un valore non nullo, ad es. 1, per ottenere un vettore $x \neq 0$ che risolve il sistema.

Nel seguito forniamo una dimostrazione rigorosa del teorema basata su questa idea, provando per induzione su $n \geq 2$, che ogni sistema del tipo (41.4), con $m < n$ e le $F_i : \mathbb{K}^n \to \mathbb{K}$ tutte funzioni lineari, ha sempre almeno una soluzione non banale.

A tal scopo, ricordiamo che ogni $F : \mathbb{K}^n \to \mathbb{K}$ è una funzione polinomiale del tipo

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

Fissato un indice j, $1 \leq j \leq n$, si può scrivere:

$$F(x) = a_jx_j + (a_1x_1 + \cdots a_{j-1}x_{j-1} + a_{j+1}x_{j+1} + \cdots + a_nx_n)$$

per cui esiste resta determinata una funzione lineare $G : \mathbb{K}^{n-1} \to \mathbb{K}$ tale che per ogni x:

$$F(x) = a_jx_j + G(\bar{x}),$$

dove $\bar{x} := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ è il vettore di $\mathbb{K}^{n-1}$ che si ottiene da x sopprimendo la coordinata j-ma. Ad esempio, se $\mathbb{K} = \mathbb{R}$ e $F : \mathbb{R}^3 \to \mathbb{R}$ è il funzionale $F(x_1, x_2, x_3) = -3x_1 + \sqrt{2}x_2 + x_3$, allora possiamo scrivere

$$F(x) = \sqrt{2}x_2 + G(x_1, x_3)$$

con $G : \mathbb{R}^2 \to \mathbb{R}$ lineare. Notiamo anche che, se $F \neq 0$, allora almeno uno degli a_j dev'essere non nullo.

Ciò premesso, consideriamo il passo base $n = 2$ della nostra dimostrazione per induzione. In tal caso necessariamente $m = 1$ e si tratta di controllare che una singola equazione

$$F_1(x) = 0$$

con $x \in \mathbb{K}^2$ ha una soluzione $(x_1, x_2) \neq (0, 0)$. Ciò è garantito dal fatto che $\mathbb{K}$ è un campo, in quanto l'equazione in questione è della forma:

$$ax_1 + bx_2 = 0,$$

per cui, o $a = b = 0$ e quindi qualsiasi coppia (x_1, x_2) risolve l'equazione (ad es. $(1, 0) \neq (0, 0)$), oppure almeno uno di tali coefficienti è non nullo: supponendo ad esempio $a \neq 0$, l'equazione si può risolvere ricavando x_1 come segue:

$$x_1 = -a^{-1}bx_2,$$

e quindi tutte e sole le soluzioni sono i vettori $(-a^{-1}bx_2, x_2)$ al variare di x_2 in $\mathbb{K}$. In particolare, in corrispondenza di $x_2 = 1$ abbiamo una soluzione non banale $(-a^{-1}b, 1)$.

Discutiamo ora il passo induttivo, supponendo che l'asserto sia vero per ogni $n \geq 2$ e proviamolo per $n + 1$. Consideriamo quindi un sistema del tipo

$$\begin{cases} F_1(x) = 0 \\ \quad\vdots \\ F_m(x) = 0 \end{cases} \tag{11}$$

dove le F_i sono applicazioni lineari $F_i : \mathbb{K}^{n+1} \to \mathbb{K}$ e $1 \leq m < n + 1$. Se tutte queste funzioni sono la funzione nulla, non cè nulla da provare, per cui possiamo supporre, a meno di scambiare l'ordine delle equazioni, che $F_1 \neq 0$: pertanto, tenendo presente quanto premesso, possiamo scrivere $F_1(x)$ nella forma

$$F_1(x) = \alpha_1 x_j + G_1(\bar{x}),$$

dove $\bar{x} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{n+1}) \in \mathbb{K}^n$, per un opportuno indice j, essendo $\alpha_1 \neq 0$. Facendo lo stesso per gli altri funzionali coinvolti nel sistema, quest'ultimo si può quindi riscrivere nella forma:

$$\begin{cases} \alpha_1 x_j + G_1(\bar{x}) = 0 \\ \alpha_2 x_j + G_2(\bar{x}) = 0 \\ \quad \vdots \quad \vdots \\ \alpha_m x_j + G_m(\bar{x}) = 0 \end{cases}$$

dove gli scalari α_i con $i > 1$ possono anche essere nulli e ogni G_i è un'applicazione lineare $G_i : \mathbb{K}^n \to \mathbb{K}$.

Ricavando x_j dalla prima equazione, e sostituendo nelle altre, il sistema risulta equivalente a:

$$\begin{cases} x_j = -\alpha_1^{-1} G_1(\bar{x}) \\ -\alpha_2 \alpha_1^{-1} G_1(\bar{x}) + G_2(\bar{x}) = 0 \\ \quad \vdots \qquad \vdots \\ -\alpha_m \alpha_1^{-1} G_1(\bar{x}) + G_m(\bar{x}) = 0 \end{cases}$$

Denotata per ogni $i > 1$ con H_i la funzione lineare $-\alpha_i \alpha_1^{-1} G_1 + G_i : \mathbb{K}^n \to \mathbb{K}$, il sistema si riscrive ancora:

$$\begin{cases} x_j = -\alpha_1^{-1} G_1(\bar{x}) \\ H_2(\bar{x}) = 0 \\ \quad \vdots \\ H_m(\bar{x}) = 0. \end{cases} \tag{12}$$

Ora, se $m = 1$ abbiamo di fatto la sola prima equazione, e certamente ammette soluzioni non banali (ottenute come nel caso $n = 2$ in corrispondenza di un vettore non nullo $\bar{x} \in \mathbb{K}^n$). Altrimenti, l'ipotesi induttiva è applicabile al sistema formato da tutte le equazioni tranne la prima, perchè esso è costituito da $m-1$ equazioni ed inoltre $m - 1 < n$. Dunque quest'ultimo sistema ammette una soluzione non banale $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n) \in \mathbb{K}^n$, ed allora il vettore $x_0 := (\bar{x}_1, \ldots, \bar{x}_{j-1}, -\alpha_1^{-1} G_1(\bar{x}), \bar{x}_j, \ldots, \bar{x}_n)$ di $\mathbb{K}^{n+1}$ è una soluzione non banale del sistema (12), ovvero del sistema (11). $\square$

Corollario 15.2. *(Teorema di classificazione) Siano n, m numeri interi positivi. 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 stante l'isomorfismo inverso $F^{-1} : \mathbb{K}^m \to \mathbb{K}^m$, è anche impossibile che $m > n$.

Osservazione 15.3. Nel seguito conveniamo di denotare con $\mathbb{K}^0$ lo spazio vettoriale banale $\{0_{\mathbb{K}}\}$ sul campo $\mathbb{K}$. Allora il risultato precedente si estende anche al caso di interi naturali qualsiasi $m, n \in \mathbb{N}$.

16. Spazi vettoriali finitamente generati

Definizione 16.1. Uno spazio vettoriale V sul campo $\mathbb{K}$ si dice **finitamente generato** oppure **di dimensione finita** se esiste una successione finita di vettori:

$$v_1, \dots, v_k, \quad k \geq 1,$$

tale che ogni vettore $v \in V$ si possa scrivere come combinazione lineare di tali vettori, ovvero nella forma

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

per opportuni scalari $\lambda_1, \dots, \lambda_k$.

In tal caso diremo che la sequenza $v_1, \dots, v_l$ costituisce un **sistema di generatori** di V. La stessa terminologia si applica anche all'insieme $\{v_1, \dots, v_k\}$.

Si osservi che *non* è richiesto nella definizione di sistema di generatori $v_1, \dots, v_k$ che la decomposizione di un generico vettore v come combinazione lineari dei v_i sia univocamente determinata.

Esempio 16.2. Ogni spazio $\mathbb{K}^n$ è finitamente generato: ciò è ovvio se $n = 0$; se $n > 0$, la base canonica $\{e_1, \dots, e_n\}$ introdotta nel §8 costituisce un sistema di generatori in forza della (5).

Esempio 16.3. Consideriamo l'insieme di funzioni:

$$V = \{f : \mathbb{R} \to \mathbb{R} | \, f \text{ è un polinomio di grado } \leq 3\}.$$

Certamente V non è vuoto, e poichè la somma $f + g$ di due polinomi di grado al massimo 3 è anch'esso un polinomio dello stesso tipo, e analogamente per λf, con λ costante, allora V è un sottospazio vettoriale di $\mathbb{R}^{\mathbb{R}}$. In particolare, V è un $\mathbb{R}$-spazio vettoriale.

Un sistema di generatori di tale spazio è dato dalla sequenza di polinomi:

$$v_0(x) = 1, v_1(x) = x, v_2(x) = x^2, \, v_3(x) = x^3,$$

per cui V è di dimensione finita. Infatti ogni polinonio $f \in V$, per definizione, è una funzione del tipo

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3,$$

dove gli $a_i \in \mathbb{R}$ sono costanti assegnate. Dunque

$$f = a_0 v_0 + a_1 v_1 + a_2 v_2 + a_3 v_3.$$

Non tutti gli spazi vettoriali sono finitamente generati. Ad esempio, consideriamo lo spazio vettoriale W costituito da *tutti* i polinomi di variabile reale:

$$W = \{f : \mathbb{R} \to \mathbb{R}|\, f \text{ è un polinomio.}\}$$

Si tratta ancora di un sottospazio vettoriale di $\mathbb{R}^{\mathbb{R}}$. Ammettiamo per assurdo che esista un sistema di generatori $f_1, \dots, f_m$ di W: allora ogni polinomio $f : \mathbb{R} \to \mathbb{R}$ sarebbe della forma

$$f = \lambda_1 f_1 + \cdots + \lambda_m f_m$$

per opportune costanti λ_i. Detto d il grado massimo dei polinomi $f_1, \dots, f_m$ seguirebbe che il grado di ogni polinomio f è al massimo d. Ciò è evidentemente impossibile, ad esempio il polinomio x^{d+1} ha grado $d + 1$.

17. Sottospazio generato da un numero finito di vettori

Definizione 17.1. 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 $v_1, \dots, v_r$.

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

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

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

$$\lambda u = \lambda(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda\lambda_1 v_1 + \cdots + \lambda\lambda_m v_m \tag{13}$$

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

Mediante questa nozione, possiamo riformulare il concetto di sistema di generatori: da una sequenza di vettori $v_1, \dots v_m$, allora abbiamo:

$$v_1, \dots, v_m \text{ è un sistema di generatori di } V \iff V = L(v_1, \dots, v_m).$$

Notiamo in particolare che, dato uno spazio vettoriale qualsiasi (anche non finitamente generato), il sottospazio $L(v_1, \ldots, v_m)$ è sempre, per costruzione, uno spazio vettoriale finitamente generato.

In corrispondenza di una successione finita $A = \{v_1, \ldots, v_m\}$ possiamo definire un'applicazione

$$\varphi_A : \mathbb{K}^m \to V$$

nel modo seguente:

$$\varphi_A(\lambda_1, \ldots, \lambda_m) := \lambda_1 v_1 + \cdots + \lambda_m v_m. \tag{14}$$

Allora abbiamo quest'altra utile caratterizzazione dei sistemi di generatori:

Proposizione 17.2. *Data una successione $A = \{v_1, \ldots, v_m\}$ di vettori di uno spazio vettoriale V, allora:*

$$v_1, \ldots, v_m \text{ è un sistema di generatori di } V \iff \varphi_A \text{ è surgettiva.}$$

DIMOSTRAZIONE: Infatti, dire che $v_1, \ldots, v_m$ è un sistema di generatori di V significa, per definizione, che ogni vettore $v \in V$ è una combinazione lineare di tali vettori, con opportuni coefficienti $\lambda_1, \ldots, \lambda_m$, ma ciò equivale a dire che esiste un vettore $(\lambda_1, \ldots, \lambda_m) \in \mathbb{K}^m$ tale che $v = \varphi_A(\lambda_1, \ldots, \lambda_m)$. $\square$

Risulta che φ_A è un'appicazione lineare: la verifica è di fatto la stessa che appena discusso per giustificare che $L(v_1, \ldots, v_m)$ è un sottospazio: infatti, dati i vettori $(\lambda_1, \ldots, \lambda_m)$ e $(\mu_1, \ldots, \mu_m)$ di $\mathbb{K}^m$, allora:

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

Il lettore può agevolmente completare la verifica della linearità di L_A tenendo conto della (13).

Notiamo anche che, per costruzione:

$$Im(\varphi_A) = L(v_1, \ldots, v_w).$$

18. VETTORI LINEARMENTE DIPENDENTI E LINEARMENTE INDIPENDENTI

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

Definizione 18.1. 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.

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.

Definizione 18.2. Si dice che i vettori $v_1, \ldots, v_m$ di V sono *linearmente dipendenti* se esiste una loro combinazione lineare nulla ma **non banale**.

Dunque, esplicitamente, $v_1, \ldots, v_m$ di V sono linearmente dipendenti se esistono m scalari **non tutti nulli** tali che:

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

Sussiste la seguente caratterizzazione:

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 $(-2\lambda_3, -\lambda_3, \lambda_3)$ 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:

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

Sussiste la seguente caratterizzazione:

Proposizione 18.4. *Siano assegnati i vettori $v_1, \ldots, v_m$ dello spazio vettoriale V. Le seguenti proprietà sono equivalenti:*

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

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

Dimostrazione: *a)* $\Rightarrow$ *b)*. Si assuma che $v_1, \dots, v_m$ 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 $j \in \{1, \dots, m\}$ 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 $v_1, \dots, v_m$, datosi che $-1 \neq 0$. $\square$

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

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

Dimostrazione: Per definizione, $v_1, \dots, v_m$ sono linearmente dipendenti se e solo se esiste una loro combinazione lineare nulla non banale, ovvero 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 18.6. 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 $\lambda_1 = \cdots = \lambda_m = 0$.

Ovviamente, sussiste la caratterizzazione, data una sequenza $A = \{v_1, \dots, v_m\}$:
$$v_1, \dots, v_m \text{ sono linearmente indipendenti} \iff \varphi_A \text{ è ingettiva.}$$

Osservazione 18.7. 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*, altrimenti si dirà *legato*.

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 $\{v_1, \dots, v_m\}$ si dice **base** di V se:

1) $v_1, \dots, v_m$ costituiscono un sistema di generatori di V.
2) $v_1, \dots, v_m$ 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
$$\lambda_1 e_1 + \cdots + \lambda_m = 0_{\mathbb{K}^m}$$
il vettore al primo membro si riscrive:
$$(\lambda_1, \dots, \lambda_m) = 0_{\mathbb{K}^m},$$
e quindi necessariamente tutti i coefficienti λ_i sono nulli.

In base a quanto stabilito nei paragrafi precedenti, posto $\mathfrak{B} = \{v_1, \dots, v_n\}$, abbiamo:
$$\mathfrak{B} \text{ è base di } V \iff \varphi_{\mathfrak{B}} \text{ è un isomorfismo.}$$

Un'applicazione molto importante di ciò e del Teorema di classificazione 15.2 è il risultato seguente, che stabilisce che due basi sono sempre necessariamente costituite dallo stesso numero di vettori:

Teorema 19.3. *Siano $\mathfrak{B} = \{v_1, \dots, v_n\}$ e $\mathfrak{B}' = \{w_1, \dots, w_m\}$ due basi di uno $\mathbb{K}$-spazio vettoriale V. Allora:*
$$n = m.$$

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

20. Esistenza di basi

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

Premettiamo qualche ulteriore osservazione in merito alle nozioni di indipendenza lineare e sottospazio generato da una sequenza di vettori.

Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e siano $v_1, \ldots, v_m$ vettori linearmente indipendenti. Allora per ogni i si ha $v_i \neq 0_V$.

Infatti, se fosse $v_i = 0$ per qualche i, si avrebbe

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_r = 0$$

ottenendo così una combinazione lineare nulla non banale dei vettori in questione. Un altro modo di giungere alla stessa conclusione è usare la Prop. 18.4 e il fatto che il vettore nullo si può sempre scrivere come combinazione lineare (ad es. banale) di una sequenza di vettori qualsiasi.

Inoltre i vettori in questione sono necessariamente tutti distinti. Infatti, se fosse $v_i = v_j$ per $i < j$, si otterrebbe:

$$0v_1 + \cdots + \cdots + 0v_{i-1} + 1v_i + \cdots + (-1)v_j + 0v_{j+1} + \cdots + 0v_m = 0.$$

Proposizione 20.1. *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. $\square$

Un'altra osservazione semplice ma utile è la seguente: siano $v_1, \ldots, v_m$ e $w_1, \ldots, w_k$ due sequenze di vettori. Allora, posto $A = \{v_1, \ldots, v_m\}$ e $B = \{w_1, \ldots, w_k\}$, si ha:

$$A \subset B \Rightarrow L(v_1, \ldots, v_m) \subset L(w_1, \ldots, v_k).$$

Infatti se $A \subset B$, allora ogni combinazione lineare dei vettori v_i è anche una combinazione lineare dei vettori w_j.

Ciò premesso, descriviamo un algoritmo mediante il quale è possibile estrarre da un sistema di generatori assegnato $v_1, \dots, v_m$ di uno spazio vettoriale V una base $\{v_{i_1}, \dots, v_{i_n}\}$ con $i_1 < \dots < i_n$. A tale scopo, introduciamo la seguente terminologia:

Definizione 20.2. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $v_1, \dots, v_m$ un sistema di generatori di V. Sia $i \in \{1, \dots, m\}$. Il vettore v_i si dirà **eliminabile** se esso è combinazione lineare dei vettori che lo precedono, ovvero se

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

Altrimenti v_i si dirà **non eliminabile**.

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

Notiamo che, se V è uno spazio vettoriale non banale, ogni sistema di generatori $v_1, \dots, v_m$ contiene qualche vettore non eliminabile. Infatti, se tali vettori fossero tutti eliminabili, avremmo che $v_1 = 0_V$, e quindi anche $v_2 = 0_V$, dovendo essere della forma $v_2 = \lambda v_1$, e così tutti i vettori v_i sarebbero nulli perchè combinazioni lineari di vettori tutti uguali a 0_V. Ne seguirebbe $V = \{0_V\}$ per definizione di sistema di generatori.

Teorema 20.3. *Sia V uno spazio vettoriale finitamente generato non banale. Dato un sistema di generatori $v_1, \dots, v_m$, siano*

$$v_{i_1}, \dots, v_{i_n}, \quad i_1 < \dots < i_n$$

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

Tale base si dice estratta dal sistema di generatori $v_1, \dots, v_m$ mediante l'algoritmo di eliminazione *o algoritmo degli* scarti successivi.

Dimostrazione: Proviamo innanzitutto che i vettori non eliminabili sono linearmente indipendenti; consideriamo una loro combinazione lineare nulla:

$$\lambda_1 v_{i_1} + \dots + \lambda_n v_{i_n} = 0. \tag{15}$$

Cominciamo col provare che $\lambda_n = 0$; se così non fosse, si potrebbe ricavare il vettore v_{i_n} dalla (16), scrivendolo come combinazione lineare degli altri:

$$v_{i_n} = -\lambda_n^{-1}\lambda_1 v_{i_1} - \dots - \lambda_n^{-1}\lambda_{n-1} v_{i_{n-1}},$$

ma ciò è impossibile, essendo v_{i_n} non eliminabile, in quanto i vettori coinvolti precedono tutti v_{i_n}. Dunque $\lambda_n = 0$ e la (16) si riscrive:

$$\lambda_1 v_{i_1} + \dots + \lambda_{n-1} v_{i_{n-1}} = 0. \tag{16}$$

A questo punto è sufficente ripetere lo stesso argomento per stabilire che $\lambda_{n-1} = 0$, sempre sfruttando la non eliminabilità di $v_{i_{n-1}}$; iterando questo ragionamento si perviene alla conclusione che tutti i coefficienti λ_i sono nulli.

Proviamo ora che i vettori di $\mathfrak{B}$ costituiscono un sistema di generatori. Si tratta di provare che

$$V = L(v_{i_1}, \dots, v_{i_n});$$

siccome la sequenza iniziale è un sistema di generatori di V, si tratta di stabilire in effetti che:

$$L(v_1, \ldots, v_m) = S,$$

dove per comodità abbiamo denotato con S il sottospazio $L(v_{i_1}, \ldots, v_{i_n})$. A questo scopo, basta provare che tutti i vettori v_i appartengono a S (essendo infatti S un sottospazio, ne consegue che ogni combinazione lineare dei v_i appartiene ancora a S). Ciò è ovvio per i vettori non eliminabili, bisogna provare ciò per i vettori eliminabili della sequenza. Siano quindi

$$v_{j_1}, \ldots, v_{j_s}, \quad j_1 < \cdots < j_s$$

i vettori eliminabili. Fissiamo l'attenzione sul primo di questi v_{j_1}: tutti i vettori che lo precedono sono necessariamente non eliminabili, ma poichè esso è eliminabile, è combinazione lineare di questi. Pertanto, essendo combinazione lineare di vettori di S, v_{j_1} appartiene a S. Consideriamo ora il secondo v_{j_2}; dei vettori che lo precedono solo v_{j_1} è eliminabile, tutti gli altri sono non eliminabili; avendo già stabilito che v_{j_1} sta in S, tutti questi vettori che precedono v_{j_2} sono quindi in S, e pertanto anche v_{j_2} è in S, sempre perchè combinazione lineare di vettori di S. Questo argomento si può iterare per tutti i vettori eliminabili della sequenza, che risultano tutti appartenere a S, concludendo così la dimostrazione. $\square$

21. Dimensione di uno spazio vettoriale

Teorema 21.1. *Sia V uno spazio vettoriale finitamente generato sul campo $\mathbb{K}$.*

Esiste un solo intero $n \in \mathbb{N}$ tale che $V \cong \mathbb{K}^n$. Se $n > 0$, tale numero è caratterizzato come il numero di vettori di una qualsiasi base di V.

Esso si chiama la **dimensione** *di V e si denota con*

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

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

Dimostrazione: Riguardo l'unicità di n, se esistesse un altro intero s per cui $V \cong \mathbb{K}^s$, avremmo di conseguenza $\mathbb{K}^s \cong \mathbb{K}^n$, il che comporta $s = n$ per il Corollario 15.2. Per quel che concerne l'esistenza, se V è banale, allora $V \cong \mathbb{K}^0$. Se V non è banale, il Teorema 20.3 garantisce che esiste almeno una base $\mathfrak{B} = \{v_1, \ldots, v_n\}$ ed in corrispondenza di tale base vi è l'isomorfismo $\varphi_{\mathfrak{B}} : \mathbb{K}^n \to V$ discusso nel §19. L'ultima affermazione è conseguenza diretta del Teorema 19.3 che garantisce a priori che due basi qualsiasi hanno lo stesso numero di vettori. $\square$

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

Ad esempio, abbiamo ovviamente $\mathtt{dim}_{\mathbb{K}}(\mathbb{K}^n) = n$, essendo $\mathbb{K}^n$ isomorfo a sè stesso; d'altra parte ciò è coerente con il fatto che la base canonica $\mathfrak{B} = \{e_1, \ldots, e_n\}$ consta esattamente di n vettori.

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

Teorema 21.2. *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 $w_1, \ldots, w_m$ 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 $w_1, \ldots, w_m$ è un sistema di generatori qualsiasi, allora $m \geq n$.*

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

$$\varphi_{\mathfrak{B}}^{-1} \circ \varphi_A : \mathbb{K}^m \to \mathbb{K}^n.$$

In base al Teorema 15.1 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 $w_1, \ldots, w_m$. Il Teorema 20.3 garantisce che da tale sistema si può estrarre una base $w_{i_1}, \ldots, w_{i_k}$, $k \leq m$. Poichè qualunque base è costituita da n vettori dev'essere $k = n$, e pertanto $m \geq n$. $\square$

Corollario 21.3. *Sia V uno spazio vettoriale finitamente generato di dimensione $n \geq 1$. Si consideri una sequenza di n vettori $v_1, \ldots, v_n$.*

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

Dimostrazione: (1) Per assurdo, si assuma che $v_1, \ldots, v_n$ siano indipendenti e che $V \neq L(v_1, \ldots, v_n)$. Allora esisterebbe almeno un vettore v_{n+1} di V tale che $v_{n+1} \notin L(v_1, \ldots, v_n)$. Per la Proposizione 20.1, seguirebbe che $v_1, \ldots, v_n, v_{n+1}$ sono linearmente indipendenti, contro la (1) del Teorema precedente.

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

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

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

Esempio 21.4. 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).

Concludiamo questo paragrafo notando che spazi vettoriali isomorfi hanno necessariamente la stessa dimensione. Piú precisamente, risulta che, se V e V' sono due $\mathbb{K}$-spazi vettoriali isomorfi, allora V è finitamente generato se e solo se lo è l'altro. In tal caso $\mathtt{dim}_{\mathbb{K}}(V) = \mathtt{dim}_K(V')$. Quest'ultima uguaglianza segue immediatamente dal fatto che, essendo $V \cong \mathbb{K}^n$, con $n = \mathtt{dim}_{\mathbb{K}}(V)$, allora si ha anche $V' \cong \mathbb{K}^n$, e questo fatto, per definizione di dimensione, dice che anche V' ha dimensione n.

Per giustificare la prima affermazione, proviamo più in generale che:

Proposizione 21.5. *Sia $F : V \to V'$ un epimorfismo tra due $\mathbb{K}$-spazi vettoriali. Allora*

$$V \text{ è finitamente generato} \Rightarrow V' \text{ è finitamente generato.}$$

Inoltre, se $v_1, \ldots, v_m$ è un sistema di generatori di V, allora $F(v_1), \ldots, F(v_m)$ costituscono un sistema di generatori di V'.

Dimostrazione: Cominciamo con l'osservare che, se $v_1, \ldots, v_m$ sono vettori di V e $F : V \to V'$ è un'applicazione lineare qualsiasi (non necessariamente surgettiva), allora:

$$F(L(v_1, \ldots, v_m)) = L(F(v_1), \ldots, F(v_m)).$$

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. 7.3. Ora, se V è finitamente generato e $v_1, \ldots, v_m$ costituiscono un sistema di generatori, allora l'uguaglianza precedente si legge:

$$F(V) = L(F(v_1), \ldots, F(v_m))$$

e, se F è un epimorfismo,

$$V' = L(F(v_1), \ldots, F(v_m))$$

in quanto la surgettività di F significa che $F(V) = V'$. $\square$

Riguardo le basi, abbiamo anche:

Proposizione 21.6. *Sia $F : V \to V'$ un isomorfismo tra spazi vettoriali con V finitamente generato, di dimensione n. Allora, se $\mathfrak{B} = \{v_1, \ldots, v_n\}$ è una base di V, anche*

$$\mathfrak{B}' = \{F(v_1), \ldots, F(v_n)\}$$

è una base di V'.

Dimostrazione: Sappiamo già da quanto precede che $\mathfrak{B}' = \{F(v_1), \ldots, F(v_n)\}$ è un sistema di generatori di V', resta da provare che tali vettori sono linearmente indipendenti. Consideriamo una combinazione lineare nulla

$$\lambda_1 F(v_1) + \cdots + \lambda_n F(v_n) = 0_{V'}.$$

Allora si ha

$$F(\lambda_1 v_1 + \cdots \lambda_n v_n) = 0_{V'},$$

e quindi il vettore $\lambda_1 v_1 + \cdots + \lambda_n v_n$ appartiene a $Ker(F)$; siccome F è ingettivo, tale nucleo è il sottospazio banale $\{0_V\}$, quindi

$$\lambda_1 v_1 + \cdots \lambda_n v_n = 0_V.$$

Stante la lineare indipendenza dei vettori in questione, consegue $\lambda_1 = \cdots = \lambda_n = 0$. $\square$

Riepilogando, se dato un $\mathbb{K}$-spazio vettoriale V si riesce a individuare un isomorfismo $F : \mathbb{K}^n \to V$ per un certo n allora abbiamo che V è finitamente generato, la sua dimensione è n, e una base di V è:

$$\mathfrak{B} = \{F(e_1), \ldots, F(e_n)\},$$

ottenuta calcolando le immagini dei vettori della base canonica di $\mathbb{K}^n$.

22. Calcolo della dimensione di alcuni spazi notevoli

Nel seguito calcoliamo la dimensione ed una base di alcuni spazi vettoriali interessanti.

Esempio 22.1. Lo spazio vettoriale $M_2(\mathbb{K})$ costituito dalle matrici quadrate di ordine 2 a coefficienti nel campo $\mathbb{K}$ ha dimensione 4.

Infatti, vi è un isomorfismo naturale

$$F : M_2(\mathbb{K}) \to \mathbb{K}^4$$

$$F \begin{pmatrix} a & b \\ c & d \end{pmatrix} := (a, b, c, d).$$

Alla stessa conclusione si perviene notando che una base di tale spazio è costituita dalle 4 matrici:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{$*$}$$

Infatti ogni matrice di $M_2(\mathbb{K})$ si può decomporre come segue:

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

Questa uguaglianza prova che le matrici in questione formano un sistema di generatori, ma implica immediatamente anche la loro lineare indipendenza: se la combinazione lineare al secondo membro è nulla, allora la matrice al primo membro dev'essere la matrice nulla $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ e quindi $a = b = c = d = 0$.

Si osservi che la stessa base si ricava dall'isomorfismo G inverso di F, $G : \mathbb{K}^4 \to M_2(\mathbb{K})$, che è dato da:

$$G(a, b, c, d) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Allora le matrici (*) sono esattamente le immagini $G(e_1), \ldots, G(e_4)$ dei vettori della base canonica di $\mathbb{K}^4$.

Più in generale, si può stabilire che $M_n(\mathbb{K})$ ha dimensione n^2, considerando l'isomorfismo definito da:

$$F(a_{ij}) := (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn})$$

dove le coordinate del vettore immagine di una matrice (a_{ij}) sono gli elementi della matrice stessa letti in ordine lessicografico, considerando prima la prima riga, poi la seconda riga e così via fino all'ultima riga.

In modo del tutto simile, il lettore può spiegarsi che $\mathtt{dim}(M_{m,n}(\mathbb{K})) = m \cdot n$.

