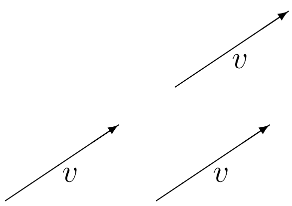
1 Spazi vettoriali

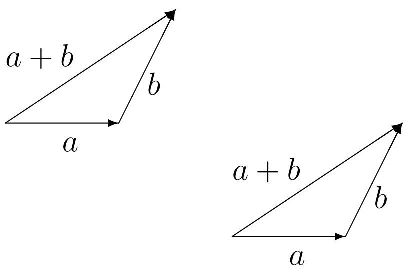
1.1 Spazi vettoriali geometrici $\mathcal{V}^n$ $(n = 1, 2, 3)$. Spazi vettoriali. Spazi vettoriali numerici $\mathbb{R}^n$

Spazi vettoriali geometrici $\mathcal{V}^n$ $(n = 1, 2, 3)$

Vettori. L'ambito del discorso è uno fra la retta, il piano, lo spazio Euclidei.
Un "vettore" è il complesso di una "lunghezza", una "direzione", un "verso". Indichiamo i vettori con lettere minuscole $a, b, \dots$ Si conviene che una lunghezza nulla, una qualsiasi direzione e verso diano un unico vettore, che si dice "vettore nullo" e si indica con $\underline{0}$; due vettori con la stessa lunghezza e direzione ma con versi diversi si dicono l'uno "opposto" dell'altro; indicato un vettore con v, si indica il suo opposto con $-v$; ovviamente $-(-v) = v$.
Un "segmento orientato" è un segmento con un verso di percorrenza, si rappresenta con una freccia che va dal primo all'ultimo estremo.
Ogni segmento orientato con estremi distinti dà una lunghezza, una direzione e un verso, quindi un vettore; un segmento orientato ridotto a un punto dà il vettore nullo; due segmenti orientati che differiscono solo per l'orientamento danno due vettori opposti.
Ogni vettore è dato da infiniti segmenti orientati: per ogni vettore v ed ogni punto, esiste un'unico segmento orientato che esce dal punto e dà v.



Somma di vettori; proprietà. Siano dati due vettori a, b. Dato un punto, dal punto esce un'unico segmento orientato che dà a, si ha così un secondo punto, dal punto esce un'unico segmento orientato che dà b, si ha così un terzo punto. Il segmento orientato dal primo al terzo punto dà un vettore, che si dice "vettore somma" di a e b e si indica con $a + b$.
Dati due diversi primi punti, si ottengono due diversi segmenti orientati, che però danno lo stesso vettore.

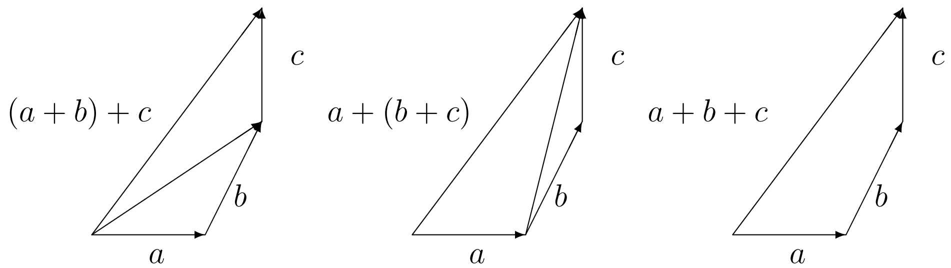


L'operazione di somma di vettori, come la somma di numeri reali, possiede le seguenti proprietà.

Associatività I due modi diversi per sommare tre vettori, lasciando invariato l'ordine, sono equivalenti:

(1) $\quad (a+b)+c = a+(b+c) \qquad$ per ogni a, b, c

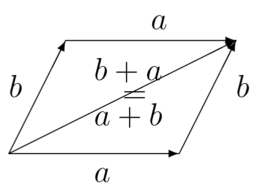
Si può dunque parlare di somma di tre vettori e scrivere la somma senza usare parentesi. $a+b+c$ si può ottenere direttamente come segue: da un punto, il vettore a dà un secondo punto, dal quale b dà un terzo punto, dal quale c dà un quarto punto, e il primo e l'ultimo punto danno il vettore $a+b+c$.



Commutatività. I due modi di sommare due vettori sono equivalenti:

(2) $\quad a+b = b+a \qquad$ per ogni a, b

Il valore comune delle due somme si può descrivere come segue. Dato un punto, dal punto escono due segmenti orientati che danno a e b, che danno due punti, dai quali escono due segmenti orientati che danno b ed a. I quattro segmenti orientati sono lati di un parallelogramma e la diagonale del parallelogramma dà il vettore somma $a+b = b+a$.



Vettore nullo. Il vettore nullo $\underline{0}$ possiede (ed è l'unico a possedere) la proprietà

(3) $\quad a+\underline{0} = a = \underline{0}+a \qquad$ per ogni a

Opposti. Per ogni vettore a, il vettore opposto $-a$ possiede (ed è l'unico a possedere) la proprietà

(4) $\quad a+(-a) = \underline{0} = (-a)+a.$

Come per i numeri reali, anche per i vettori queste proprietà permettono di risolvere le equazioni espresse in termini di somma:

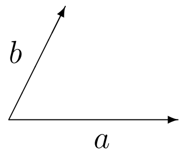
per ogni a, b l'equazione $x+a=b$ ha un'unica soluzione, $x = b+(-a)$;

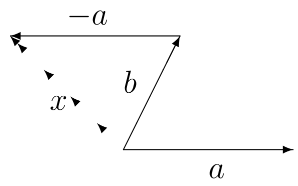
Infatti, ciascuna delle seguenti uguaglianze implica la successiva

$$x+a=b; \qquad x+a+(-a) = b+(-a); \qquad x+\underline{0} = b+(-a); \qquad x = b+(-a);$$

quindi, se c'è una soluzione, essa è $b + (-a)$.
Si lascia al lettore di verificare che $b + (-a)$ è una soluzione.

$b + (-a)$ si scrive in breve $b - a$ e si dice "differenza di b meno a".





Prodotto di numeri per vettori. Sia a un vettore. Si definisce il prodotto di un numero intero relativo per a ponendo

$n\, a = a + a + \cdots + a$ (n volte), per ogni $n = 1, 2, \dots,$

$0\, v = \overrightarrow{0}$;

$n\, a = (-a) + (-a) + \cdots + (-a)$ ($|n|$ volte), per ogni $n = -1, -2, \dots,$

Si definsce il prodotto di un numero razionale per a ponendo per ogni $n = 1, 2, \dots$

$$\frac{1}{n}\, a = (\text{unica soluzione dell'equazione } nx = a),$$

e per ogni $m = 0, \pm 1, \pm 2, \dots$ ed $n = 1, 2, \dots$

$$\frac{m}{n}\, a = m\, (\frac{1}{n}\, a).$$

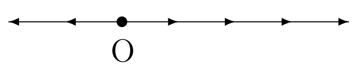
Ogni numero reale può essere dato da una successione di numeri razionali non negativi. Si definsce il prodotto di un numero reale per a mediante una successione di prodotti di numeri razionali per a. Ad esempio, $\sqrt{2} = 1.414\dots$ si può vedere come "limite" della successione di numeri razionali

$$1,\ 1.4,\ 1.41,\ 1.414,\ \dots$$

e si definisce $\sqrt{2}\, v$ come "limite" della successione di vettori

$1\, a,\ 1.4\, a,\ 1.41\, a,\ 1.414\, a,\ \dots$

Un'illustrazione dei vettori $\frac{m}{2}\, a$ per $-2 \leq m \leq 4$; tutti i vettori sono rappresentati da segmenti orientati uscenti dal punto O.



Proprietà. Le operazioni prodotto di numeri per numeri, somma di numeri per numeri, prodotto di numeri per vettori, somma di vettori sono legate dalla proprietà

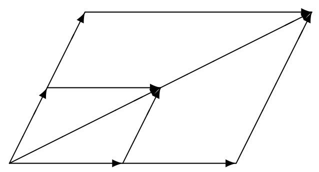
(5) $r\, (a + b) = r\, a + r\, b$;

(6) $(r + s)\, a = r\, a + s\, a$;

(7) $\quad (rs)\, a = r\,(s\, a)$;

(8) $\quad 1\, a = a$.