Esempio 22.2. Fissato $n \geq 0$, l'insieme

$$\mathbb{R}_n[x] = \{p : \mathbb{R} \to \mathbb{R} | \, p \text{ è un polinomio di grado } \leq n\},$$

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

$$1, x, x^2, \ldots, x^n \tag{$*$}$$

con coefficienti esattamente $a_0, \ldots, a_n$; tali polinomi costituiscono quindi un sistema generatori di $\mathbb{R}_n[x]$. I polinomi (*) costituiscono di fatto una base; 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 22.3. 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 $\mathbb{K}^n$. Pertanto, se $G : \mathbb{K}^n \to \mathbb{K}$ è un altro vettore di $Hom(\mathbb{K}^n, \mathbb{K})$, con $G(x) = b_1x_1 + \cdots + b_nx_n$, allora:

$$F = G \iff F(e_i) = G(e_i) \text{ per ogni } i = 1, \ldots, n. \qquad (**)$$

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:

$$\texttt{dim}_{\mathbb{K}} Hom(\mathbb{K}^n, \mathbb{K}) = n.$$

Ad esempio, una base di $Hom(\mathbb{R}^3, \mathbb{R})$ è costituita dalle 3 proiezioni:

$$p_1(x_1, x_2, x_3) = x_1,\ p_2(x_1, x_2, x_3) = x_2,\ p_3(x_1, x_2, x_3) = x_3.$$

Notiamo che il fatto che tale spazio ha dimensione 3 si può giustificare definendo direttamente un isomorfismo

$$\Phi : Hom(\mathbb{R}^3, \mathbb{R}) \to \mathbb{R}^3$$

ponendo:

$$\Phi(F) := (F(e_1), F(e_2), F(e_3)).$$

In forza della (**), questa applicazione è certamente ingettiva; ma è anche ovviamente surgettiva perchè per ogni vettore $(a_1, a_2, a_3) \in \mathbb{R}^3$ vi è l'applicazione lineare

$F(x) = a_1x_1 + a_2x_2 + a_3x_3$, per la quale $\Phi(F) = (a_1, a_2, a_3)$. Inoltre, è agevole controllare che Φ è lineare: infatti date $F, G \in Hom(\mathbb{R}^3, \mathbb{R})$, allora

$$\begin{aligned}&\Phi(F+G) = ((F+G)(e_1), (F+G)(e_2), (F+G)(e_3)) = \\ &(F(e_1)+G(e_1), F(e_2)+G(e_2), F(e_3)+G(e_3)) = \\ &(F(e_1), F(e_2), F(e_3)) + (G(e_1), G(e_2), G(e_3)) = \\ &\Phi(F) + \Phi(G).\end{aligned}$$

In modo del tutto analogo il lettore verifichi che $\Phi(\lambda F) = \lambda\Phi(F)$ per ogni $\lambda \in \mathbb{R}$.

Questo approccio si può generalizzare al caso generale, ottenendo, per ogni campo $\mathbb{K}$, e per ogni $n \geq 1$, un isomorfismo

$$\Phi : Hom(\mathbb{K}^n, \mathbb{K}) \to \mathbb{K}^n,$$

dato da

$$\Phi(F) := (F(e_1), \ldots, F(e_n)).$$

Ciò fornisce un'altra giustificazione del fatto che $Hom(\mathbb{K}^n, \mathbb{K})$ ha dimensione n.

23. Sottospazi di spazi di dimensione finita

Teorema 23.1. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $n > 0$ e sia W un sottospazio di V. Allora*

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

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

Dimostrazione: 1) Ovviamente possiamo supporre $W \neq \{0\}$. Ragioniamo per assurdo, assumendo che W non sia finitamente generato. Proviamo che ciò ha come conseguenza la seguente proprietà:

$$\forall k \in \mathbb{N}^* \quad \exists v_1, \ldots, v_k \in W \text{ linearmente indipendenti.} \tag{$*$}$$

Naturalmente questo è in contraddizione con il fatto che V ha dimensione finita (Teorema 21.2). Proviamo la proprietà (*) per induzione su k. Poichè in W esiste almeno un vettore $v_1 \neq 0$, la (*) è vera per $k = 1$. Supposta la (*) vera per $k \geq 1$, esistono $v_1, \ldots, v_k$ vettori in W linearmente indipendenti. Ora, avendo assunto che W non sia finitamente generato, dev'essere $W \neq L(v_1, \ldots, v_k)$. Scelto un vettore $v_{k+1} \in W$, $v_{k+1} \notin L(v_1, \ldots, v_k)$, abbiamo che $v_1, \ldots, v_k, v_{k+1}$ sono indipendenti in forza della Proposizione 20.1. Pertanto la proprietà (*) è vera per $k + 1$.

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 $k \leq n$ per il Teorema 21.2. Se $k = n$, i vettori di $\mathfrak{B}'$ sono n e quindi formano necessariamente anche una base di V in virtù del Corollario 21.3; 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 $n > 0$ e sia $\mathfrak{B} = \{u_1, \dots, u_n\}$ una base di V. Siano assegnati dei vettori linearmente lindipendenti

$$v_1, \dots, v_k$$

. Allora esistono $n - k$ vettori $u_{j_1}, \dots, u_{j_{n-k}}$ della base $\mathfrak{B}$ tali che

$$\mathfrak{B}' = \{v_1, \dots, v_k, u_{j_1}, \dots, u_{j_{n-k}}\}$$

è una base di V.

Dimostrazione: Consideriamo la sequenza costituita da tutti i vettori

$$v_1, \dots, v_k, u_1, \dots, u_n.$$

Tale sequenza è ancora un sistema di generatori, da cui si può estrarre una base $\mathfrak{B}'$ mediante l'algoritmo di eliminazione. Poichè $v_1, \dots, v_k$ sono indipendenti, ciascuno di essi non è eliminabile. Dunque la base $\mathfrak{B}'$ è del tipo richiesto. □

25. Teoremi principali per il calcolo di dimensioni

Il risultato seguente è di importanza basilare in tutte le applicazioni dell'algebra lineare.

Teorema 25.1. *(Teorema del rango o della dimensione)*
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 anche $Im(f)$ è finitamente generato ed inoltre

$$\mathtt{dim}(V) = \mathtt{dim}(Ker(f)) + \mathtt{dim}(Im(f)). \tag{17}$$

Dimostrazione: Osserviamo in prima istanza che l'applicazione ridotta di f,

$$f : V \to Im(f)$$

è ancora lineare ed è surgettiva per costruzione; dunque è un epimorfismo. Poichè V è finitamente generato, tale dev'essere $Im(f)$, cfr. la Prop. 21.5.

Proviamo ora la (17). Esaminiamo dapprima il caso in cui f è ingettiva. In tal caso la funzione ridotta $f : V \to Im(f)$ è anche ingettiva, e quindi è un isomorfismo. Pertanto

$$\mathtt{dim}(V) = \mathtt{dim}(Im(f));$$

d'altra parte per l'ingettività, $Ker(f) = \{0_V\}$ e quindi $\mathtt{dim}(Ker(f)) = 0$ e pertanto la tesi (17) 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 $\mathfrak{B}_0 = \{u_1, \ldots, u_k\}$ 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 $\mathfrak{B}_0$ ad una base

$$\mathfrak{B} = \{u_1, \ldots, u_n\}$$

di V. Consideriamo quindi il sottospazio

$$W = L(u_{k+1}, \ldots, u_n).$$

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. Consideriamo la restrizione a W della stessa funzione ristretta $f : V \to Im(f)$:

$$g = f|_W : W \to Im(f),$$

che è ancora un'applicazione lineare (Prop. 14.1). Proveremo che g è un isomorfismo. Da ciò segue che $\mathtt{dim}(Im(f)) = \mathtt{dim}(W) = n-k$ e quindi la (17).

Riguardo la surgettività, basta osservare che $f(W) = Im(f)$; infatti

$$Im(f) = f(V) = f(L(u_1, \ldots, u_n)) = L(f(u_1), \ldots, f(u_k), f(u_{k+1}), \ldots, f(u_n)) =$$

$$= L(0, \ldots, 0, f(u_{k+1}), \ldots, f(u_n)) = L(f(u_{k+1}), \ldots, f(u_n)) = f(W).$$

In questi passaggi abbiamo usato la linearità di f ed il fatto che $f(u_1) = \cdots = f(u_k) = 0$ perchè tali vettori appartengono al nucleo di f.

Riguardo l'ingettività di g, proviamo che il nucleo di g è banale. Sia $w \in W$ tale che $g(w) = 0$; allora $f(w) = 0$, e quindi tale vettore appartiene anche a $Ker(f)$. Ora, siccome i vettori $u_1, \ldots, u_n$ sono linearmente indipendenti, ciò comporta subito che $w = 0$. Infatti, abbiamo

$$\lambda_1 u_1 + \cdots + \lambda_k u_k = w = \alpha_1 u_{k+1} + \cdots + \alpha_{n-k} u_n$$

per opportuni scalari λ_i e α_j, da cui

$$\lambda_1 u_1 + \cdots + \lambda_k u_k - \alpha_1 u_{k+1} - \cdots - \alpha_{n-k} u_n = 0,$$

ma ciò implica che tutti i gli scalari in questione devono essere nulli, ovvero $w = 0$. $\square$

Definizione 25.2. Sia $f : V \to V'$ un'applicazione lineare tra $\mathbb{K}$-spazi vettoriali, con V finitamente generato. La dimensione di $Im(f)$ prende il nome di **rango** di f e si denoterà con $rg(f)$.

26. Il teorema di Grassmann

In questa sezione discutiamo la dimensione di prodotti diretti di spazi vettoriali e di somme di sottospazi. Cominciamo trattando il primo caso.

Teorema 26.1. *Siano V_1 e V_2 due $\mathbb{K}$-spazi vettoriali di dimensione finita. Allora*

$$\mathtt{dim}(V_1 \times V_2) = \mathtt{dim}(V_1) + \mathtt{dim}(V_2).$$

Dimostrazione: Posto $n = \mathtt{dim}(V_1)$ e $m = \mathtt{dim}(V_2)$, siano assegnati due isomorfismi $f_1 : V_1 \to \mathbb{K}^n$ e $f_2 : V_2 \to \mathbb{K}^m$. A partire da questi, possiamo definire in modo standard un'applicazione:

$$f : V_1 \times V_2 \to \mathbb{K}^n \times \mathbb{K}^m,$$

ponendo

$$f(v_1, v_2) := (f_1(v_1), f_2(v_2)) \quad \forall (v_1, v_2) \in V_1 \times V_2.$$

È un semplice esercizio (di carattere generale) verificare che si tratta di una bigezione, stante la bigettività di entrambe le funzioni coinvolte. Tale applicazione è lineare; infatti, dati due vettori (v_1, v_2) e (w_1, w_2) di $V_1 \times V_2$, allora:

$$\begin{aligned}
&f((v_1, v_2) + (w_1, w_2)) = f((v_1 + w_1, v_2 + w_2)) = (f_1(v_1 + w_1), f_2(v_2 + w_2)) = \\
&= (f_1(v_1) + f_1(w_1), f_2(v_2) + f_2(w_2)) = (f_1(v_1), f_2(v_2)) + (f_1(w_1), f_2(w_2)) = \\
&= f(v_1, v_2) + f(w_1, w_2).
\end{aligned}$$

Analogamente si verifica che $f(\lambda(v_1, v_2)) = \lambda f((v_1, v_2))$ per ogni scalare λ. Dunque abbiamo stabilito che:

$$V_1 \times V_2 \cong \mathbb{K}^n \times \mathbb{K}^m.$$

Quindi per completare la dimostrazione basta verificare che:

$$\mathbb{K}^n \times \mathbb{K}^m \cong \mathbb{K}^{n+m}.$$

Ciò è banale se $n = 0$ oppure $m = 0$; ad esempio se $n = 0$, l'applicazione

$$(0, x) \mapsto x$$

realizza un isomorfismo $\mathbb{K}^0 \times \mathbb{K}^m \to \mathbb{K}^m$. Nel caso in cui n, m siano entrambi positivi, 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 verifica che si tratta di un isomorfismo. □

Discutiamo ora il caso della somma di due sottospazi, provando un altro risultato fondamentale:

Teorema 26.2. *(Formula di Grassmann) Sia V uno spazio vettoriale di dimensione finita. Siano U e W due sottospazi di V. Allora:*

$$\mathtt{dim}(U + W) = \mathtt{dim}(U) + \mathtt{dim}(W) - \mathtt{dim}(U \cap W). \tag{18}$$

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 appartere sia a U che a W; se ne deduce che:

$$Ker(F) = \{(u, -u) |\, u \in U \cap W\}. \tag{19}$$

Da qui segue facilmente che:

$$Ker(F) \cong U \cap W.$$

Infatti, un isomorfismo tra questi spazi è dato semplicemente dall'applicazione

$$u \in U \cap W \mapsto (u, -u) \in Ker(F).$$

Applicando quindi a F il Teorema del rango, otteniamo:

$$\mathtt{dim}(U+W) = \mathtt{dim}(Im(F)) = \mathtt{dim}(U\times W)-\mathtt{dim}(Ker(F)) = \mathtt{dim}(U)+\mathtt{dim}(W)-\mathtt{dim}(U\cap W).$$

□

Corollario 26.3. *Sia V uno spazio vettoriale di dimensione finita e siano U, W sottospazi di V. Allora se la somma $U + W$ è diretta, si ha*

$$\mathtt{dim}(U + W) = \mathtt{dim}(U) + \mathtt{dim}(W).$$

27. Componenti di un vettore rispetto ad una base

Consideriamo un $\mathbb{K}$-spazio vettoriale V. Notiamo che sussiste la seguente caratterizzazione delle basi di V:

Proposizione 27.1. *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 $v_1, \ldots, v_n$.

Ciò è conseguenza immediata del fatto che la condizione a) è caratterizzata mediante l'applicazione lineare

$$\varphi_{\mathfrak{B}} : \mathbb{K}^n \to V,$$

studiata nel §19 e definita da

$$\varphi_{\mathfrak{B}}(x_1, \ldots, x_n) = x_1 v_1 + \cdots + x_n v_n = \sum_{i=1}^{n} x_i v_i,$$

mediante la condizione che questa sia un isomorfismo. Ma la bigettività di $\varphi_{\mathfrak{B}}$ si traduce nel fatto che:

$$\forall v \in V \ \exists | (x_1, \ldots, x_n) \in \mathbb{K}^n : \quad v = \sum_{i=1}^{n} x_i v_i,$$

che è la condizione b) enunciata sopra.

Ha senso allora dare la seguente:

Definizione 27.2. Si chiameranno **componenti** di un vettore $v \in V$ **rispetto alla base** $\mathfrak{B}$ gli scalari $x_1, \ldots, x_n$ tali che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

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 di $\mathfrak{B}$ è essenziale nella definizione delle componenti.

Nel seguito, data una base $\mathfrak{B}$, denoteremo con

$$\psi_{\mathfrak{B}} : V \to \mathbb{K}^n$$

l'inverso dell'isomorfismo $\varphi_{\mathfrak{B}}$. Tale funzione associa quindi ad ogni vettore v la n-pla $(x_1, \ldots, x_n)$ delle sue componenti rispetto alla base in questione.

Esempio 27.3. Nel caso in cui $V = \mathbb{K}^n$ e scegliamo la base canonica $\mathfrak{B}_0 = \{e_1, \ldots, e_n\}$, allora le componenti di ogni $x = (x_1, \ldots, x_n)$ sono esattamente $(x_1, \ldots, x_n)$. Ciò segue dalla identità fondamentale (5).