Un'illustrazione della (5):



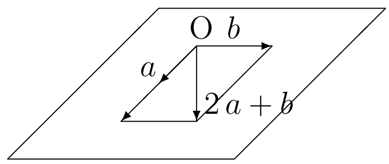
Indichiamo con $\mathcal{V}^1, \mathcal{V}^2, \mathcal{V}^3$ rispettivamente gli insiemi dei vettori della retta, piano e spazio Euclidei, con le operazioni di somma di vettori e di prodotto di numeri per vettori. Diciamo che questi insiemi con queste operazioni sono gli "spazi vettoriali geometrici".

Rappresentazione. Per "vedere" tutti i vettori, è utile rappresentarli con segmenti orientati uscenti da un punto fissato. Fissato un punto O, si ha che ogni segmento orientato uscente da O dà un vettore e ogni vettore è dato da un'unico segmento orientato uscente da O. Si ha così una corrispondenza biunivoca fra segmenti orientati uscenti da O e vettori. Per semplicità di linguaggio, a volte riferiremo termini e concetti dati per vettori anche a segmenti orientati e viceversa.

Dalle definizioni delle operazioni sui vettori si può dedurre che

sia dato un vettore $a \neq \underline{0}$; allora:
il segmento orientato a sta su un'unica retta per O;
per ogni numero r, il segmento orientato $r\, a$ sta sulla retta;
ogni segmento orientato sulla retta è del tipo $r\, a$ per qualche r;

siano dati due vettori a, b con direzioni diverse; allora:
i segmenti orientati a, b stanno su un'unico piano per O;
per ogni due numeri r, s il segmento orientato $r\, a + s\, b$ sta sul piano;
ogni segmento orientato sul piano è del tipo $r\, a + s\, b$ per qualche r, s.



dati tre segmenti orientati a, b, c che non stanno su un piano per O,
ogni segmento orientato dello spazio è del tipo $r\, a + s\, b + t\, c$ per qualche $r, s, t \in \mathbb{R}$.

Spazi vettoriali.

Uno "spazio vettoriale" reale è
- un insieme V di elementi qualsiasi detti "vettori" indicati con lettere minuscole $a, b, \ldots$, con un dato elemento detto "vettore nullo" indicato con $\underline{0}$ e per ogni vettore a un vettore "opposto" indicato con $-a$;
- un'operazione "somma", che ad ogni $a, b \in V$ associa un $a + b \in V$;
- un'operazione "prodotto esterno", che ad ogni $r \in \mathbb{R}$ e $a \in V$ associa un $r\,a \in V$;
tali che valgano le proprietà evidenziate per gli spazi vettoriali geometrici, cioè:

proprietà dell'operazione somma:

$$\begin{array}{rll} (1)\ \text{associativa} & (a+b)+c = a+(b+c) & \\ (2)\ \text{commutativa} & a+b = b+a & \\ (3)\ \text{del vettor nullo} & a+\underline{0} = a = \underline{0}+a & \\ (4)\ \text{degli opposti} & a+(-a) = \underline{0} = (-a)+a & \qquad\qquad (a,b,c \in V) \end{array}$$

proprietà di legame fra le operazioni

$$\begin{array}{rll} (5)\ \text{semidistributiva} & r\,(a+b) = r\,a + r\,b & \\ (6)\ \text{semidistributiva} & (r+s)\,a = r\,a + s\,a & \\ (7)\ \text{semiassociativa} & r\,(s\,a) = (rs)\,a & \\ (8)\ \text{del numero 1} & 1\,a = a & \qquad (a, b \in V,\ r, s, 1 \in \mathbb{R}). \end{array}$$

In particolare, per la proprietà associativa, ha senso scrivere una somma di tre o più vettori senza usare parentesi; la somma di sequenza di vettori $a_1, a_2, \ldots, a_m \in V$ si scrive $a_1 + a_2 + \ldots + a_m$ in breve

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

Il "multiplo" con coefficiente $r \in \mathbb{R}$ di un vettore a è il vettore $r\,a$; la "combinazione lineare" con coefficienti $r, s \in \mathbb{R}$ di due vettori a, b è il vettore

$$r\,a + s\,b;$$

in generale, la "combinazione lineare" con coefficienti $r_1, r_2, \ldots, r_m \in \mathbb{R}$ di m vettori $a_1, a_2, \ldots, a_m$ è il vettore

$$r_1\,a_1 + r_2\,a_2 + \ldots + r_m\,a_m = \sum_{i=1}^{m} r_i\,a_i.$$

Spazi vettoriali numerici $\mathbb{R}^n$

Sia n un intero positivo fissato. Una $n-$pla ordinata di numeri reali, in breve una $n-$pla di numeri reali, ha una prima, seconda, ..., $n-$ma componente numero reale. Due $n-$ple sono uguali se e solo se ciascuna componente dell'una è uguale alla rispettiva componente dell'altra. L'insieme delle $n-$ple di numeri reali si indica con $\mathbb{R}^n$.
Spesso, se si indica una $n-$pla con una lettera, allora si indicano le sue componenti con la stessa lettera con indici $1, \ldots, n$. Dunque per ogni due $n-$ple a, b,

$a = b$ se e solo se $a_i = b_i$ per ogni $i = 1, \dots, n$.

La $n-$pla con tutte le componenti uguali a 0 si dice $n-$pla nulla e si indica con $\underline{0}$:

$(\underline{0})_i = 0$ per ogni $i = 1, \dots, n$.

Prendendo gli opposti delle componenti di $n-$pla a, si ha la $n-$pla opposta di a, che si indica con $-a$:

$(-a)_i = -a_i$ per ogni $i = 1, \dots, n$.

Sommando le componenti di una $n-$pla a con le rispettive componenti di una $n-$pla b, si ottiene la $n-$pla somma di a e b, che si indica con $a + b$:

$(a + b)_i = a_i + b_i$ per ogni $i = 1, \dots, n$.

Moltiplicando un numero r per le componenti di una $n-$pla a si ottiene la $n-$pla prodotto di r per a, che si indica con $r\,a$:

$(r\,a)_i = r\,a_i$ per ogni $i = 1, \dots, n$.

Le operazioni sui numeri reali soddisfano le proprietà (1), ..., (8); le operazioni sulle $n-$ple sono definite componente per componente come le rispettive operazioni sui numeri reali; così le operazioni sulle $n-$ple soddisfano le proprietà (1), .., (8). Quindi l'insieme $\mathbb{R}^n$ delle $n-$ple di numeri reali, con queste operazioni di somma e di prodotto esterno, è uno spazio vettoriale.

Ad esempio, per ogni $r \in \mathbb{R}$, $a, b \in \mathbb{R}^n$,

$r\,(a + b) = r\,a + r\,b$ equivale a

$(r\,(a + b))_i = (r\,a + r\,b)_i$ per ogni i, equivale a

$r\,(a + b)_i = (r\,a)_i + (r\,b)_i$ per ogni i, equivale a

$r\,(a_i + b_i) = r\,a_i + r\,b_i$ per ogni i, che vale perchè in $\mathbb{R}$.

Identificazione di $\mathcal{V}^n$ ed $\mathbb{R}^n$.

$\mathcal{V}^1$ ed $\mathbb{R}$. Un "riferimento" per $\mathcal{V}^1$ è un vettore non nullo $i \neq \underline{0}$; ogni vettore v si scrive in un unico modo come multiplo di i

$v = v_1\,i$

si dice che v_1 è la "coordinata" di v rispetto ad i.

Associando a ciascun vettore la sua coordinata si ha un'applicazione biiettiva, che permette di identificare $\mathcal{V}^1$ con $\mathbb{R}$.

$\mathcal{V}^2$ ed $\mathbb{R}^2$. Un "riferimento" per $\mathcal{V}^2$ è dato da una coppia di vettori i, j aventi direzioni diverse; ogni vettore v si scrive in un unico modo come combinazuione lineare di i e j

$v = v_1\,i + v_2\,j$

si dice che la coppia (v_1, v_2) è la "coordinata" di v rispetto a i, j.

Si osservi che

$\underline{0}$ ha coordinata $(0, 0)$;

i ha coordinata $(1,0)$;

j ha coordinata $(0,1)$.