Dunque in tal caso $\psi_{\mathfrak{B}_0} = Id_{\mathbb{K}^n}$ è semplicemente l'isomorfismo identico.

28. Matrice associata ad un insieme di vettori

Sia assegnato uno spazio vettoriale V di dimensione $n > 0$ e si fissi una sua base $\mathfrak{B} = \{v_1, \ldots, v_n\}$. Nel seguito, se $(x_1, \ldots, x_n) = \psi_{\mathfrak{B}}(v)$ è la n-pla delle componenti di un vettore $v \in V$, essa verrà identificata, laddove è conveniente, con il vettore colonna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e si chiamerà il *vettore delle componenti di v* rispetto alla base assegnata.

Definizione 28.1. Sia $u_1, \ldots, u_k$ 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

$$(\psi_{\mathfrak{B}}(u_1) \quad \psi_{\mathfrak{B}}(u_2) \quad \cdots \quad \psi_{\mathfrak{B}}(u_k)). \tag{20}$$

Lemma 28.2. *Sia $f : V \to V'$ un isomorfismo tra $\mathbb{K}$-spazi vettoriali entrambi di dimensione finita. Allora per ogni sottospazio W di V si ha*

$$\mathtt{dim}(W) = \mathtt{dim}(f(W)).$$

Dimostrazione: Basta osservare che la ridotta della restrizione

$$f_{|W} : W \to f(W)$$

della funzione f a W è un isomorfismo tra W e $f(W)$. □

L'utilità della nozione di matrice associata ad un insieme di vettori è stabilita dal risultato seguente.

Teorema 28.3. *Sia $u_1, \ldots, u_k$ una successione finita di vettori di V, $k \geq 1$. Sia $\mathfrak{B}$ una base di V e sia A la matrice associata a tali vettori rispetto alla base $\mathfrak{B}$. Allora*

$$\mathtt{dim}(L(u_1, \ldots, u_k)) = rg(A).$$

Dimostrazione: Poichè $\psi_{\mathfrak{B}}$ è un isomorfismo, in forza del Lemma precedente abbiamo

$$\mathtt{dim}(L(u_1, \ldots, u_k)) = \mathtt{dim}L(\psi_{\mathfrak{B}}(u_1), \ldots, \psi_{\mathfrak{B}}(u_k)) = \mathtt{dim}L(A_{(1)}, \ldots, A_{(k)}) = rg(A),$$

dove $A_{(1)}, \ldots, A_{(k)}$ sono le colonne di A. □

Corollario 28.4. *Siano V uno 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*

$$\mathfrak{B}' \text{ è una base di } V \iff \det(A) \neq 0.$$

Dimostrazione: Infatti, sappiamo che la sequenza di n vettori in questione forma una base se e solo se sono è un sistema di generatori, il che accade se e solo se il sottospazio $L(u_1, \ldots, v_n)$ ha dimensione n, il che è vero se e solo se $rg(A) = n$ per il teorema precedente. □

29. Esistenza ed unicità di applicazioni lineari

Come applicazione della nozione di componenti di un vettore, 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.

Teorema 29.1. *Siano V e W due $\mathbb{K}$-spazi vettoriali, con V finitamente generato. Sia*

$$v_1, \ldots, v_k$$

un sistema di generatori 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: a) Siano $f, g : V \to W$ lineari e si assuma che $f(v_i) = g(v_i)$ per ogni $i \in \{1, \ldots, k\}$. Sia $v \in V$; allora possiamo scrivere tale vettore nella forma $v = \sum_{i=1}^{n} x_i v_i$, per opportuni scalari $x_1, \ldots, x_k$, da cui per la linearità di entrambe le applicazioni:

$$f(v) = f(x_1v_1+\cdots+x_nv_n) = x_1f(v_1)+\cdots+x_nf(v_n) = x_1g(v_1)+\cdots+x_ng(v_n) = g(v)$$

e stante l'arbitrarietà di v abbiamo che $f = g$. □

Il risultato seguente mostra che si possono costruire applicazioni lineari prescrivendone i valori assunti sui vettori di una base:

Teorema 29.2. *(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 $w_1, \ldots, w_n$, definiamo un'applicazione $f : V \to W$ nel modo seguente:

$$f(v) := x_1w_1+\cdots+x_nw_n, \qquad \text{dove } (x_1, \ldots, x_n) \text{ sono le componenti di } v \text{ rispetto a } \mathfrak{B}.$$

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 $i \in \{1, \ldots, n\}$, 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. Verifichiamo questo in due modi diversi.

Primo modo: posto $A = \{w_1, \ldots, w_n\}$, abbiamo l'applicazione lineare

$$\varphi_A : \mathbb{K}^n \to W, \quad \varphi_A(x_1, \ldots, x_n) = x_1 w_1 + \cdots + x_n w_n.$$

Allora basta osservare che f coincide con la funzione composta

$$f = \varphi_A \circ \psi_{\mathfrak{B}}.$$

Dunque f è lineare perchè composta di applicazioni lineari.

Secondo modo: dati due vettori v, v' di V di componenti rispettivamente $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ allora $v + v'$ ha componenti $(x_1 + y_1, \ldots, x_n + y_n)$ (ciò si può verificare direttamente, ma è conseguenza diretta della linearità di $\psi_{\mathfrak{B}} : V \to \mathbb{K}^n$). Quindi per definizione di f:

$$f(v + w) = (x_1 + y_1)w_1 + \cdots + (x_n + y_n)w_n =$$
$$= (x_1 w_1 + \cdots + x_n w_n) + (y_1 w_1 + \cdots + y_n w_n) = f(v) + f(w).$$

In modo analogo si verifica che $f(\lambda v) = \lambda f(v)$ per ogni $\lambda \in \mathbb{K}$ e $v \in V$. $\square$

30. Relazione tra matrici ed applicazioni lineari

Siano V e W due spazi vettoriali non banali sullo stesso campo $\mathbb{K}$, entrambi finitamente generati, aventi dimensioni n ed m rispettivamente.

Definizione 30.1. Siano fissate due basi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C} = \{w_1, \ldots, w_m\}$ di V e di W rispettivamente. Sia $f : V \to W$ un'applicazione lineare. Si chiamerà **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 $\mathfrak{C}$. Tale matrice si denoterà con uno dei simboli

$$M_{\mathfrak{B}}^{\mathfrak{C}}(f), \ M_{\mathfrak{B},\mathfrak{C}}(f), \ M(f)(\mathfrak{B}, \mathfrak{C}).$$

Dunque, per definizione (cf. (20)):

$$M_{\mathfrak{B},\mathfrak{C}}(f) = (\psi_{\mathfrak{C}}(f(v_1)) \quad \psi_{\mathfrak{C}}(f(v_2)) \quad \cdots \quad \psi_{\mathfrak{C}}(f(v_n))). \tag{23}$$

essendo al solito $\psi_{\mathfrak{C}} : W \to \mathbb{K}^m$ l'isomorfismo che associa ad ogni vettore di W il vettore delle sue componenti rispetto alla base $\mathfrak{C}$. Chiaramente, $M_{\mathfrak{B},\mathfrak{C}}(f)$ è una matrice di tipo $m \times n$.

Esempio 30.2. Data un'applicazione lineare $f : \mathbb{K}^n \to \mathbb{K}^m$, la matrice ad essa associata rispetto alle basi canoniche $\mathfrak{B}_0$ e $\mathfrak{C}_0$ è

$$A = (f(e_1) \ \cdots \ f(e_n)),$$

ovvero le colonne di A sono semplicemente le immagini $f(e_i)$ dei vettori della base canonica di $\mathbb{K}^n$.

Ciò segue direttamente dalla caratterizzazione delle componenti rispetto alla base canonica, cfr. Es. 27.3. Ad esempio, se $f : \mathbb{K}^4 \to \mathbb{K}^3$ è definita da:

$$f(x_1, x_2, x_3, x_4) := (x_1 + x_2 + x_4, 2x_1 - 2x_4, 3x_1 + x_2 - x_4),$$

allora la matrice ad essa associata, rispetto alle basi canoniche, è:

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

Il risultato che segue giustifica il termine *rango* che abbiamo attribuito alla dimensione dell'immagine di un'applicazione lineare:

Teorema 30.3. *Siano fissate due basi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C} = \{w_1, \ldots, w_m\}$ di V e di W rispettivamente. Sia $f : V \to W$ un'applicazione lineare. Allora*

$$rg(M_{\mathfrak{B},\mathfrak{C}}(f)) = rg(f).$$

Dimostrazione: Si tratta di applicare il Teorema 28.3:

$$rg(f) = \mathtt{dim}(Im(f)) = \mathtt{dim}L(f(v_1), \ldots, f(v_n)) = rg(M_{\mathfrak{B},\mathfrak{C}}(f)).$$

□

Vale 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 spazi vettoriali assegnati è esso stesso uno spazio vettoriale (§13).

Teorema 30.4. *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*

$$M_{\mathfrak{B}}^{\mathfrak{C}} : Hom(V, W) \to M_{m,n}(\mathbb{K})$$

che fa corrispondere ad ogni $f \in Hom(V, W)$ la matrice associata $M_{\mathfrak{B}}^{\mathfrak{C}}(f)$ ad f, è un isomorfismo di spazi vettoriali.

Dimostrazione: Assumiamo che la base $\mathfrak{B}$ sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$. Per prima cosa stabiliamo che l'applicazione $M_{\mathfrak{B}}^{\mathfrak{C}}$ è lineare. Siano $f, g \in Hom(V, W)$. Allora

$$M_{\mathfrak{B}}^{\mathfrak{C}}(f+g) = (\, \psi_{\mathfrak{C}}((f+g)(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}((f+g)(v_n)) \,) =$$
$$(\, \psi_{\mathfrak{C}}(f(v_1) + g(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(f(v_n) + g(v_n)) \,) =$$
$$(\, \psi_{\mathfrak{C}}(f(v_1)) + \psi_{\mathfrak{C}}(g(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(f(v_n)\,) + \psi_{\mathfrak{C}}(g(v_n)) \,) = M_{\mathfrak{B}}^{\mathfrak{C}}(f) + M_{\mathfrak{B}}^{\mathfrak{C}}(g)$$

dove abbiamo tenuto conto della definizione di $f + g$ e sfruttato la linearità di $\psi_{\mathfrak{C}}$. Analogamente:

$$M_{\mathfrak{B}}^{\mathfrak{C}}(\lambda f) = (\, \psi_{\mathfrak{C}}((\lambda f)(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}((\lambda f)(v_n)) \,) = (\, \psi_{\mathfrak{C}}(\lambda f(v_1)) \quad \cdots \quad \psi_{\mathfrak{C}}(\lambda f(v_n)) \,) =$$
$$(\, \lambda \psi_{\mathfrak{C}}(f(v_1)) \quad \cdots \quad \lambda \psi_{\mathfrak{C}}(f(v_n)) \,) = \lambda M_{\mathfrak{B}}^{\mathfrak{C}}(f).$$

Proviamo ora che $M^{\mathfrak{C}}_{\mathfrak{B}}$ è ingettiva; sia $f \in Ker(M^{\mathfrak{C}}_{\mathfrak{B}})$; ciò significa che la matrice $M^{\mathfrak{C}}_{\mathfrak{B}}(f)$ è la matrice nulla e quindi tutte le sue colonne sono nulle; dunque in particolare $\psi_{\mathfrak{C}} f(v_i) = 0$ per ogni i. Ciò significa che $f(v_i) = 0$ per ogni $i = 1, \dots, n$. Ciò è sufficiente per concludere che $f = 0$, ovvero che f coincide con la funzione nulla $0 : V \to W$, in quanto i valori assunti dalle due funzioni lineari in questione sui vettori della base $\mathfrak{B}$ coincidono (Teorema 29.1).

Resta da stabilire la surgettività di M; sia $A \in M_{m,n}(\mathbb{K})$ e consideriamo i seguenti vettori di W:

$$u_i := \psi_{\mathfrak{C}}^{-1}(A_{(i)}), \quad i = 1, \dots, n.$$

Applicando il Teorema 29.2, abbiamo che esiste una ed una sola applicazione lineare $f : V \to W$ tale che

$$f(v_i) = u_i, \quad i = 1, \dots, n.$$

Risulta allora per costruzione che $M^{\mathfrak{C}}_{\mathfrak{B}}(f) = A$, infatti:

$$M^{\mathfrak{C}}_{\mathfrak{B}}(f) = (\, \psi_{\mathfrak{C}}(f(v_1)) \ \cdots \ \psi_{\mathfrak{C}}(f(v_n)) \,) = (\, \psi_{\mathfrak{C}}(u_1) \ \cdots \ \psi_{\mathfrak{C}}(u_n) \,) = (\, A_{(1)} \ \cdots \ A_{(n)} \,) = A.$$

□

Come applicazione, abbiamo:

Corollario 30.5. *Se V e W sono spazi vettoriali non bnali di dimensione n ed m, allora anche $Hom(V, W)$ è finitamente generato e*

$$\mathtt{dim}(Hom(V, W)) = mn.$$

Dimostrazione: Si tratta di osservare che la dimensione dello spazio di matrici $M_{m,n}(\mathbb{K})$ è proprio mn. Ciò può giustificarsi ad esempio notando che esso è isomorfo al prodotto diretto

$$\mathbb{K}^m \times \cdots \times \mathbb{K}^m$$

di n copie di $\mathbb{K}^m$, mediante la funzione

$$A \mapsto (A_{(1)}, \dots, A_{(n)})$$

che associa ad ogni matrice di tipo $m \times n$ la n-pla delle sue colonne, ciascuna delle quali è considerata come vettore di $\mathbb{K}^m$. Alternativamente, si può costruire direttamente un isomorfismo tra $M_{m,n}(\mathbb{K})$ e $\mathbb{K}^{mn}$ trasformando una matrice $A = (a^i_j)$ in un vettore

$$(a^1_1, \cdots, a^1_n, \cdots, a^m_1, \cdots, a^m_n)$$

che si ottiene scrivendo di seguito tutti gli mn ingressi della matrice, in ordine lessicografico, dalla prima all'ultima riga. □

Mettiamo ora in evidenza un'utile caratterizzazione della matrice associata ad un'applicazione lineare. Ricordiamo che ogni matrice $A \in M_{m,n}(\mathbb{K})$ determina in modo canonico un'applicazione lineare

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

definita mediante il prodotto righe per colonne:

$$\forall x \in \mathbb{K}^n \quad L_A(x) := Ax. \tag{24}$$

Teorema 30.6. *Siano fissate due basi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C} = \{w_1, \ldots, w_m\}$ degli spazi vettoriali V e di W. Sia $f : V \to W$ un'applicazione lineare. La matrice $A := M_{\mathfrak{B}}^{\mathfrak{C}}(f)$ è l'unica matrice A di tipo $m \times n$ tale che il seguente diagramma di applicazioni risulti commutativo:*

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \psi_{\mathfrak{B}} \downarrow & & \downarrow \psi_{\mathfrak{C}} \\ \mathbb{K}^n & \xrightarrow{L_A} & \mathbb{K}^m \end{array}$$

cioè tale che:

$$L_A \circ \psi_{\mathfrak{B}} = \psi_{\mathfrak{C}} \circ f. \tag{25}$$

Dimostrazione: Posto $A = M_{\mathfrak{B}\mathfrak{C}}(f)$, verifichiamo la (25). Poichè ambo i membri di tale uguaglianza sono applicazioni lineari $V \to \mathbb{K}^m$, è sufficiente provare che esse assumono gli stessi valori sui vettori della base $\mathfrak{B}$ (usiamo ancora il Teorema 29.1). Ricordando che per ogni vettore v_i della base, le sue componenti, rispetto alla base stessa, sono $(0, \ldots, 0, 1, 0, \ldots, 0)$, cioè che:

$$\psi_{\mathfrak{B}}(v_i) = e_i, \tag{26}$$

dove $\{e_i\}$ è la base canonica di $\mathbb{K}^n$ (cfr. (22)), abbiamo:

$$(L_A \circ \psi_{\mathfrak{B}})(v_i) = L_A(e_i) = Ae_i = A_{(i)} = \psi_{\mathfrak{C}}(f(v_i)) = (\psi_{\mathfrak{C}} \circ f)(v_i)$$

e la (25) è provata.

Riguardo infine l'unicità di A, supponiamo che esista un'altra matrice B che soddisfi la stessa proprietà, cioè tale che:

$$L_B \circ \psi_{\mathfrak{B}} = \psi_{\mathfrak{C}} \circ f.$$

Allora

$$L_A \circ \psi_{\mathfrak{B}} = L_B \circ \psi_{\mathfrak{B}}$$

da cui, essendo $\psi_{\mathfrak{B}}$ una bigezione, segue immediatamente

$$L_A = L_B$$

e ciò garantisce senz'altro che

$$A = B,$$

in quanto si ha:

$$A_{(i)} = L_A(e_i) = L_B(e_i) = B_{(i)}.$$

□

Il risultato appena provato ha una semplice interpretazione esplicita in termini di componenti, come regola che fornisce il legame tra le componenti di un vettore $v \in V$ e le componenti della sua immagine $f(v)$:

Corollario 30.7. *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 la matrice associata ad f rispetto a $\mathfrak{B}$ e $\mathfrak{C}$.

Per ogni vettore $v \in V$, se il vettore delle componenti di v rispetto a $\mathfrak{B}$ è $x = (x_1, \ldots, x_n)$, allora il vettore y delle componenti di $f(v)$ rispetto a $\mathfrak{C}$ è dato da

$$y = Ax. \tag{27}$$

Dimostrazione: La dimostrazione è una semplice riscrittura del risultato precedente: dalla relazione

$$L_A \circ \psi_{\mathfrak{B}} = \psi_{\mathfrak{C}} \circ f$$

ricaviamo infatti

$$y = \psi_{\mathfrak{C}}(f(v)) = (L_A \circ \psi_{\mathfrak{B}})(v) = L_A(x) = Ax.$$

□

Spesso questo fatto si esprime dicendo che l'applicazione lineare f ha equazione

$$f : y = Ax$$

rispetto alla due basi fissate $\mathfrak{B}$ e $\mathfrak{C}$.

Esempio 30.8. Nel caso di un'aplicazione lineare $f : \mathbb{K}^n \to \mathbb{K}^m$, se A denota la matrice associata rispetto alle basi *canoniche* $\mathfrak{B}_0$ e $\mathfrak{C}_0$ degli spazi di partenza e di arrivo, allora l'equazione

$$y = Ax$$

esprime direttamente il legame tra $x \in \mathbb{K}^n$ ed il vettore $y = f(x)$. Dunque in tal caso:

$$f = L_A.$$

Ciò peraltro segue dalla (25) perchè $\psi_{\mathfrak{B}_0} = Id_{\mathbb{K}^n}$ e analogamente $\psi_{\mathfrak{C}_0} = Id_{\mathbb{K}^m}$.

31. Caratterizzazioni degli isomorfismi

In questo paragrafo approfondiamo il caso di un'applicazione lineare tra spazi vettoriali aventi la stessa dimensione.

Teorema 31.1. *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 ingettiva*
b) *f è surgettiva*
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 (17) risulta:

$$f \text{ ingettiva} \iff \mathtt{dim}(Ker(f)) = 0 \iff n = \mathtt{dim}(Im(f)) \iff W = Im(f),$$

dove l'ultima equivalenza è giustificata dal Teorema 23.1, applicato al sottospazio $Im(f)$ di W. $\square$

Come applicazione di questo risultato abbiamo la seguente caratterizzazione degli isomorfismi, come quelle applicazioni lineari che trasformano basi in basi; più precisamente:

Teorema 31.2. *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: L'implicazione $a) \Rightarrow b)$ è già nota (Prop. 21.6), mentre $b) \Rightarrow c)$ è ovvia. Resta da provare che $c) \Rightarrow a)$. Sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base di V e assumiamo che 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 provare che f è surgettiva. Infatti abbiamo:

$$Im(f) = L(f(v_1), \ldots, f(v_n)) = W.$$

$\square$

Concludiamo questo paragrafo con un'altra caratterizzazione, operativamente conveniente, in termini di matrici:

Teorema 31.3. *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 31.1 e 30.3:

$$f \text{ isomorfismo} \iff rg(f) = n \iff rg(M^{\mathfrak{C}}_{\mathfrak{B}}(f)) = n \iff M^{\mathfrak{C}}_{\mathfrak{B}}(f) \in GL(n, \mathbb{K}).$$

$\square$

32. Relazione tra composizione di applicazioni lineari e prodotto di matrici

Teorema 32.1. *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:*

$$M_{\mathfrak{B}}^{\mathfrak{D}}(g \circ f) = M_{\mathfrak{C}}^{\mathfrak{D}}(g) M_{\mathfrak{B}}^{\mathfrak{C}}(f). \tag{28}$$

Dimostrazione: Sfruttiamo la caratterizzazione della matrice associata a $g \circ f$ fornita dal Teorema 30.6. Denotate rispettivamente con n, m e s le dimensioni di V, W e U, e posto:

$$A := M_{\mathfrak{B}}^{\mathfrak{C}}(f),\ B = M_{\mathfrak{C}}^{\mathfrak{D}}(g),\ C = BA$$

dobbiamo provare che la matrice C rende commutativo il diagramma:

$$\begin{array}{ccc} V & \xrightarrow{g\circ f} & U \\ \psi_{\mathfrak{B}} \downarrow & & \downarrow \psi_{\mathfrak{D}} \\ \mathbb{K}^n & \xrightarrow{L_C} & \mathbb{K}^s \end{array}$$

cioè che:

$$L_C \circ \psi_{\mathfrak{B}} = \psi_{\mathfrak{D}} \circ (g \circ f).$$

Abbiamo a disposizione le analoghe relazioni soddisfatte dalle matrici A e B:

$$L_A \circ \psi_{\mathfrak{B}} = \psi_{\mathfrak{C}} \circ f, \qquad L_B \circ \psi_{\mathfrak{C}} = \psi_{\mathfrak{D}} \circ g.$$

Ora, osserviamo innanzitutto che:

$$L_C = L_{BA} = L_B \circ L_A.$$

Infatti, per ogni $x \in \mathbb{K}^n$ si ha, stante l'associatività del prodotto righe per colonne:

$$L_{BA}(x) = (BA)x = B(Ax) = B(L_A(x)) = L_B(L_A(x)) = L_B \circ L_A(x).$$

Ciò premesso, possiamo completare la nostra verifica:

$$L_C \circ \psi_{\mathfrak{B}} = L_B \circ (L_A \circ \psi_{\mathfrak{B}}) = (L_B \circ \psi_{\mathfrak{C}}) \circ f = \psi_{\mathfrak{D}} \circ (g \circ f),$$

come volevasi. □

Prima di procedere facciamo un'utile osservazione riguardante la matrice associata all'applicazione identica:

Esempio 32.2. 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 la (26):

$$M_{\mathfrak{B}}^{\mathfrak{B}}(Id_V) = \big(\psi_{\mathfrak{B}}(Id_V(v_1)) \quad \cdots \quad \psi_{\mathfrak{B}}(Id_V(v_n))\big)$$
$$\big(\psi_{\mathfrak{B}}(v_1) \quad \cdots \quad \psi_{\mathfrak{B}}(v_n))\big) = \big(e_1 \quad \cdots \quad e_n\big) = I_n.$$

Corollario 32.3. *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*

$$M^{\mathfrak{B}}_{\mathfrak{C}}(f^{-1}) = M^{\mathfrak{C}}_{\mathfrak{B}}(f)^{-1}. \tag{29}$$

Dimostrazione: Dalle relazioni

$$f \circ f^{-1} = Id_W, \ f^{-1} \circ f = Id_V$$

segue, in forza del Teorema precedente:

$$M^{\mathfrak{C}}_{\mathfrak{B}}(f) M^{\mathfrak{B}}_{\mathfrak{C}}(f^{-1}) = M^{\mathfrak{C}}_{\mathfrak{C}}(Id_W) = I_n, \ M^{\mathfrak{B}}_{\mathfrak{C}}(f^{-1}) M^{\mathfrak{C}}_{\mathfrak{B}}(f) = M^{\mathfrak{B}}_{\mathfrak{B}}(Id_V) = I_n$$

da cui l'asserto. □

33. Cambiamenti di base

Sia V uno spazio vettoriale di dimensione $n > 0$ sul campo $\mathbb{K}$ e siano assegnate due basi $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{B}' = \{v'_1, \ldots, v'_n\}$ di V.

Definizione 33.1. 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 $v'_1, \ldots, v'_n$ rispetto a $\mathfrak{B}$.

Esempio 33.2. 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:

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

Proposizione 33.3. *Assegnate due basi $\mathfrak{B}$ e $\mathfrak{B}'$ di 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

$$M_{\mathfrak{B},\mathfrak{B}'} = (\psi_{\mathfrak{B}}(v_1') \ \cdots \ \psi_{\mathfrak{B}}(v_n')) = (\psi_{\mathfrak{B}}(Id_V(v_1')) \ \cdots \ \psi_{\mathfrak{B}}(Id_V(v_n'))) = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V).$$

Poichè Id_V è un isomorfismo, segue dal Corollario 31.3 che $M_{\mathfrak{B},\mathfrak{B}'}$ è invertibile. L'ultima affermazione segue dal Corollario 32.3 applicato a Id_V. □

Il risultato seguente illustra come cambiano le coordinate di un vettore effettuando un cambiamento di base:

Teorema 33.4. *Dato un vettore $v \in V$, siano $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $x' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$ i vettori delle componenti di v rispetto alle basi $\mathfrak{B}$ e a $\mathfrak{B}'$. Denotata con M la matrice di passaggio da $\mathfrak{B}$ a $\mathfrak{B}'$, sussiste la relazione*

$$x = Mx'. \tag{30}$$

Dimostrazione: Siccome $M = M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V)$, basta applicare il Corollario 30.7 all'applicazione identica: per ogni $v \in V$ sappiamo infatti che il legame tra le coordinate v e di $v = Id_V(v)$ rispetto alla due basi (dove va considerata prima $\mathfrak{B}'$ e poi $\mathfrak{B}$) è proprio dato dalla (30). □

Studiamo infine come cambia la matrice associata ad un'applicazione lineare facendo scelte diverse delle basi negli spazi di partenza e di arrivo:

Teorema 33.5. *(del cambiamento di base)*
Siano V, W spazi vettoriali entrambi non banali sullo stesso campo $\mathbb{K}$, entrambi finitamente generati ed aventi dimensioni n ed m rispettivamente. Sia $f : V \to W$ lineare. Si fissino due basi $\mathfrak{B}$, $\mathfrak{B}'$ di V e due basi $\mathfrak{C}$, $\mathfrak{C}'$ di W. Allora risulta

$$M^{\mathfrak{C}'}_{\mathfrak{B}'}(f) = C^{-1} M^{\mathfrak{C}}_{\mathfrak{B}}(f)\, B \tag{31}$$

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 32.1, si ottiene quindi:

$$M^{\mathfrak{C}'}_{\mathfrak{B}'}(f) = M^{\mathfrak{C}'}_{\mathfrak{B}'}(Id_W \circ (f \circ Id_V)) = M^{\mathfrak{C}'}_{\mathfrak{C}}(Id_W) M^{\mathfrak{C}}_{\mathfrak{B}'}(f \circ Id_V) =$$
$$M^{\mathfrak{C}'}_{\mathfrak{C}}(Id_W) M^{\mathfrak{C}}_{\mathfrak{B}}(f) M^{\mathfrak{B}}_{\mathfrak{B}'}(Id_V) = C^{-1} M^{\mathfrak{C}}_{\mathfrak{B}}(f) B.$$

□

34. 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 34.1. 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

$$g(u, \lambda v + \mu w) = \lambda g(u, v) + \mu g(v, w)$$

che si ottiene direttamente dalla 2). Ciò giustifica la terminologia "bilineare": ovvero 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.

Definizione 34.2. Si chiama *spazio vettoriale Euclideo* ogni coppia (V, g) costituita da uno spazio vettoriale reale ed un fissato prodotto scalare su V.

Osserviamo anche quanto segue riguardo il prodotto scalare del vettore nullo con gli altri vettori:

Proposizione 34.3. *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 abbiamo:

$$g(v, 0_V) = g(v, 0_V + 0_V) = g(v, 0_V) + g(v, 0_V).$$

□

Osservazione 34.4. La proprietà (3) si può riformulare come segue: per ogni v si ha $g(v, v) \geq 0$ e

$$g(v, v) = 0 \Rightarrow v = 0.$$

Discutiamo alcuni esempi.

Definizione 34.5. 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 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, oltre la verifica diretta, possiamo notare che la definizione di g_0 può anche scriversi 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 conveniamo che ogni vettore di $y \in \mathbb{R}^n$ è un vettore colonna (per cui ${}^t x$ è un vettore riga). Ciò premesso, la bilinearità segue direttamente dalle proprietà standard del prodotto tra matrici e dell'operazione di trasposizione (che è lineare):

$$g_0(\lambda x + \mu y, z) = {}^t(\lambda x + \mu y) z = (\lambda\, {}^t x + \mu\, {}^t y) z = \lambda\, {}^t x\, z + \mu\, {}^t y\, z = \lambda g_0(x, z) + \mu g_0(y, z).$$

Esempio 34.6. È possibile munire lo spazio $M_n(\mathbb{R})$ delle matrici quadrate di ordine n di un prodotto scalare ponendo

$$g(A, B) := tr({}^t A B).$$

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 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 $C = {}^tAA = (c^i_j)$, 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 la i-ma riga della trasposta di A coincide con la sua i-ma colonna, pensata come riga, ovvero coincide con il vettore riga ${}^tA_{(i)}$; dunque c^i_i si può esprimere mediante il prodotto scalare standard:

$$c^i_i = {}^tA_{(i)} A_{(i)} = g_o(A_{(i)}, A_{(i)}).$$

In conclusione

$$g(A, A) = tr(C) = \sum_{i=1}^{n} g_o(A_{(i)}, A_{(i)}).$$

Dunque $g(A, A) \geq 0$, e se $g(A, A) = 0$, allora per ogni colonna $g_o(A_{(i)}, A_{(i)}) = 0$ e quindi $A_{(i)} = 0$ perchè g_0 è un prodotto scalare. Pertanto $A = 0$.