Associando a ciascun vettore la sua coordinata si ha un'applicazione biiettiva, che permette di identificare $\mathcal{V}^2$ con $\mathbb{R}^2$. Questa biiezione è compatibile con le operazioni di somma e di prodotto esterno su $\mathcal{V}^2$ ed $\mathbb{R}^2$, nel senso che per ogni v', v'', v si ha

$v' + v'' = v$ in $\mathcal{V}^2$ se e solo se $v' + v'' = v$ in $\mathbb{R}^2$.

Infatti, posto

$v' = v'_1 i + v'_2 j$

$v'' = v''_1 i + v''_2 j$

$v = v_1 i + v_2 j,$

si ha:

$v' + v'' = v$ in $\mathcal{V}^2$ significa $v'_1 i + v'_2 j + v''_1 i + v''_2 j = v_1 i + v_2 j$, che equivale a

$v'_1 + v''_1 = v_1$ e $v'_2 + v''_2 = v_2$ che equivale a

$(v'_1, v'_2) + (v''_1, v''_2) = (v_1, v_2)$ che significa $v' + v'' = v$ in $\mathbb{R}^2$.

$\mathcal{V}^3$ **ed** $\mathbb{R}^3$. Un "riferimento" per $\mathcal{V}^3$ è dato da una terna di vettori i, j, k rappresentati da segmenti orientati uscenti da un punto O, non complanari; ogni vettore v si scrive in un unico modo come combinazione lineare di i, j, k

$v = v_1 i + v_2 j + v_3 k$

si dice che la terna (v_1, v_2, v_3) è la "coordinata" di v rispetto a i, j, k.

Si osservi che

$\underline{0}$ ha coordinata $(0,0,0)$;

i ha coordinata $(1,0,0)$;

j ha coordinata $(0,1,0)$;

k ha coordinata $(0,0,1)$.

Associando a ciascun vettore la sua coordinata si ha un'applicazione biiettiva, che permette di identificare $\mathcal{V}^3$ con $\mathbb{R}^3$. Questa biiezione è compatibile con le operazioni di somma e di prodotto esterno su $\mathcal{V}^3$ ed $\mathbb{R}^3$.

Possiamo dunque tradurre problemi e risoluzioni espressi in termini di vettori geometrici in problemi e risoluzioni espressi in termini di vettori numerici.

Esempio. Fissato in $\mathcal{V}^2$ un riferimento i, j, consideriamo i vettori

$a = (3,1)$, $b = (1,2)$, $c = (2,2)$.

Ci chiediamo se c è combinazione di a e b.

L'equazione

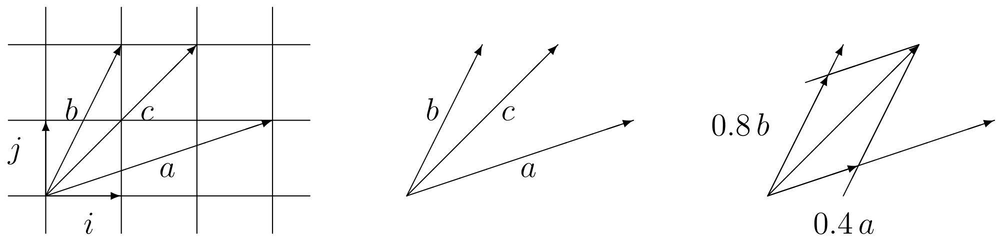
$x\,(3,1) + y\,(1,2) = (2,2)$

equivale al sistema di equazioni

$$\begin{cases} 3\,x + y = 2 \\ x + 2\,y = 2 \end{cases}$$

che ha soluzione $x = \frac{2}{5}$, $y = \frac{4}{5}$.

Un'illustrazione del corrispettivo geometrico del problema e del risultato:



Esempio. Il sistema di 3 uguaglianze fra numeri reali

$$\begin{cases} 2x + 3y = p_1 \\ 4x + 5y = p_2 \\ 6x + 7y = p_3 \end{cases}$$

equivale all'unica uguagliaza fra terne di numeri reali

$$(2x + 3y, 4x + 5y, 6x + 7y) = (p_1, p_2, p_3)$$

e può essere riscritto come

$$x(2, 4, 6) + y(3, 5, 7) = (p_1, p_2, p_3);$$

osserviamo che $(2, 4, 6)$ e $(3, 5, 7)$ non sono uno multiplo reale dell'altro.

Fissato un riferimento per $\mathcal{V}^3$, e identificati $a = (2, 4, 6)$, $b = (3, 5, 7)$, $p = (p_1, p_2, p_3)$ con vettori in $\mathcal{V}^3$, possiamo vedere che: (1) al variare di x, y in $\mathbb{R}$ il 1° membro dà i vettori su un piano; (2) il sistema delle 3 equazioni nelle incognite x, y ha qualche soluzione se e solo se il piano per a, b contiene il vettore p; (3) in tal caso la soluzione è unica.

Riassunto, notazioni, approfondimenti

Abbiamo considerato le seguenti strutture:

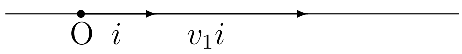
- spazi vettoriali $\mathcal{V}^n$ $(n = 1, 2, 3)$ dei vettori nella retta, piano, spazio Euclidei;

- spazi vettoriali $\mathbb{R}^n$ di $n-$ple $(n = 1, 2, 3, 4, \ldots)$;

ed abbiamo dato la definizione di

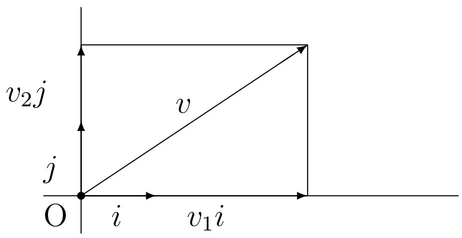
- spazio vettoriale V, come insieme con operazioni che soddisfano certe proprietà.

$\mathcal{V}^n$ **ed** $\mathbb{R}^n$. Gli spazi vettoriali geometrici $\mathcal{V}^n$, fissato un riferimento, possono essere identificati coi corrispettivi spazi vettoriali numerici $\mathbb{R}^n$ $(n = 1, 2, 3)$ come segue:



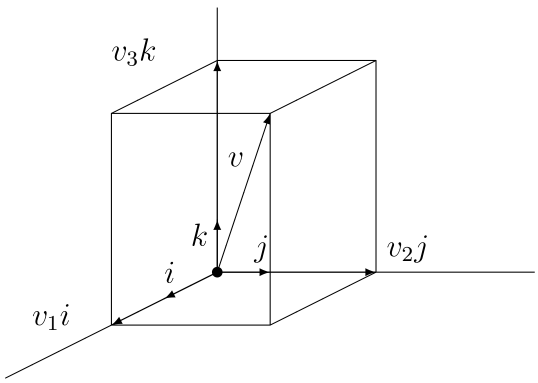
$\mathcal{V}^1 \ni v_1 i \sim v_1 \in \mathbb{R}$

(vettori come segmenti orientati uscenti da un punto fissato O; riferimento: un vettore $i \neq \underline{0}$)



$\mathcal{V}^2 \ni v = v_1 i + v_2 j \sim (v_1, v_2) \in \mathbb{R}^2$

(vettori come segmenti orientati uscenti da un punto fissato O; riferimento: una coppia di vettori i, j non allineati)



$\mathcal{V}^3 \ni v = v_1 i + v_2 j + v_3 k \sim (v_1, v_2, v_3) \in \mathbb{R}^3$

(vettori segmenti orientati uscenti da un punto fissato; riferimento: una terna di vettori i, j, k non complanari)

$\mathbb{R}^n$, **notazione.** Per il calcolo, la usuale scrittura in riga di coppie ordinate è meno agevole della scrittura come colonne; ad esempio:

$$(1, 2) + (3, 4) = (3, 6); \qquad \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

$$3\,(1,2) = (3,6); \quad 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Per questa ragione, di regola scriveremo le coppie, terne, ... $n-$ple come colonne

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad \ldots$$

Spazi vettoriali, regole di calcolo. Di seguito mostriamo come dalle proprietà che caratterizzano le operazioni in uno spazio vettoriale V seguano alcune proprietà familiari dal calcolo letterale. Di seguito, a, b, c indicano vettori in V e r un numero in $\mathbb{R}$.