Esempio 34.7. Definiamo un prodotto scalare su $V = Hom(\mathbb{R}^n, \mathbb{R})$, lo spazio di tutti i funzionali lineari $f : \mathbb{R}^n \to \mathbb{R}$, ponendo:

$$< f, g >:= f(e_1)g(e_1) + \cdots + f(e_n)g(e_n),$$

dove $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{R}^n$. La proprietà di simmetria è chiara, riguardo la bilinearità basta tener conto della definizione di somma tra funzionali e prodotto per uno scalare:

$$< \lambda f + \mu g, h >= \sum_{i=1}^{n} (\lambda f + \mu g)(e_i)h(e_i) = \sum_{i=1}^{n} (\lambda f(e_i) + \mu g(e_i))h(e_i) =$$

$$= \lambda \sum_{i=1}^{n} f(e_i)h(e_i) + \mu \sum_{i=1}^{n} g(e_i)h(e_i) = \lambda < f, h > + \mu < g, h > .$$

Infine, per ogni funzionale $f : \mathbb{R}^n \to \mathbb{R}$ si ha $< f, f >= \sum_{i=1}^{n} f(e_i)^2 \geq 0$; se tale quantità è zero, allora

$$f(e_i) = 0$$

per ogni vettore della base. Ciò è sufficiente per concludere che f è la funzione nulla (Teorema 29.1).

Esempio 34.8. 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 (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 34.9. 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_0) uno spazio vettoriale Euclideo e sia

$$f : V \to W$$

un monomorfismo dove V è un qualsiasi spazio vettoriale reale. Allora g_0 induce un prodotto scalare g su V, definito come segue:

$$g(u, v) := g_0(f(u), f(v)).$$

Tale prodotto scalare si chiama il *pull-back* di g_0 mediante f e si denota di solito con $g = f^*(g_0)$. Si lasciano le verifiche come esercizio.

Definizione 34.10. 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:

$$||\lambda v||_g = |\lambda| ||v||_g \tag{32}$$

Definizione 34.11. 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:

$$\frac{v}{||v||_g}.$$

Esempio 34.12. I vettori della base canonica $\{e_i\}$ sono tutti versori dello spazio Euclideo $(\mathbb{R}^n, g_0)$. Inoltre essi sono mutuamente ortogonali, nel senso che

$$g_0(e_i, e_j) = 0 \quad \text{se } i \neq j.$$

Nel prossimo paragrafo studiamo in generale basi che hanno questa proprietà e che sono uno strumento basilare nella teoria degli spazi vettoriali Euclidei.

35. Basi ortonormali

Definizione 35.1. Sia (V, g) uno spazio vettoriale Euclideo di dimensione $n \geq 1$. Una sequenza di 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.

Se è soddisfatta solo quest'ultima condizione e i vettori v_i sono tutti non nulli (non necessariamente versori), si parla invece di *base ortogonale.*

Ogni base ortonormale è effettivamente una base di V: ciò è conseguenza del risultato seguente, che ne garantisce la lineare indipendenza (pertanto trattandosi di n vettori è una base):

Proposizione 35.2. *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

$$\lambda_j = 0.$$

$\square$

Proviamo ora l'esistenza di questo tipo di basi.

Teorema 35.3. *Sia (V, g) uno spazio vettoriale Euclideo di dimensione $n \geq 1$. Si fissi una base*

$$\mathfrak{B} = \{v_1, \dots, v_k\}.$$

Allora esiste una base ortonormale

$$\mathfrak{B}' = \{v'_1, \dots, v'_n\}$$

tale che, per ogni k, $1 \leq k \leq n$ si abbia:

$$L(v_1, \dots, v_k) = L(v'_1, \dots, v'_k).$$

Prima di procedere con la dimostrazione osserviamo preliminarmente che, se (V, g) è uno spazio vettoriale Euclideo, ogni sottospazio $W \subset V$ è esso stesso uno spazio vettoriale Euclideo in modo canonico: la restrizione

$$g|_{W \times W} : W \times W \to \mathbb{R}$$

del prodotto scalare di V è manifestamente un prodotto scalare anche su W.

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 supponiamo l'asserto vero per spazi vettoriali Euclidei di dimensione $n - 1$. Consideriamo quindi il sottospazio $W := L(v_1, \dots, v_{n-1})$; esso ha dimensione $n - 1$ e $\{v_1, \dots, v_{n-1}\}$ è una sua base. Per l'ipotesi induttiva, W ammette una base ortonormale

$$v'_1, \dots, v'_{n-1}$$

tale che

$$L(v_1, \dots, v_k) = L(v'_1, \dots, v'_k)$$

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:

$$u_n := v_n - g(v_n, v'_1)v'_1 - \cdots - g(v_n, v'_{n-1})v'_{n-1}.$$

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

$$v_n \in L(v'_1, \dots, v'_{n-1}) = L(v_1, \dots, v_{n-1})$$

contro la lineare indipendenza dei vettori della base $\mathfrak{B}$.

Inoltre risulta che u_n è ortogonale a tutti vettori $v'_1, \dots, v'_{n-1}$: fissato infatti uno di questi, v'_j, risulta:

$$\begin{aligned} g(u_n, v'_j) &= g(v_n - g(v_n, v'_1)v'_1 - \dots - g(v_n, v'_{n-1})v'_{n-1}, v'_j) = \\ &= g(v_n, v'_j) - g(v_n, v'_j)g(v'_j, v'_j) = \\ &= g(v_n, v'_j) - g(v_n, v'_j) = 0 \end{aligned}$$

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 $v'_1, \dots, v'_n$ 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 35.4. 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:

$$v'_1 = \frac{v_1}{||v_1||} = \frac{1}{\sqrt{2}}(1,0,1) = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}).$$

$$u_2 = v_2 - g_0(v_2, v'_1)v'_1 = (0,1,-1) + \frac{\sqrt{2}}{2}v'_1 = (\frac{1}{2}, 1, -\frac{1}{2})$$

$$v'_2 = \frac{u_2}{||u_2||} = (\frac{1}{3}\sqrt{\frac{3}{2}}, \frac{2}{3}\sqrt{\frac{3}{2}}, -\frac{1}{3}\sqrt{\frac{3}{2}})$$

$$u_3 = v_3 - g_0(v_3, v'_1)v'_1 - g_0(v_3, v'_2)v'_2 = (0,3,0) - 0 - (1,2,-1) = (-1,1,1)$$

$$v'_3 = \frac{u_3}{||u_3||} = (-\frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}).$$

36. 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 36.1. *Sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ 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 $(x_1, \ldots, x_n)$ 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 standard disponendo di una base ortonormale:

Proposizione 36.2. *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:*

$$g(u, v) = \sum_{i=1}^{n} x_i y_i = {}^t x\, y. \tag{33}$$

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 e ancora del fatto che gli unici prodotti scalari non nulli $g(v_i, v_j)$ sono quelli del tipo $g(v_i, v_i)$:

$$g(u, v) = g(\sum_{i=1}^{n} x_i v_i, v) = \sum_{i=1}^{n} x_i g(v_i, v) = \sum_{i=1}^{n} x_i g(v_i, \sum_{k=1}^{n} y_k v_k) = \sum_{i=1}^{n} x_i y_i g(v_i, v_i) = \sum_{i=1}^{n} x_i y_i.$$

Questo risultato mostra che effettuare il 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.

Nel prossimo paragrafo interpreteremo questo fatto mediante il concetto di isometria tra spazi vettoriali Euclidei.

37. Isometrie lineari

In questo paragrafo denotiamo con (V, g) e (W, g') due spazi vettoriali Euclidei della stessa dimensione $n \geq 1$.

Definizione 37.1. 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:

$$g(u, v) = g'(F(u), F(v)) \tag{34}$$

per ogni $u, v \in V$.

Teorema 37.2. *Siano (V, g) e (W, g') spazi vettoriali Euclidei della stessa dimensione $n \geq 1$. Ogni isometria lineare $F : V \to W$ è un isomorfismo. Inoltre F^{-1} è anch'essa un'isometria lineare.*

Dimostrazione: Sia $v \in Ker(F)$; allora

$$g(v, v) = g'(F(v), F(v)) = 0,$$

da cui $v = 0$. Resta provato che F è un monomorfismo, e ciò basta per poter affermare che è un isomorfismo, in quanto gli spazi V e W hanno la stessa dimensione (Teorema 31.1).

Riguardo l'ultima affermazione, la (34) si può riscrivere

$$g'(F(u), F(v)) = g(F^{-1}(F(u)), F^{-1}(F(v)))$$

e, poichè ogni vettore di W, stante la surgettività di F, è del tipo $F(u)$ per un certo $u \in V$, abbiamo che anche F^{-1} conserva i prodotti scalari. $\square$

Questo concetto specializza la nozione di isomorfismo tra spazi vettoriali nel contesto euclideo; due spazi vettoriali Euclidei (V, g) e (W, g') si diranno *isometrici* se esiste un isometria lineare tra essi.

Esempio 37.3. Fissata una base ortonormale $\mathfrak{B} = \{v_1, \ldots, v_n\}$ 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 $(\mathbb{R}^n, g_0)$, dove g_0 è il prodotto scalare standard.

Infatti per ogni coppia di vettori u e v aventi vettori di componenti x e y sappiamo che (Prop. 36.2):

$$g(u, v) = {}^t x\, y = g_0(x, y) = g_0(\psi_{\mathfrak{B}}(u), \psi_{\mathfrak{B}}(v)).$$

Utilizzando questo esempio otteniamo il seguente risultato di classificazione:

Teorema 37.4. *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 37.5. *Sia $F : V \to W$ un'applicazione lineare. Sia $\mathfrak{B} = \{v_1, \ldots, v_n\}$ una base ortonormale di V. Sono equivalenti:*

a) F è un'isometria.

b) $\{F(v_1), \ldots, F(v_n)\}$ è base ortonormale di W.

Notiamo che questa caratterizzazione si può combinare con il Teorema di esistenza ed unicità delle applicazioni lineari per costruire isometrie lineari: supponiamo assegnate due basi ortonormali $\mathfrak{B} = \{v_1, \ldots, v_n\}$ e $\mathfrak{C} = \{w_1, \ldots, w_n\}$ degli spazi (V, g) e (W, g'). Allora vi è l'unica applicazione lineare $F : V \to W$ che trasforma la prima base nella seconda, cioè tale che:

$$F(v_i) = w_i, \quad i = 1, \ldots, n.$$

Essa è automaticamente un isomorfismo per il Teorema 31.2. Ora, essendo $\mathfrak{C}$ una base ortonormale, il risultato precedente (mediante l'implicazione $b) \Rightarrow a)$), garantisce che tale F è un'isometria lineare.

Dimostrazione:

$a) \Rightarrow b)$ È immediata perchè F conserva il prodotto scalare.

$b) \Rightarrow a)$ Supponiamo che $\{F(v_1), \ldots, F(v_n)\}$ 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 $(x_1, \ldots, x_n)$ allora

$$F(v) = F(x_1 v_1 + \cdots + x_n v_n) = x_1 F(v_1) + \cdots + x_n F(v_n).$$

(alternativamente si può anche notare che la matrice $M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ coincide in questo con la matrice identica I_n e quindi applicare il Cor. 30.7).

Ciò premesso, la tesi segue subito dalla (33) 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 $g_0(x, y)$ dei vettori delle componenti di u e v. $\square$

38. Matrici ortogonali

Definizione 38.1. Una matrice quadrata $A \in M_n(\mathbb{R})$ si dice *ortogonale* se

$$ {}^t A A = I_n. \tag{35}$$

Per definizione I_n è ortogonale.

Proposizione 38.2. *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 (35) 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 (35) 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 38.3. *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 (35):

Proposizione 38.4. *Per ogni matrice $A \in M_n(\mathbb{R})$, le proprietà seguenti sono equivalenti:*

a) A è ortogonale.
b) Le colonne $A_{(1)}, \dots, A_{(n)}$ costituiscono una base ortonormale di $\mathbb{R}^n$.
c) Le righe $A^{(1)}, \dots, A^{(n)}$ costituiscono una base ortonormale di $\mathbb{R}^n$.

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

$$c^i_j = ({}^tA)^{(i)} \cdot A_{(j)} = {}^tA_{(i)} \cdot A_{(j)}$$

che si riscrive:

$$c^i_j = g_0(A_{(i)}, A_{(j)}).$$

Da qui la conclusione è immediata. $\square$

Esempio 38.5. Le matrici ortogonali di ordine 2 sono determinabili esplicitamente come segue. Notiamo intanto che per ogni vettore $(a, b) \in \mathbb{R}^2$ i vettori (x, y) ortogonali ad esso sono tutti e soli quelli le cui coordinate soddisfano l'equazione lineare:

$$ax + by = 0.$$

Pertanto, se $(a, b) \neq 0$, allora sono tutti e soli i vettori del tipo

$$(-\lambda b, \lambda a)$$

al variare di $\lambda \in \mathbb{R}$. Ora, in base alla caratterizzazione precedente una matrice

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

è ortogonale se e solo se le colonne costituiscono una base ortonormale di $\mathbb{R}^2$, il che accade se e solo se sono sodisfatte le relazioni:

$$a^2 + b^2 = 1,\ c = -\lambda b, d = \lambda a,\ c^2 + d^2 = 1$$

da cui si trae subito che

$$O(2) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} :\ a^2 + b^2 = 1 \right\} \cup \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} :\ a^2 + b^2 = 1 \right\}.$$

Usando le funzioni trigonometriche, queste matrici si possono riscrivere:

$$O(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\} \cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}.$$

Tenendo con di questa descrizione esplicita, il lettore può verificare in modo diretto che per ogni matrice $A \in O(2)$, l'applicazione lineare $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ ad essa associata è un'isometria lineare (sottintenendo sempre il prodotto scalare standard su $\mathbb{R}^2$). Ciò è vero in realtà per ogni matrice $A \in O(n)$ in quanto:

$$g_0(L_A(x), L_A(y)) = {}^t(Ax)Ay = {}^tx\, {}^tAA\, y = {}^tx\, I_n\, y = g_o(x, y).$$

Nel paragrafo seguente approfondiamo in generale qual è il legame tra le due nozioni in esame.

39. Relazione tra isometrie lineari e matrici ortgonali

Teorema 39.1. *Sia $F : (V, g) \to (W, g')$ un'applicazione lineare. 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 $\mathfrak{B} = \{v_1, \dots, v_n\}$, abbiamo che le colonne della matrice $A := M^{\mathfrak{C}}_{\mathfrak{B}}(F)$ sono gli n vettori

$$\psi_{\mathfrak{C}}(F(v_1)), \dots, \psi_{\mathfrak{C}}(F(v_n))$$

di $\mathbb{R}^n$. Ora, in base al Teorema 37.5, F è un'isometria se e solo se se i vettori $\{F(v_1), \dots, F(v_n)\}$ costituiscono una base ortonormale di W, e ciò è vero se e solo se i vettori $\{\psi_{\mathfrak{C}}(F(v_1)), \dots, \psi_{\mathfrak{C}}(F(v_n))\}$ 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 $A \in O(n)$. $\square$

40. Ortogonalità in uno spazio vettoriale Euclideo

Il concetto di perpendicolarità tra vettori si estende a sottospazi in modo naturale, come segue:

Definizione 40.1. 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 40.2. Se $W = L(w_1, \dots, w_k)$ allora risulta:

$$W^\perp = \{v \in V | g(v, w_i) = 0 \ \ \forall i = 1, \dots, k\}.$$

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 $w = x_1 w_1 + \dots + x_k w_k$ e quindi

$$g(v, w) = \sum_{i=1}^{k} x_i g(v, w_i) = 0.$$

Teorema 40.3. *Per ogni sottospazio W di uno spazio vettoriale Euclideo (V, g) si ha:*

1) $W^\perp$ è un sottospazio vettoriale di V.