- Legge di cancellazione:

$a + c = b + c$ implica $a = b$;

infatti, ciascuna delle seguenti uguaglianze implica la successiva (si lascia al lettore di motivare ciascun passaggio):

$a + c = b + c$;

$(a + c) + (-c) = (b + c) + (-c)$;

$a + (c + (-c)) = b + (c + (-c))$;

$a + \underline{0} = b + \underline{0}$;

$a = b$.

- Legge di annullamento del prodotto esterno:

$r\,a = \underline{0}$ se e solo se $r = 0$ oppure $a = \underline{0}$.

Infatti, da una parte si ha:

$0\,a + 0\,a = \underline{0} + 0\,a$ (si lascia al lettore giustificarla) e per la legge di cancellazione $0\,a = \underline{0}$;

in modo analogo si prova che $r\,\underline{0} = \underline{0}$.

Dall'altra parte, ciascuna delle seguenti uguaglianze implica la successiva (si lascia al lettore di motivare i passaggi)

$r\,a = \underline{0}$ e $r \neq 0$,

$\frac{1}{r}\,(r\,a) = \frac{1}{r}\,\underline{0}$,

$(\frac{1}{r}\,r)\,a) = \underline{0}$,

$1\,a = \underline{0}$

$a = \underline{0}$.

1.2 Dipendenza/Indipendenza lineare

Definizione, primi esempi, proposizioni.

Premessa. I vettori di $\mathbb{R}^2$

$$a = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad c = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

soddisfano la relazione

c è combinazione lineare, con coefficienti $\frac{2}{5}, \frac{4}{5}$, di a, b:

$\frac{2}{5}\, a + \frac{4}{5}\, b = c$;

questa relazione equivale alla relazione

$2\, a + 4\, b - 5\, c = \underline{0}$, cioè

$\underline{0}$ è combinazione lineare, con coefficienti $2, 4, -5$ di a, b, c.

Da questa relazione si deduce che anche b, c sono combinazione lineare degli altri:

$b = -\frac{2}{4}\, a + \frac{5}{4}\, c$

$a = -\frac{4}{2}\, b + \frac{5}{2}\, c.$

Definizione. Una sequenza $a_1, a_2, \dots, a_m$ di vettori di uno spazio vettoriale V è linearmente dipendente se l'uguaglianza

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = \underline{0}$$

vale per qualche $\alpha_1, \alpha_2, \dots, \alpha_m$ non tutti nulli.

Si dice che una sequenza di vettori è linearmente indipendente se non è linearmente dipendente, cioè:

Definizione. Una sequenza $a_1, a_2, \dots, a_m$ di vettori di uno spazio vettoriale V è linearmente indipendente se l'uguaglianza

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = \underline{0}$$

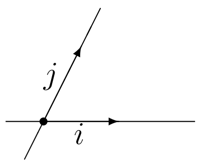
vale solo per $\alpha_1, \alpha_2, \dots, \alpha_m$ tutti nulli.

Esempi

(0) I vettori a, b, c di $\mathbb{R}^2$ considerati sopra sono linearmente dipendenti.

(1) In $\mathcal{V}^1$, un qualsiasi vettore $i \neq \underline{0}$ è linearmente indipendente; infatti, l'uguaglianza $\alpha i = \underline{0}$ vale solo per $\alpha = 0$.

(2) In $\mathcal{V}^2$:
un qualsiasi vettore $i \neq \underline{0}$ è linearmente indipendente;
una coppia di vettori i, j non allineati è linearmente indipendente; ce ne si può convincere da un punto di vista geometrico oppure da un punto di vista algebrico;



nel secondo modo: identifichiamo $\mathcal{V}^2$ con $\mathbb{R}^2$ usando i, j come riferimento, allora la relazione

$$\alpha\, i + \beta\, j = \underline{0}$$

equivale alla relazione

$$\alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

dalla quale, uguagliando le 1° componenti si ha $\alpha\, 1 + \beta\, 0 = 0$ cioè $\alpha = 0$ e uguagliando le 2° componenti si ha $\alpha\, 0 + \beta\, 1 = 0$ cioè $\beta = 0$.
aggiungendo ad i, j un qualsiasi vettore v si ha una sequenza i, j, v linearmente dipendente; infatti vale una relazione

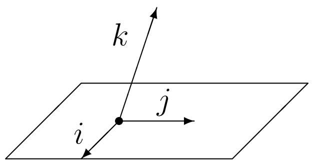
$$v_1\, i + v_2\, j = v$$

da cui

$$v_1\, i + v_2\, j - v = \underline{0}$$

e $v_1, v_2, -1$ non sono tutti nulli.

(3) In $\mathcal{V}^3$:
un qualsiasi vettore $i \neq \underline{0}$ è linearmente indipendente;
una coppia di vettori i, j non allineati è linearmente indipendente;
una terna di vettori i, j, k non complanari è linearmente indipendente;



ce ne si può convincere da un punto di vista geometrico oppure da un punto di vista algebrico; nel secondo modo: identifichiamo $\mathcal{V}^3$ con $\mathbb{R}^2$ usando i, j, k come riferimento, allora la relazione

$$\alpha\, i + \beta\, j + \gamma\, k = \underline{0}$$

equivale alla relazione

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

dalla quale, uguagliando le 1°, 2°, 3° componenti si ha rispettivamente $\alpha = 0$, $\beta = 0$, $\gamma = 0$.

(4) In $\mathbb{R}^n$:
la sequenza $e_1, e_2, \ldots, e_n$ dei vettori unità, definiti da

$$e_i = n\text{−pla con } i\text{−ma componente 1 e altre componenti 0,}$$

è linearmente indipendente;
ogni sottosequenza della sequenza $e_1, e_2, \ldots, e_n$ è linearmente indipendente.

(6) In uno spazio vettoriale V, consideriamo un vettore v;
se $v \neq \underline{0}$, allora $\alpha\, v = \underline{0}$ implica $\alpha = 0$ quindi v è linearmente indipendente;
$1\, \underline{0} = \underline{0}$, quindi $\underline{0}$ è linearmente dipendente.

Proposizione. Per ciascuna sequenza $a_1, a_2, \ldots, a_m$ di vettori di uno spazio vettoriale V, sono equivalenti:

(1) $a_1, a_2, \ldots, a_m$ è linearmente dipendente;

(2) $a_1 = \underline{0}$ oppure esiste un a_i $(i \geq 2)$ che è combinazione lineare dei precedenti.

1 *Dimostrazione.*

(1) implica (2). Per ipotesi, vale una relazione

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_m a_m = \underline{0}$$

con $\alpha_1, \alpha_2, \ldots, \alpha_m$ non tutti nulli; sia i massimo tale che $\alpha_i \neq 0$; allora:
se $i = 1$, si ha $\alpha_1 a_1 + 0a_2 + \cdots + 0a_m = \underline{0}$ cioè $\alpha_1 a_1 = \underline{0}$ con $\alpha_1 \neq 0$, da cui $a_1 = \underline{0}$.
se $i \geq 2$, si ha

$$\alpha_1 a_1 + \cdots + \alpha_i a_i + 0\, a_{i+1} + \cdots + 0\, a_m = \underline{0}$$

cioè

$$\alpha_1 a_1 + \cdots + \alpha_i a_i = \underline{0} \text{ con } \alpha_i \neq 0,$$

da cui

$$a_i = -\tfrac{\alpha_1}{\alpha_i}\, a_1 - \cdots - \tfrac{\alpha_{i-1}}{\alpha_i}\, a_{i-1}.$$

(2) implica (1).
Se $a_1 = \underline{0}$, allora vale $1\, a_1 + 0\, a_2 + \cdots + 0\, a_m = \underline{0}$, quindi $a_1, \ldots, a_m$ è linearmente dipendente.
Se esiste un a_i $(i \geq 2)$ tale che $a_i = \alpha_1\, a_1 + \cdots + \alpha_{i-1}\, a_{i-1}$, allora

$$\alpha_1\, a_1 + \cdots + \alpha_{i-1}\, a_{i-1} - a_i + 0\, a_{i+1} + \cdots + 0\, a_m = \underline{0},$$

quindi $a_1, \ldots, a_m$ è linearmente dipendente ■

Applicazione. In $\mathcal{V}^2$, ogni sequenza $a_1, \ldots, a_m$ di almeno 3 vettori è linearmente dipendente. Infatti:
se $a_1 = \underline{0}$, allora la sequenza è linearmente dipendente;
se $a_1 \neq \underline{0}$ e a_2 sta sulla retta di a_1, allora $a_2 = \alpha\, a_1$, quindi la sequenza è linearmente dipendente;
se a_1 e a_2 non sono allineati, allora $a_3 = \alpha_1\, a_1 + \alpha_2\, a_2$ (cfr. esempio 2) quindi la sequenza è linearmente dipendente.