2) $V = W \oplus W^\perp$.

DIMOSTRAZIONE: Riguardo 1), notiamo che evidentemente $0 \in W^{\perp}$, per cui tale insieme non è vuoto. Verifichiamo che ogni combinazioni lineare $\lambda u + \mu v$ di vettori u e v di $W^{\perp}$ appartiene allo stesso insieme; a tal scopo, considerato un vettore arbitrario $w \in W$, è sufficiente osservare che:

$$g(\lambda u + \lambda v, w) = \lambda g(u, w) + \mu g(\lambda v, w) = 0$$

perchè i prodotti scalari $g(u, w)$ e $g(v, w)$ sono entrambi nulli.

2) Risulta

$$W \cap W^{\perp} = \{0\}$$

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 in base alla definizione si ha che $\{0\}^{\perp} = V$ e quindi se il sottospazio è banale, l'asserto è senz'altro vero.

Fissiamo quindi una base ortonormale $\mathfrak{B} = \{w_1, \dots, w_k\}$ di W (rispetto alla restrizione del prodotto scalare g a W), dove $k = \mathtt{dim}(W) \geq 1$. Dato un vettore qualsiasi $v \in V$, consideriamo il seguente vettore di W:

$$w := g(v, w_1)w_1 + \cdots + g(v, w_n)w_n$$

e poniamo

$$u := v - w.$$

Risulta 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 $j \in \{1, \dots, k\}$ abbiamo:

$$g(u, w_j) = g(v, w_j) - g(\sum_{i=1}^{k} g(v, w_i)w_i, w_j) = g(v, w_j) - g(v, w_j) = 0.$$

Conclusione:

$$v = w + u$$

con $w \in W$ e $u \in W^{\perp}$ e ciò prova, stante l'arbitrarietà di v, che $V = W + W^{\perp}$. $\square$

Applicando la formula di Grassmann consegue:

Corollario 40.4. *Per ogni sottospazio W di uno spazio vettoriale Euclideo (V, g) si ha:*

$$\mathtt{dim}(W^{\perp}) = \mathtt{dim}(V) - \mathtt{dim}(W).$$

Come altra conseguenza del Teorema precedente, risulta che fissato un sottospazio W 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.

Osservazione 40.5. 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 $u \in W^{\perp}$ e $u \in U$, si ha che $g(u, v) = 0$ perchè u è anche un vettore di W.

41. 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 41.1. Sia chiama **sottopazio affine** di $\mathbb{R}^n$ ogni sottoinsieme del tipo

$$P_o + W := \{P_o + w |\, w \in \mathbb{R}^n\}$$

dove $P_0 \in \mathbb{R}^n$ è un punto fissato e $W \subset \mathbb{R}^n$ è un sottospazio vettoriale.

Diremo che $P_o + W$ 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 41.2. *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. $\square$

Si introduce anche la seguente terminologia:

Definizione 41.3. Dato un sottospazio affine di giacitura $S = P_0 + W$, 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 41.4. 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:

$$S = \left\{ P = (x_1, \ldots, x_n) : A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\} \tag{36}$$

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 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 una notazione utile che permette di caratterizzare in modo conveniente l'appartenenza di un punto ad un dato sottospazio affine:

Definizione 41.5. Dati due punti A, B di $\mathbb{R}^n$, il vettore

$$\overrightarrow{AB} := B - A$$

si chiamerà vettore applicato di estremi A e B.

Proposizione 41.6. *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. □

Utilizzeremo questo risultato per dare la seguente importante caratterizzazione dei sottospazi affini, che afferma che l'esempio discusso sopra ne costituisce il prototipo:

Teorema 41.7. *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

$$Ax = b.$$

Inoltre, vera a) oppure b), la giacitura W di S è lo spazio delle soluzioni del sistema lineare omogeneo

$$Ax = 0$$

e si ha $\mathtt{dim}(S) = n - rg(A)$.

DIMOSTRAZIONE: Abbiamo già discusso che $b) \Rightarrow a)$ nell'Esempio 41.4, insieme alle ultime affermazioni. Resta da provare che $a) \Rightarrow b)$. Supponiamo che S sia il sottospazio affine $S = P_0 + W$ e consideriamo un sistema di generatori $u_1, \ldots, u_k$ del sottospazio $W^\perp$, di modo che

$$W^\perp = L(u_1, \ldots, u_k).$$

Abbiamo che un punto $P = (x_1, \ldots, x_n)$ 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 $W^\perp$, ma ciò avviene se e solo se

$$\overrightarrow{P_0P} \cdot u_i = 0, \quad i = 1, \ldots, k. \tag{37}$$

(si ricordi l'Oss. 40.2). Ciò fornisce un sistema di equazioni che descrive S: infatti posto $P_0 = (x_0^1, \ldots, x_0^n)$ e $u_i = (u_i^1, \ldots, u_i^n)$, le (37) si riscrivono

$$(P - P_o) \cdot u_i = 0 \quad i = 1, \ldots, k$$

che si esplicitano come segue:

$$\begin{cases} u_1^1 x_1 + \cdots + u_1^n x_n = b_1 \\ \cdots \quad \cdots \quad \cdots \\ u_k^1 x_1 + \cdots + u_k^n x_n = b_k \end{cases}$$

dove abbiamo denotato con b_i il prodotto scalare $P_0 \cdot u_i$. Si tratta di un sistema lineare nelle incognite $(x_1, \ldots, x_n)$, compatibile perchè $(x_0^1, \ldots, x_0^n)$ ne è una soluzione per costruzione, il cui insieme delle soluzioni è esattamente S. □

Per sintetizzare il fatto che un sottospazio affine S è l'insieme delle soluzioni di un sistema lineare, si scrive di solito

$$S : Ax = b$$

e le equazioni del sistema si dicono le equazioni di S. Questa scrittura corrisponde esplicitamente alla (38).

Definizione 41.8. I sottospazi affini di dimensione 1 di $\mathbb{R}^n$ si dicono *rette*, mentre sottospazi di dimensione 2 si dicono *piani*. Inoltre si chiamano *iperpiani* i sottospazi affini di dimensione $n-1$.

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.

Esempio 41.9. 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.

42. 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 $P = (x_1, \ldots, x_n)$; 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 42.1. 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 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 42.2. 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 $(x_1, \ldots, x_n)$. 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 42.3. Il riferimento *standard* di $\mathbb{R}^n$ è $\mathfrak{R}_0 = (0, \mathfrak{B}_0)$ 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 42.4. *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 42.5. Naturalmente 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 $P(x_1, \ldots, x_n)$ e $Q(x'_1, \ldots, x'_n)$ coincidono se e solo se hanno le stesse coordinate, cioè $x_i = x'_i$ per ogni i, mentre fissata $(x^1_o, \ldots, x^n_o)$ una n-pla di scalari, vi è esattamente un punto P di coordinate $(x^1_o, \ldots, x^n_o)$.

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

$$P = O + u_1 - 2u_2 = (1,2) + (3,1) - 2(\sqrt{2}, -1) = (4 - 2\sqrt{2}, 5).$$

In presenza di un riferimento qualsiasi, il Teorema 41.7 ammette la seguente generalizzazione:

Teorema 42.6. *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

$$S = \left\{ P(x_1, \ldots, x_n) : A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \right\} \tag{38}$$

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

$$Ax = 0.$$

Si ha inoltre $\mathtt{dim}(S) = n - rg(A)$.

La giustificazione di questo risultato risiede nella seguente proposizione, che concerne la formula di cambiamento delle coordinate affini di un punto per un cambio di riferimento.

Dato un punto P di coordinate $(x_1, \ldots, x_n)$ rispetto ad un fissato riferimento affine, scriveremo $P(x)$, dove con x il vettore $(x_1, \ldots, x_n)$ costituito dalle coordinate. Inoltre per ogni vettore v scriveremo $v(x)$ per intendendere che esso ha vettore di componenti x rispetto alla base del riferimento.

Proposizione 42.7. *(Cambiamento di riferimento) Siano* $\mathfrak{R} = (O, \mathfrak{B})$ *e* $\mathfrak{R}' = (O', \mathfrak{B}')$ *due riferimenti affini di* $\mathbb{R}^n$. *Per ogni punto* P, *se* $P(x)$ *rispetto a* $\mathfrak{R}$ *e* $P(x')$ *rispetto a* $\mathfrak{R}$, *allora si ha:*

$$x = Cx' + d \tag{39}$$

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 (39) 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 (39) si ricava uguagliano i vettori delle componenti dei vettori ad ambo i membri, rispetto alla base $\mathfrak{B}$. Per quel che concerne il primo membro, tale vettore è il vettore delle coordinate x di P rispetto a $\mathfrak{R}$; riguardo il secondo, si tratta di sommare le componenti di $\overrightarrow{OO'}$ e di $\overrightarrow{O'P}$. Il vettore $\overrightarrow{OO'}$ ha componenti date dal vettore d delle coordinate di O' rispetto ad $\mathfrak{R}$. Infine, applicando la (30), il vettore $\overrightarrow{O'P}$ ha componenti Cx', in quanto esso ha componenti x' nella base $\mathfrak{B}'$.

Nel caso di riferimenti Cartesiani, si utilizza il seguente fatto di natura generale:

Proposizione 42.8. *Siano* $\mathfrak{B}$ *e* $\mathfrak{B}'$ *due basi ortonormali di uno spazio vettoriale Euclideo n-dimensionale* (V, g). *Allora*

$$M_{\mathfrak{B},\mathfrak{B}'} \in O(n).$$

Dimostrazione: Infatti, basta ricordare che la matrice $M_{\mathfrak{B},\mathfrak{B}'}$ coincide con la matrice associata all'applicazione identica $Id_V : V \to V$ rispetto alle basi $\mathfrak{B}'$ e $\mathfrak{B}$; siccome questa è un'isometria lineare, tale matrice è ortogonale per il Teorema 39.1. □

Siamo ora in grado di provare il Teorema 42.6. Fissiamo un riferimento affine $\mathfrak{R}$ e 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, utilizzando le coordinate rispetto al riferimento standard $\mathfrak{R}_0 = (0, \mathfrak{B}_0)$:

$$S = \{P(x) : \ Ax = b\}.$$

Utilizzando la (39), rispetto al riferimento $\mathfrak{R}'$ lo stesso insieme è descritto come segue:

$$S = \{P(x') : \ A(Cx' + d) = b\}$$

cioè

$$S = \{P(x') : \ A'x' = b'\},$$

dove abbiamo posto $A' := AC$ e $b' := b - Ad$. Riguardo la giacitura W di S, in modo simile, sfruttando ancora la (30) abbiamo che

$$W = \{v = (x_1, \ldots, x_n) \in \mathbb{R}^n : Ax = 0\} = \{v(x') : \ ACx' = 0\} = \{v(x') : \ A'x' = 0\}.$$

43. 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:

$$P : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \\ a''x + b''y + c''z + d'' = 0 \end{cases}, \quad \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \neq 0.$$

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, si ottiene quindi che i punti di r sono tutti e soli i punti di coordinate

$$\begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad t \in \mathbb{R}.$$

e queste equazioni sono le equazioni parametriche della retta.

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 di $t = 0$.

44. 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; essa è incentrata sui concetti definiti di seguito:

Definizione 44.1. 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, se S e T sono sottospazi affini incidenti, con giaciture U e W, allora

$$S \subset T \iff U \subset W.$$

Infatti, fissato un punto $P \in S \cap T$, possiamo rappresentare i due sottospazi nella forma

$$S = P + U, \; T = P + W \tag{40}$$

e quindi è evidente che, se $U \subset W$, allora $S \subset T$. L'altra implicazione si lascia come esercizio.

Riguardo il parallelismo consegue:

Proposizione 44.2. *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. □

Nel seguito scriveremo $S//T$ per indicare il fatto che S e T sono sottospazi paralleli.

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:

$$\alpha \cap \beta = r \text{ retta}, \; \alpha \cap \beta = \emptyset.$$

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, discutiamo alcuni risultati di natura generale su cui si basa lo studio della geometria di posizione in $\mathbb{R}^n$.

Proposizione 44.3. *Siano S e T sottospazi affini incidenti di $\mathbb{R}^n$. Allora $S \cap T$ è un sottospazio affine.*

Dimostrazione: Fissato un punto $P \in S \cap T$ rappresentiamo ancora S e T mediante la (40). Allora risulta

$$S \cap T = P + (U \cap W)$$

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 (41.6) di appartenza di un punto ad un dato sottospazio affine. Infatti, se Q appartiene a $S \cap T$ allora si ha sia $\overrightarrow{PQ} \in U$ che $\overrightarrow{PQ} \in W$. Ma ciò garantisce che $\overrightarrow{PQ} \in U \cap W$ e quindi che $Q \in P + (U \cap W)$. Resta così giusticata l'inclusione $\subset$. L'altra è simile e si lascia al lettore. □

Sussiste anche la seguente caratterizzazione dei sottospazi incidenti:

Proposizione 44.4. *Siano $S = A + U$ e $T = B + W$ due sottospazi affini di $\mathbb{R}^n$. Allora*

$$S \cap T \neq \emptyset \iff \overrightarrow{AB} \in U + W.$$

La dimostrazione (un utile esercizio che si può svolgere agevolmente utilizzando ad es. l'identià di Chasles) si omette per brevità.

Ai fini 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 44.5. *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:*

$$S \vee T = A + U + W + L(\overrightarrow{AB}).$$

Vale infine la seguente formula per calcolare la dimensione di $S \vee T$, che è la versione affine della formula di Grassmann (18):

Teorema 44.6. *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 $S \cap T = \emptyset$: $\mathtt{dim}(S \vee T) = \mathtt{dim}(S) + \mathtt{dim}(T) + 1 - \mathtt{dim}(U \cap W)$.

Omettiamo la dimostrazione, che può farsi senza difficoltà usando la formula di Grassmann vettoriale e tenendo conto della Prop. 44.4. Ci limitiamo a illustrare come applicare questo risultato per provare in modo rigoroso le affermazioni fatte sopra sulle rette e i piani di $\mathbb{R}^3$.

Proposizione 44.7. *Due piani di $\mathbb{R}^3$ non sono mai sghembi. Quindi se α e β sono piani distinti, si ha*

$$\alpha // \beta \iff \alpha \cap \beta = \emptyset.$$

Inoltre, se α e β sono incidenti, allora $\alpha \cap \beta$ è una retta.