In termini di indipendenza lineare, sui ha:

Proposizione. Per ciascuna sequenza $a_1, a_2, \ldots, a_m$ di vettori di uno spazio vettoriale V, sono equivalenti:

(1) $a_1, a_2, \ldots, a_m$ è linearmente indipendente;

(2) $a_1 \neq \underline{0}$ e nessun a_i $(i \geq 2)$ è combinazione lineare dei precedenti.

Due proprietà

1- Due sequenze di vettori che differiscono solo per l'ordine dei vettori, sono entrambe linearmente dipendenti o entrambe linearmente indipendenti (deriva dalla commutatività dell'operazione di somma di vettori).

2- Se una sequenza di vettori è linearmente indipendente, allora anche ogni sua sottosequenza lo è. (Idea: sia $a, b, c, \ldots$ linearmente indipendente; l'uguaglianza $\beta, b + \gamma, c + \cdots = \underline{0}$ può essere riscritta $0\, a + \beta, b + \gamma, c + \cdots = \underline{0}$, dunque si ha $0 = \beta = \gamma = \cdots = 0$; abbiamo provato che $b, c, \ldots$ è linearmente indipendente.)

Una procedura per vettori in $\mathbb{R}^n$

Descriviamo una procedura per risolvere in modo tendenzialmente efficiente il problema di stabilire la proprietà di dipendenza/indipendenza lineare di una sequenza di vettori in uno spazio vettoriale $\mathbb{R}^n$. La procedura consiste nell'applicazione di operazioni sulla sequenza di vettori che lasciano invariata la proprietà e permettono di passare da una sequenza ad una con un vettore in meno, avvicinandosi così al caso di un solo vettore, per il quale il problema ha una soluzione ovvia.

Di regola, consideriamo ciascun vettore di $\mathbb{R}^n$ come una colonna con n componenti e ciascuna sequenza $a, b, \ldots, d$ di m vettori di $\mathbb{R}^n$ come una matrice con n righe ed m colonne, e rappresentiamo la matrice come una tabella

$$\begin{array}{cccc} a & b & \ldots & d \\ \hline a_1 & b_1 & & d_1 \\ \vdots & \vdots & & \vdots \\ a_n & b_n & & d_n \end{array}$$

Fatti.

1- Per un singolo vettore v, si ha: v è linearmente dipendente o indipendente secondo che v sia uguale o diverso da $\underline{0}$ (vero in ogni spazio vettoriale);

2- C'è un caso in cui si può cancellare un vettore, lasciando invariata la proprietà.

> Se $a, b, c, \ldots$ sono vettori in $\mathbb{R}^n$ tali che in una certa componente, a è non 0, mentre $b, c, \ldots$ sono 0, allora $a, b, c, \ldots$ è linearmente indipendente se e solo se $b, c, \ldots$ è linearmente indipendente.

Un esempio: una sequenza di 3 vettori a, b, c di $\mathbb{R}^4$ del tipo

$$\begin{array}{ccc} a & b & c \\ \hline * & * & * \\ \blacklozenge & 0 & 0 \\ * & * & * \\ * & * & * \end{array}$$

dove il simbolo $\blacklozenge$ indica un numero $\neq 0$ e i simboli $*$ indicano numeri, è linearmente indipendente se e solo se la sequenza dei 2 vettori b, c

b	c
$*$	$*$
0	0
$*$	$*$
$*$	$*$

è linearmente indipendente.

2 Dimostrazione. Se $a, b, c, \ldots$ è linearmente indipendente, allora anche $b, c, \ldots$ lo è. Proviamo che se $b, c, \ldots$ è linearmente indipendente, allora anche $a, b, c, \ldots$ lo è. La uguaglianza

$$\alpha\, a + \beta\, b + \gamma\, c + \cdots = \underline{0}$$

equivale a n uguaglianze in $\mathbb{R}$, una per ogni componente. Per ipotesi, c'è una componente nella quale a è un certo numero $\blacklozenge \neq 0$, mentre $b, c, \ldots$ sono 0; in tale componente si ha

$$\alpha\, \blacklozenge + \beta\, 0 + \gamma\, 0 + \cdots = 0, \qquad \text{da cui} \qquad \alpha = 0;$$

dunque $(*)$ equivale a

$$\alpha = 0 \qquad \beta\, b + \gamma\, c + \cdots = \underline{0}.$$

Essendo $b, c, \ldots$ linearmente indipendente, si ha $\beta = \gamma = \cdots = 0$. Dunque $a, b, c, \ldots$ è linearmente indipendente.

3- Ci sono delle operazioni su una sequenza di vettori che lasciano invariata la proprietà di indipendenza lineare e possono essere usate per semplificare la sequenza. In particolare, si ha:

> Per ogni sequenza di vettori di uno spazio vettoriale, l'operazione di sommare un multiplo di un vettore a un altro vettore lascia invariata la proprietà di indipendenza lineare della sequenza.

Dimostrazione. Proviamo che per ogni sequenza $a, b, c, \ldots$ di vettori ed ogni $r \in \mathbb{R}$, l'operazione di sommare r volte un vettore a un altro vettore, per semplicità il 1° al 2°, lascia invariata l'indipendenza lineare.

Supponiamo che $a, b, c, \ldots$ sia linearmente indipendente e proviamo che $a, b + ra, c, \ldots$ lo è.

$$x\, a + y\,(b + ra) + z\, c + \ldots = \underline{0} \text{ equivale a}$$

$$(x + yr)\, a + y\, b + z\, c + \ldots = \underline{0},$$

implica

$$x + yr = y = z = \cdots = 0, \text{ da cui}$$

$$x = y = z = \cdots = 0.$$

Supponiamo che $a, b + ra, c, \ldots$ sia linearmente indipendente e proviamo che $a, b, c, \ldots$ lo è. Basta osservare che la seconda sequenza si ottiene dalla prima sommando $-r$ volte il 1° vettore al 2°.

Procedura *Esempio 1.*

a	b	c	d
1	1	3	1
1	2	5	2
1	3	7	4
1	4	9	8
1	5	11	16

a	b	c	d
1	0	0	0
1	1	2	1
1	2	4	3
1	3	6	7
1	4	8	15

b	c	d
0	0	0
1	2	1
2	4	3
3	6	7
4	8	15

b	c	d
0	0	0
1	0	0
2	0	1
3	0	4
4	0	11

c	d
0	0
0	0
0	1
0	4
0	11

(qui i simboli a, b, c, d vanno intesi come variabili, il cui valore varia durante la procedura) Conclusione:

i vettori iniziali a, b, c, d sono linearmentre dipendenti.

1° passo: $a_1 \neq 0$, annullate b_1, c_1, d_1 sommando $-a$, $-3\,a$, $-a$ a b, c, d;
2° passo: cancellato a;
3° passo: $b_2 \neq 0$, annullate c_2, d_2 sommando $-2\,b$, $-b$ a c, d;
4° passo: cancellato b;
5° passo: riconosciuto che $c = \underline{0}, d$ sono linearmente dipendenti.

Commento. Qui si è illustrata la procedura, così come l'avrebbe applicata una macchina; in realtà, dopo il 1° passo si sarebbe potuto osservare che $c = 2\,b$ quindi che i vettori a, b, c, d dopo il 1° passo son linearmente dipendenti, e dedurre che i vettori a, b, c, d iniziali lo sono.

Esempio 2.

a	b	c
1	1	1
1	0	0
0	1	0
0	0	1

b	c
1	1
0	0
1	0
0	1

c
1
0
0
1

Conclusione:

i vettori iniziali a, b, c sono linearmentre indipendenti.