Dimostrazione: Consideriamo due piani α e β e supponiamo per assurdo che siano sghembi; per definizione, $\alpha \cap \beta = \emptyset$ ed inoltre α e β non sono paralleli, per cui le corrispondenti giaciture U e W sono diverse. Consegue che $U \cap W$ ha dimensione 1: infatti tale sottospazio non può essere banale perchè per la formula di Grassmann vettoriale si avrebbe $\mathtt{dim}(U+W) = 4$, e non puó avere dimensione 2 essendo $U \neq W$. Applicando quindi la formula di Grassmann affine nel caso 2), si avrebbe

$$\mathtt{dim}(S \vee T) = 2 + 2 + 1 - 1 = 4$$

pervenendo ad un assurdo. Supponendo ora $\alpha \neq \beta$, è chiaro che se α e β sono paralleli, allora sono disgiunti (cfr. Prop. 44.2), e viceversa, se sono disgiunti, allora devono necessariamente essere paralleli per quanto appena provato. 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 due caratterizzazioni delle rette sghembe di $\mathbb{R}^3$. Definiamo due rette *complanari* se esiste un piano che contiene entrambe.

Proposizione 44.8. *Siano r e s rette distinte di $\mathbb{R}^3$. Allora*

r e s sono complanari $\iff$ r e s sono incidenti oppure sono parallele.

Dunque:

r e s sono sghembe $\iff$ r e s non sono complanari.

Se r e s sono complanari, allora esiste un solo piano che contiene entrambe.

Dimostrazione: Denotiamo ancora con U e W le giaciture delle due rette. Abbiamo intanto che r e s sono complanari se e solo se $r \vee s$ è un piano, perchè a priori tale spazio non può essere una retta, avendosi $r \neq s$. Da ciò segue anche l'ultima affermazione dell'enunciato. Quindi si tratta di caratterizzare quando $\mathtt{dim}(r \vee s) = 2$;

applicando la formula di Grassmann, tenendo conto delle due eventualità 1) oppure 2) si ricava che:

$$\texttt{dim}(r \vee s) = 2 \iff \texttt{dim}(r \cap s) = 0 \text{ oppure } \texttt{dim}(U \cap W) = 1$$

ma, essendo U e W entrambi 1-dimensionali consegue

$$\texttt{dim}(r \vee s) = 2 \iff \texttt{dim}(r \cap s) = 0 \text{ oppure } U = W$$

ovvero

$$\texttt{dim}(r \vee s) = 2 \iff r \cap s = \{P\} \text{ oppure } r//s.$$

e di qui la conclusione. □

Si ha quest'altra utile caratterizzazione:

Proposizione 44.9. *Siano r e s rette distinte di $\mathbb{R}^3$. Allora sono equivalenti:*

a) r e s sono sghembe;

b) r e s non sono parallele ed esistono due piani paralleli distinti α e β tali che

$$r \subset \alpha \text{ e } s \subset \beta.$$

Dimostrazione: L'implicazione $b) \Rightarrow a)$ segue dal fatto che α e β sono disgiunti e quindi tali sono a maggior ragione r e s. Proviamo che $a) \Rightarrow b)$. Assumendo r e s sghembe, e $r = A + U$ e $s = B + W$, dove $U = L(v)$ e $W = L(w)$, allora per definizione r e s non sono parallele; in particolare i vettori v e w devono essere linearmente indipendenti. Pertanto $L(u, v) = U + W$ è un sottospazio di dimensione 2 di $\mathbb{R}^3$. Per provare l'asserto basta considerare i piani paralleli:

$$\alpha := A + L(u, v), \quad \beta := B + L(u, v),$$

che per costruzione contengono rispettivamente r e s. Notiamo che tali piani sono distinti: ad esempio, il punto B non appartiene a α perchè altrimenti si avrebbe $\overrightarrow{AB} \in U + W$ contro il fatto che $r \cap s = \emptyset$ (cfr. Prop. 44.4). □

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 44.10. *Siano r una retta e α un piano di $\mathbb{R}^3$, e si assuma che r non sia contenuta in α. Allora o r e α sono incidenti, ed in tal caso $r \cap \alpha$ è un punto, oppure sono paralleli. Inoltre*

$$r//\alpha \iff r \cap \alpha = \emptyset.$$

45. Perpendicolarità e condizioni analitiche

Definizione 45.1. 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 $U^\perp \subset (W^\perp)^\perp$, ma $(W^\perp)^\perp = W$.

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. 40.4.

Esempio 45.2. 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 45.1 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 $\mathfrak{R} = (O, \{u_1, u_2, u_3\})$.

Osserviamo innanzitutto che dato un piano α di equazione:

$$\alpha : ax + by + cz + d = 0,$$

allora il vettore $n(a, b, c)$ di componenti (a, b, c), ovvero

$$n = au_1 + bu_2 + cu_3,$$

è sempre **perpendicolare** al piano in questione.

Infatti, sappiamo che la giacitura W del piano è descritta dall'equazione

$$W : ax + by + cz = 0,$$

che può riscriversi utilizzando il prodotto scalare:

$$W : (a, b, c) \cdot (x, y, z) = 0.$$

Quindi un vettore v di componenti (x, y, z) vi appartiene se e solo se $v \cdot n = 0$ (ciò è vero perchè la base $\mathfrak{B}$ è ortonormale, si ricordi la (33)).

Ciò premesso, ricaviamo i seguenti criteri:

Condizioni di parallelismo e perpendicolarità retta-piano

Consideriamo una retta r ed un piano α di equazioni:

$$\begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad \alpha : ax + by + cz + d = 0.$$

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 45.3. 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.

46. Sottospazio passante per punti assegnati

Spesso è utlile individuare un sottospazio affine imponendo il passaggio per un certo numero di punti assegnati; il seguente risultato descrive il più piccolo di tali sottospazi:

Teorema 46.1. *Siano $P_0, P_1, \cdots, P_k$ punti di $\mathbb{R}^n$, $k \geq 1$. Allora esiste il più piccolo (per inclusione) sottospazio affine S passante per tutti questi punti; esso è dato da*

$$S = P_0 + L(\overrightarrow{P_oP_1}, \ldots, \overrightarrow{P_oP_k}).$$

Si omette per brevità la dimostrazione, che è un esercizio.

Esempio 46.2. Per due punti distinti P_0 e P_1 di $\mathbb{R}^n$ passa una ed una sola retta.

Si tratta infatti della retta $r = P_0 + L(\overrightarrow{P_oP_1})$. Questo è uno degli assiomi della geometria Euclidea sintetica (in cui la nozione di retta è primitiva, e tale assioma, insieme ad altri, ne descrive le proprietà essenziali); nel modello $\mathbb{R}^n$ è una conseguenza della definizione di retta.

Dati i punti $P_0, \ldots, P_k$, diremo essi sono *allineati* se esiste una retta che passa per tutti.

Corollario 46.3. *Per tre punti non allineati P_0, P_1, P_2 passa uno ed un solo piano.*

Infatti, il più piccolo sottospazio che contiene i tre punti non può essere una retta per ipotesi; poichè esso è dato da

$$\alpha = P_0 + L(\overrightarrow{P_oP_1}, \overrightarrow{P_oP_2}) \tag{41}$$

deve aversi $\mathtt{dim}L(\overrightarrow{P_oP_1}, \overrightarrow{P_oP_2}) = 2$, ovvero si tratta di un piano. Qualsiasi piano β che passi per i tre punti dovrà contenere α e quindi coincidere con α.

In $\mathbb{R}^3$ possiamo scrivere un'equazione del piano in questione, note le coordinate dei punti in un fissato riferimento Cartesiano. Posto $P_0(x_o, y_o, z_0)$, $P_1(x_1, y_1, z_1)$ e $P_2(x_2, y_2, z_2)$, allora si ha

$$\alpha : \begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0.$$

Infatti, stante la (41), abbiamo che un punto $P(x, y, z)$ appartiene ad α se e solo se $\overrightarrow{P_oP}$ è combinazione lineare dei vettori $\overrightarrow{P_oP_1}$ e $\overrightarrow{P_oP_2}$. Il vettore $\overrightarrow{P_oP}$ ha componenti $(x - x_0, y - y_0, z - z_0)$ e analogamente per $\overrightarrow{P_oP_1}$ e $\overrightarrow{P_oP_2}$, e di qui si ottiene l'equazione di cui sopra.

Concludiamo questo paragrafo citando il seguente risultato, che generalizza il famoso *quinto postulato* o *postulato delle rette parallele* euclideo, che è uno degli assiomi principali nella trattazione sintetica della geometria piana:

Proposizione 46.4. *Dato un sottospazio affine S di $\mathbb{R}^n$, per ogni punto P passa uno ed un solo sottospazio T parallelo a S e tale che*

$$\mathtt{dim}(T) = \mathtt{dim}(S).$$

Il sottospazio in questione è infatti $T = P + W$ dove W è la giacitura di S.

47. Distanze e angoli

Mediante il prodotto scalare è possibile dare una definizione semplice di distanza tra punti:

Definizione 47.1. 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. Per discuterla, è opportuno premettere il seguente risultato valido in qualsiasi spazio vettoriale Euclideo:

Teorema 47.2. *(disuguaglianza di Cauchy-Schwarz). Siano u, v vettori arbitrari di uno spazio vettoriale Euclideo (V, g). Allora si ha*

$$|g(u, v)| \leq ||u||||v||. \tag{42}$$

Dimostrazione: La disuguaglianza da provare si può riformulare come:

$$g(u, v)^2 \leq ||u||^2||v||^2,$$

cioè

$$g(u, v)^2 \leq g(u, u)g(v, v). \tag{43}$$

Questa disuguaglianza è banale se $v = 0$, per cui assumeremo $v \neq 0$. Useremo il fatto che, essendo g definita positiva, per ogni λ, μ scalari si ha

$$g(\lambda u + \mu v, \lambda u + \mu v) \geq 0.$$

Per la bilinearità e simmetria di g, il primo membro si sviluppa come segue:

$$\lambda^2 g(u, u) + 2\lambda\mu g(u, v) + \mu^2 g(v, v) \geq 0.$$

Per ottenere la (43) basta scegliere $\lambda = g(v, v)$ e $\mu = -g(u, v)$: sostituendo infatti si ricava

$$g(v, v)^2 g(u, u) - 2g(v, v)g(u, v)^2 + g(u, v)^2 g(v, v) \geq 0$$

e quindi dividendo per la quantità positiva $g(v, v)$ si ha la (43). □

Corollario 47.3. *(Disuguaglianza triangolare) Per ogni u, v vettori di uno spazio vettoriale Euclideo (V, g) si ha:*

$$||u + v|| \leq ||u|| + ||v||. \tag{44}$$

La dimostrazione si lascia la lettore; è sufficiente provare che

$$||u + v||^2 \leq (||u|| + ||v||)^2.$$

Si utilizza la disuguaglianza di Cauchy-Schwarz unitamente alla seguente identità

$$||u + v||^2 = ||u||^2 + 2g(u, v) + ||v||^2. \tag{45}$$

Quest'ultima si ricava sviluppando $g(u + v, u + v)$ come fatto sopra.

La versione geometrica di questo risultato è:

Proposizione 47.4. *(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 47.5. 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 (32) 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.

La disuguaglianza di Cauchy-Schwarz permette di introdurre anche in modo analitico la nozione di angolo tra vettori:

Definizione 47.6. 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 (42), 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 $\lambda < 0$.

Date due rette r e s di $\mathbb{R}^n$ (incidenti o no), restano definiti, mediante questa nozione, due angoli tra esse: per definizione, ciascuno di essi è l'angolo formato da una coppia di vettori non nulli u e v paralleli alle due rette (se si cambia u in λu e/o v in μv, il coseno dell'angolo al più cambia segno).

Tornando alle distanze, in questo contesto il Teorema di Pitagora assume la seguente formulazione:

Teorema 47.7. *Siano A, B, C punti di $\mathbb{R}^n$. Allora*

$$d(B,C)^2 = d(A,B)^2 + d(A,C)^2 - 2g(\overrightarrow{AB}, \overrightarrow{AC}). \tag{46}$$

DIMOSTRAZIONE: Si ha infatti, utilizzando la (45):

$$\begin{aligned} d(B,C)^2 &= ||\overrightarrow{BC}||^2 = ||\overrightarrow{BA} + \overrightarrow{AC}||^2 = ||\overrightarrow{BA}||^2 + 2g(\overrightarrow{BA}, \overrightarrow{AC}) + ||\overrightarrow{AC}||^2 = \\ &= d(A,B)^2 - 2g(\overrightarrow{AB}, \overrightarrow{AC}) + d(A,C)^2. \end{aligned}$$

□

48. Distanza tra punti e sottospazi

Teorema 48.1. *Siano S un sottospazio affine di $\mathbb{R}^n$ e A un punto non appartenente a S. Allora esiste uno ed un solo punto $H \in S$ tale che il vettore $\overrightarrow{AH}$ è perpendicolare a S. Inoltre*

$$d(A,H) < d(A,P) \text{ per ogni punto } P \in S,\ P \neq H.$$

Il punto H in questo enunciato 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 tra A e il sottospazio** S, che si denota con $d(A,S)$.

Nel caso in cui $A \in S$ si pone invece $d(A,S) = 0$.

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 non nullo $\overrightarrow{P_0A}$ come

$$\overrightarrow{P_0A} = w + u$$

dove $w \in W$ e $u \in W^\perp$. In corrispondenza del vettore 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:

$$\overrightarrow{AH} = H - A = P_0 + w - A = w + P_0 - A = w + (-w - u) = -u \in W^\perp.$$

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 (46), poichè i vettori $\overrightarrow{AH}$ e $\overrightarrow{HP}$ sono perpendicolari (in quanto $\overrightarrow{HP} \in W$):

$$d(A,P)^2 = d(A,H)^2 + d(H,P)^2 > d(A,H)^2$$

essendo $d(H,P) > 0$. $\square$

Come caso rilevante, stabiliamo una formula per calcolare la distanza tra un punto ed un iperpiano:

Teorema 48.2. *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

$$d(A, \alpha) = \frac{|a_1x_1^0 + \cdots + a_nx_n^0 + d|}{\sqrt{a_1^2 + \cdots + a_n^2}}.$$

Dimostrazione: Si osservi che la formula in questione dà correttamente $d(A, \alpha) = 0$ nel caso in cui A sia un punto di S. Consideriamo quindi il caso $A \notin S$. Si tratta di esplicitare la distanza $P(A, H)$ dove il punto $H \in S$ è quello caratterizzato nel Teorema precedente; denotato con n il vettore di componenti $(a_1, \ldots, a_n)$ rispetto alla base del riferimento, come si è visto nel caso dei piani di $\mathbb{R}^3$ abbiamo che n è perpendicolare ad α. Dunque, essendo $\overrightarrow{AH}$ perpendicolare a S, il punto H è della forma

$$H = A + tn$$

per un opportuno scalare $t \in \mathbb{R}$, da determinarsi imponendo il passaggio di S per H. Poichè l'equazione di α si può reintepretare come

$$\alpha : \overrightarrow{OP} \cdot n + d = 0$$

dove $P(x_1, \ldots, x_n)$ è un punto generico di $\mathbb{R}^n$, abbiamo che H appartiene a S se e solo se

$$\overrightarrow{OH} \cdot n = -d$$

ovvero

$$\overrightarrow{OA} \cdot n + t\, n \cdot n = -d,$$

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 la formula enunciata. $\square$

49. Isometrie

Definizione 49.1. 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 49.2. *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

$$\tau_b(X) := X + b.$$

Dunque ogni isometria è una funzione composta

$$f = \tau_b \circ L_A$$

di una traslazione e di un'isometria lineare.

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).