1° passo: visto che $a_2 \neq 0$ e $b_2 = c_2 = 0$, cancellato a;
2° passo: visto che $b_3 \neq 0$ e $c_3 = 0$, cancellato b;
3° passo: riconosciuto che $c \neq \underline{0}$ è linearmente indipendente.

Procedura, Pseudocodice (grezzo)

Input: una sequenza di vettori in un $\mathbb{R}^n$;
Output: "lin. dip." o "lin. indip."

Leggi 1° vettore;
Casi
Se 1° vett $= \underline{0}$, scrivi "lin. dip.";
Se 1° vett $\neq \underline{0}$ ed è il solo vettore della sequenza, scrivi "lin. indip."
Altrimenti, scegli una componente non 0 del 1° vettore, annulla* quella componente negli altri vettori, cancella il 1° vettore.
Fine Casi

Ripeti

*tramite operazioni elementari.

$m > n$ **vettori in** $\mathbb{R}^n$.

Geometricamente, si "vede" che:
2 o più vettori in $\mathcal{V}^1 \simeq \mathbb{R}$ sono linearmente dipendenti;
3 o più vettori in $\mathcal{V}^2 \simeq \mathbb{R}^2$ sono linearmente dipendenti;
4 o più vettori in $\mathcal{V}^3 \simeq \mathbb{R}^3$ sono linearmente dipendenti.

In generale, si ha

2 *Teorema.* Ogni sequenza di $m > n$ vettori di $\mathbb{R}^n$ è linearmente dipendente.

Dimostrazione. Siano $a_1, \ldots, a_{n+1}, \ldots$ vettori in $\mathbb{R}^n$. Li consideriamo come colonne di una matrice e li rappresentiamo con una tabella

$$\begin{array}{ccccc} a_1 & \ldots & a_n & a_{n+1} & \ldots \\ \hline * & \ldots & * & * & \ldots \\ \vdots & & \vdots & \vdots & \\ * & \ldots & * & * & \ldots \end{array}$$

Applichiamo la procedura.
Se $a_1 = \underline{0}$, allora i vett. sono lin dip.; altrimenti, a_1 ha una componente $\neq 0$, supponiamo per semplicità la 1°, con operazioni elementari la annulliamo negli altri vettori.
Se $a_2 = \underline{0}$, allora i vett. sono lin dip.; altrimenti, a_2 ha una componente $\neq 0$, supponiamo per semplicità la 2°, con operazioni elementari la annulliamo negli altri vettori.
...
Se $a_n = \underline{0}$, allora i vett. sono lin dip.; altrimenti, a_n ha una componente $\neq 0$, che deve essere la n°, con operazioni elementari la annulliamo negli altri vettori.
A questo punto si ha una sequenza del tipo

$$\begin{array}{ccccc} a_1 & \ldots & 0 & a_{n+1} & \ldots \\ \hline \blacklozenge & \ldots & 0 & 0 & \ldots \\ \vdots & \ddots & \vdots & \vdots & \\ * & \ldots & \blacklozenge & 0 & \ldots \end{array}$$

con $a_{n+1} = \underline{0}$. Concludiuoamo che i vettori iniziali $a_1, \ldots, a_{n+1}, \ldots$ sono linearmente dipendenti.

1.3 Dimensione, basi, coordinate

Altri esempi di spazi vettoriali.

- Spazio vettoriale nullo. Uno spazio vettoriale deve contenere almeno un vettore: il vettore nullo $\underline{0}$. Per l'insieme $\{\underline{0}\}$ si ha un'unica possibile operazione di somma di vettori e un'unica possibile operazione di prodotto di numeri per vettori

$$\underline{0} + \underline{0} = \underline{0},$$

$$r\,\underline{0} = \underline{0}, \quad \forall r \in \mathbb{R},$$

e queste operazioni soddisfano tutte le proprietà (1), ..., (8). Quindi, c'è uno ed un solo spazio vettoriale con un solo elemento: l'insieme $\{\underline{0}\}$ con le operazioni ovvie sopra scritte, lo si dice "spazio vettoriale nullo". Geometricamente, lo spazio vettoriale nullo può essere pensato come l'insieme dei vettori che stanno in un punto, e può essere indicato con $\mathcal{V}^0$.

- Spazio vettoriale delle successioni. Una successione a di numeri reali ha una componente numero reale a_i per ogni intero positivo $i = 1, 2, \ldots$; due successioni sono uguali se e solo se ciascuna componente dell'una è uguale alla rispettiva componente dell'altra. L'insieme delle successioni di numeri reali si indica con $\mathbb{R}^\infty$.
Così come si sono definite la $n-$pla nulla, l'opposta di una $n-$pla e le operazioni di somma di $n-$ e di prodotto di numeri per $n-$ple, analogamente si definiscono la successione nulla, l'opposta di una successione, le operazioni di somma di successioni e di prodotto di numeri per successioni e si prova che l'insieme $\mathbb{R}^\infty$ con queste operazioni è uno spazio vettoriale.
In $\mathbb{R}^\infty$ si hanno infinite successioni unità

$$e_1 = (1, 0, 0, \ldots)$$

$$e_2 = (0, 1, 0, \ldots)$$

$$\vdots$$

Dimensione.

Possiamo distinguere gli spazi vettoriali geometrici $\mathcal{V}^n$ senza usare termini geometrici come segmenti orientati, punto, retta, piano e spazio, usando invece solo termini del linguaggio degli spazi vettoriali, specialmente l'indipendenza lineare:

$\mathcal{V}^0$: non c'è alcun vettore indipendente;
$\mathcal{V}^1$: c'è un, ma non più di un, vettore indipendente;
$\mathcal{V}^2$: ci sono due, ma non più di due, vettori indipendenti;
$\mathcal{V}^3$: ci sono tre, ma non più di tre, vettori indipendenti.

Per convenzione, diciamo che

una sequenza vuota di vettori è indipendente.

Al posto di dire "numero dei termini" di una sequenza, diciamo in breve "lunghezza" di una sequenza. Dunque: in $\mathcal{V}^n$ la massima lunghezza di una sequenza di vettori indipendente è n. Ciò suggerisce la seguente

Definizione. Si dice che uno spazio vettoriale V ha "dimensione finita" se l'insieme delle lunghezze delle sequenze di vettori indipendenti in V possiede massimo; tale massimo si dice "dimensione" di V e si indica con $\dim(V)$.

Esempi

- Per ciascun ente geometrico, la dimensione dello spazio vettoriale sull'ente coincide con quella dell'ente: $\dim(\mathcal{V}^n) = n$.

- Per ogni intero positivo n fissato, in $\mathbb{R}^n$ c'è una sequenza di n vettori indipendente, ad esempio quella dei vettori unità $e_1, e_2, \dots, e_n$, ma ogni sequenza di $m > n$ vettori è dipendente, quindi $\dim(\mathbb{R}^n) = n$.

- Nello spazio vettoriale $\mathbb{R}^\infty$, per ogni intero positivo n c'e una sequenza di n vettori indipendente, quindi $\mathbb{R}^\infty$ non ha dimensione finita.

Basi, coordinate.

In ciascun spazio vettoriale geometrico si può fissare una sequenza di vettori di riferimento e tramite di essa assegnare ad ogni vettore una sequenza coordinata di numeri. Ciò suggerisce la seguente

Definizione. Si dice che una sequenza $b_1, \dots, b_n$ di vettori di uno spazio vettoriale V è una "base" di V se ogni $v \in V$ si scrive in un unico modo come combinazione lineare

$$v = v_1\, b_1 + \cdots + v_n\, b_n;$$

la $n-$pla dei coefficienti si dice "coordinata" di v rispetto alla base.

Esempi

- Lo spazio vettoriale nullo $\{\underline{0}\}$ ha un'unica base: l'insieme vuoto $\emptyset$; la coordinata di $\underline{0}$ rispetto alla base è la sequenza vuota di numeri reali.

- In ciascuno spazio vettoriale geometrico $\mathcal{V}^n$, ciascuna sequenza di n vettori riferimento è una base e per ciascun vettore la $n-$pla coordinata in senso geometrico coincide con la $n-$pla coordinata in senso algebrico.

- In $\mathbb{R}^2$ vale l'identità

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

cioè: ogni coppia $v \in \mathbb{R}^2$ si scrive come combinazione lineare delle due coppie unità e_1, e_2 e i coefficienti sono le componenti di v, così la scrittura è unica; e_1, e_2 è una base, detta "base canonica", di $\mathbb{R}^2$ e le coordinate di ciascun $v \in \mathbb{R}^2$ rispetto ad essa sono le componenti di v.

- Identificando $\mathbb{R}^2$ con $\mathcal{V}^2$ si vede che $\mathbb{R}^2$ ha infinite basi.

- Sia n un intero positivo fissato. In $\mathbb{R}^n$ vale l'identità

$$v = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n$$

(si lascia al lettore di provarla), cioè: ogni $v \in \mathbb{R}^n$ si scrive come combinazione lineare delle $n-$ple unità $e_1, e_2, \dots, e_n$ e i coefficienti sono le componenti $v_1, v_2, \dots, v_n$ di v, così la scrittura è unica; $e_1, e_2, \ldots, e_n$ è una base, detta "base canonica", di $\mathbb{R}^n$ e le coordinate di ciascun $v \in \mathbb{R}^n$ rispetto ad essa sono le componenti di v.

- Si prova che non è possibile scrivere ogni successione di numeri reali come combinazione lineare di un numero finito di successioni fissate una volta per tutte, dunque lo spazio vettoriale $\mathbb{R}^\infty$ delle successioni non ha alcuna base.

1 *Proposizione.* Se $b_1, ..., b_n$ è una base di uno spazio vettoriale V allora $b_1, ..., b_n$ è linearmente indipendente.

Infatti, essendo $b_1, \ldots, b_n$ una base di V, l'uguaglianza

$$\beta_1 \, b_1 + \cdots + \beta_n \, b_n = \underline{0}$$

deve valere per un'unica sequenza di coefficienti, dunque $\beta_1 = \ldots, \beta_n = 0$, e $b_1, \ldots, b_n$ è indipendente.

Poichè ogni sequenza di $m > n$ vettori in $\mathbb{R}^n$ è dipendente, si ha

Ogni base di $\mathbb{R}^n$ ha al più n vettori.

Dimensione, basi.

In ciascun spazio vettoriale geometrico $\mathcal{V}^n$, ogni sequenza indipendente di n vettori è una base e viceversa ogni sequenza base è indipendente ed ha n vettori. Questo fatto vale in generale:

2 *Teorema 1.* Sia V uno spazio vettoriale di dimensione n. Se $b_1, \ldots, b_n$ sono n vettori indipendenti in V allora $b_1, \ldots, b_n$ sono una base di V.

Dimostrazione.
- Proviamo che ogni $v \in V$ si scrive come combinazione lineare di $b_1, \ldots, b_n$. La sequenza

$$b_1, \ldots, b_n, v$$

ha $n + 1$ vettori, per ipotesi il massimo numero di vettori indipendenti in V è n, quindi la sequenza $b_1, \ldots, b_n, v$ è dipendente; per la proposizione, si ha $b_1 = \underline{0}$ oppure uno dei vettori della sequenza è combinazione lineare dei precedenti; ancora per la proposizione, essendo $b_1, \ldots, b_n$ indipendente, si ha $b_1 \neq \underline{0}$ e nessun b_i è combinazione lneare dei precedenti; dunque v è combinazione lineare di $b_1, \ldots, b_n$:

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

- Proviamo che ogni $v \in V$ si scrive in un unico modo come combinazione lineare di $b_1, \ldots, b_n$. Siano

$$v = \sum_{i=1}^{n} \beta_i' \, b_i,$$

$$v = \sum_{i=1}^{n} \beta_i'' \, b_i;$$

Sottraendo alla 1° uguaglianza la 2° uguaglianza, si ha

$$\underline{0} = \sum_{i=1}^{n} (\beta_i' - \beta_i'')\, b_i;$$

essendo $b_1, \ldots, b_n$ indipendenti, si ha $\beta_i' - \beta_i'' = 0$ per ogni i, cioè $\beta_i' = \beta_i''$ per ogni i. ■

Come conseguenza, si ha

Ogni sequenza di n vettori indipendenti di $\mathbb{R}^n$ è una base di $\mathbb{R}^n$.

Teorema 2. Sia V uno spazio vettoriale. Se in V esiste una base $b_1, \ldots, b_n$ di n vettori, allora V ha dimensione finita n.

Non diamo dimostrazione. Come conseguenza, si ha

Ogni base di $\mathbb{R}^n$ è ha n vettori.

I due teoremi si possono riassumere come segue:

Teorema. Una sequenza di vettori di uno spazio vettoriale V di dimensione finita n è una base di V se e solo se è una sequenza indipendente di n vettori.

2 Sistemi lineari, spazi vettoriali, matrici

2.1 Sistemi lineari

n equazioni e incognite ($n = 2, 3$)

I sistemi lineari di 2 equazioni in 2 incognite x, y sono quelli del tipo

$$(2.1) \qquad \begin{cases} a_1 x + b_1 y = p_1 \\ a_2 x + b_2 y = p_2 \end{cases} ;$$

i dati di cui consiste il sistema si possono raccogliere nella matrice

$$\left[\begin{array}{cc|c} a_1 & b_1 & p_1 \\ a_2 & b_2 & p_2 \end{array} \right].$$

In genere, un tale sistema ha una ed una sola soluzione. Un metodo di risoluzione, detto metodo di "eliminazione" consiste in linea di massima in: (1) sommare alla 2° equazione un multiplo della 1° equazione, in modo da eliminare la x dalla 2° equazione; (2) dalla 2° equazione ricavare il valore della y; (3) sostituire il valore della y nella 1° equazione, e dalla 1° equazione ricavare la x. Ma il sistema può anche avere infinite soluzioni, o nessuna soluzione.

Esempio 1. Risoluzione di un sistema col metodo di eliminazione.

$$\begin{cases} x + 2\,y &= 4 \\ 3\,x + 5\,y &= 9 \end{cases} \qquad \begin{array}{rcr} x + 2\,y &=& 4 \\ -y &=& -3 \end{array} \qquad \begin{array}{rcr} x + 6 &=& 4 \\ y &=& 3 \end{array} \qquad \begin{array}{rcr} x &=& -2 \\ y &=& 3 \end{array}$$

il sistema ha l'unica soluzione $x = -2,\ y = 3$.

Esempio 2.

$$\begin{cases} x + y = 1 \\ 2\,x + 2\,y = 2 \end{cases}$$

ogni soluzione della 1° equazione è anche soluzione della 2° equazione, il sistema ha le stesse soluzioni della 1° equazione, ha infinite soluzioni.

Esempio 3.

$$\begin{cases} x + y = 1 \\ 2\,x + 2\,y = 3 \end{cases}$$

nessuna soluzione della 1° equazione è anche soluzione della 2° equazione, il sistema non ha soluzioni.

Interpretazione vettoriale. Ogni sistema lineare di 2 equazioni in 2 incognite si può riscrivere come 1 equazione in 2 incognite con coefficienti e termine noto coppie

$$(2.2) \qquad x \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + y \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \quad \text{in breve} \quad x\,a + y\,b = p;$$

i dati di cui consiste il sistema si possono raccogliere nella matrice

$$\left[\begin{array}{cc|c} a & b & p \end{array} \right], \text{ in breve } \left[\begin{array}{c|c} A & p \end{array} \right]$$

dove A è la matrice con colonne a, b.

Le seguenti proprietà sono equivalenti

il sistema (2.1) ha un'unica soluzione, per ogni $p_1, p_2 \in \mathbb{R}$;

l'equazione (2.2) ha un'unica soluzione, per ogni $p \in \mathbb{R}^2$;

a, b è una base di $\mathbb{R}^2$;

a, b sono indipendenti.

Esempio. Le seguenti proprietà sono equivalenti

$\begin{cases} x + 2\,y &= p_1 \\ 3\,x + 5\,y &= p_2 \end{cases}$ ha un'unica soluzione, per ogni $p_1, p_2 \in \mathbb{R}$;

$x \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ ha un'unica soluzione, per ogni $p \in \mathbb{R}^2$;

$\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ è una base di $\mathbb{R}^2$;

$\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ sono indipendenti.

L'ultima proprietà è vera, dunque tutte le proprietà sono vere.

I sistemi lineari di 3 equazioni in 3 incognite x, y, z sono quelli del tipo

$$\begin{cases} a_1x + b_1y + c_1z = p_1 \\ a_2x + b_2y + c_2z = p_2 \\ a_3x + b_3y + c_3z = p_3 \end{cases} \tag{3.1}$$

i dati di cui consiste il sistema si possono raccogliere nella matrice

$$\left[\begin{array}{ccc|c} a_1 & b_1 & c_1 & p_1 \\ a_2 & b_2 & c_2 & p_2 \\ a_3 & b_3 & c_3 & p_3 \end{array}\right].$$

In genere, un tale sistema ha una ed una sola soluzione. Il metodo di risoluzione per eliminazione consiste in linea di massima in:

(1) sommare alla 2° e 3° equazione un multiplo della 1°, in modo da eliminare la x;
(2) sommare alla 3° equazione un multiplo della 2°, in modo da eliminare la y;
(3) dalla 3° equazione ricavare il valore della z;
(4) sostituire il valore della z nella 2° equazione e ricavare la y;
(5) sostituire il valore della z e della y nella 1° equazione e ricavare la x.

Ma il sistema può anche avere infinite soluzioni, o nessuna soluzione.

Esempio 4. Risoluzione di un sistema col metodo di eliminazione.

$$\begin{cases} x + 2\,y + 3\,z = 7 \\ 2\,x + 3\,y + 4\,z = 8 \\ 3\,x + 4\,y + 6\,z = 9 \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 2 & 3 & 4 & 8 \\ 3 & 4 & 6 & 9 \end{array}\right]$$

sommiamo alla 2° e 3° equazione dei multipli della 1°, per eliminare la x:

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 7 \\ 0 & -1 & -2 & -6 \\ 0 & -2 & -3 & -12 \end{array}\right]$$

sommiamo alla 3° equazione un multiplo della 2°, per eliminare la y:

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 7 \\ 0 & -1 & -2 & -6 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\begin{cases} x + 2\,y + 3\,z & = & 7 \\ \quad -y - 2\,z & = & -6 \\ \qquad\quad z & = & 0 \end{cases}$$

dalla 3° equazione ricaviamo $z = 0$;

sostituiamo $z = 0$ nella 2° equazione, otteniamo $-y = -6$, ricaviamo $y = 6$;

sostituiamo $z = 0$ e $y = 6$ nella 1° equazione, otteniamo $x + 12 = 7$, ricaviamo $x = -5$.

Esempio 5.

$$\begin{cases} x + y + z = 5 \\ x + 2\,y + 3\,z = 6 \\ 2\,x + 3\,y + 4\,z = 7 \end{cases}$$

Da una parte si ha

$$(x + y + z) + (x + 2\,y + 3\,z) = 2\,x + 3\,y + 4\,z$$

e dall'altra parte si ha

$$5 + 6 \neq 7;$$

dunque il sistema non ha alcuna soluzione.

Interpretazione vettoriale. Ogni sistema lineare di 3 equazioni in 3 incognite si può riscrivere come 1 equazione in 3 incognite con coefficienti a, b, c e termine noto $p \in \mathbb{R}^3$

(3.2) $\qquad x\,a + y\,b + z\,c = p;$

i dati di cui consiste il sistema si possono raccogliere nella matrice di 3 righe e 3+1 colonne

$$\left[\begin{array}{ccc|c} a & b & c & p \end{array}\right], \text{ in breve } \left[\begin{array}{c|c} A & p \end{array}\right].$$

Le seguenti proprietà sono equivalenti

il sistema (3.1) ha un'unica soluzione, per ogni tre termini noti;

l'equazione (3.2) ha un'unica soluzione, per ogni termine noto;

a, b, c è una base di $\mathbb{R}^3$;

a, b, c è indipendente.

m equazioni, n incognite

Un sistema lineare di m equazioni in n incognite $x_1, x_2, \ldots, x_n$ è una sequenza di equazioni del tipo

$$(4)\qquad \begin{cases} a_{11}x_1 & +\cdots & +a_{1n}x_n & = p_1 \\ & & \vdots & \\ a_{m1}x_1 & +\cdots & +a_{mn}x_n & = p_m \end{cases}$$

con gli a_{ij} e i b_i costanti in $\mathbb{R}$. Il sistema è rappresentato dalla matrice di m righe ed $n+1$ colonne

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & p_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & p_m \end{array}\right].$$

Abbiamo descritto informalmente, su esempi di sistemi di n equazioni in n incognite con $n = 2, 3$, il metodo di eliminazione, per trovare la soluzione di un sistema nel "caso generale". Di seguito descriviamo su esempi una versione del metodo di eliminazione più generale e più flessibile, per rispondere alle seguenti domande su un sistema lineare qualsiasi:

il sistema ha soluzioni? se si, ha un'unica soluzione? se si, qual'è?

Al posto di scrivere "un'incognita compare con coefficiente $\neq 0$ in un'equazione" scriviamo in breve "un'incognita compare in un'equazione".

Il passo principale consiste nello scegliere in un'equazione un'incognita che compare e, sommando multipli dell'equazione alle altre equazioni, eliminare l'incognita dalle altre equazioni. Indichiamo l'operazione di sommare alla i° equazione r per la j° ($r \in \mathbb{R}$) con

$$\xrightarrow{i^\circ + r\cdot j^\circ}$$

Se a un certo passo si trova un'equazione in cui non compare alcuna incognita e il termine noto è $\neq 0$, allora si conclude che il sistema non ha alcuna soluzione.
Se questa circostanza non si presenta mai, si giunge a un sistema in cui a ciascuna equazione è associata un'incognita che compare solo nell'equazione.
Da ciascuna equazione si ricava l'incognita associata, lasciando libere le eventuali altre incognite; se ci sono incognite libere, il sistema ha infinite soluzioni, altrimenti ha un'unica soluzione.

Esempio 1.

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 3 & 5 & 9 \end{array}\right] \xrightarrow{2^\circ - 3\cdot 1^\circ} \left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & -1 & -3 \end{array}\right] \xrightarrow{1^\circ + 2\cdot 2^\circ} \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & -1 & -3 \end{array}\right];$$

1° equazione: $x = -2$;
2° equazione: $-y = -3$, da cui $y = 3$.

Esempio 2.

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 2 \end{array}\right] \xrightarrow{2^\circ - 2\cdot 1^\circ} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

2° equazione: $0\,x + 0\,y = 0$, soddisfatta per ogni x, y; la cancelliamo.
1° equazione: $x + y = 1$; ricaviamo $x = 1 - y$, con y libera (infinite soluzioni).

Esempio 3.

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 3 \end{array}\right] \xrightarrow{2°-2\cdot 1°} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]$$

2° equazione: $0\,x+0\,y=1$, non ha alcuna soluzione; il sistema non ha alcuna soluzione.

Esempio 4.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 2 & 3 & 4 & 8 \\ 3 & 4 & 6 & 9 \end{array}\right]$$

$$\xrightarrow[\ ]{\substack{2°-2\cdot 1° \\ 3°-3\cdot 1°}} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 0 & -1 & -2 & -6 \\ 0 & -2 & -3 & -12 \end{array}\right]$$

$$\xrightarrow[\ ]{\substack{1°+2\cdot 2° \\ 3°-2\cdot 2°}} \left[\begin{array}{ccc|c} 1 & 0 & -1 & -5 \\ 0 & -1 & -2 & -6 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[\ ]{\substack{1°+3° \\ 2°+2\cdot 3°}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & -1 & 0 & -6 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

1° equazione: $x=-5$; 2° equazione: $-y=-6$, da cui $y=6$; 3° equazione: $z=0$.

Esempio 6.

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \\ 1 & 1 & 0 & 4 \end{array}\right] \qquad \text{nella 1° equaz. scegliamo la } y,$$

$$\xrightarrow{3°-1°} \left[\begin{array}{ccc|c} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \\ 1 & 0 & -1 & 2 \end{array}\right] \qquad \text{nella 2° equaz. scegliamo la } x,$$

$$\xrightarrow{3°-2°} \left[\begin{array}{ccc|c} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \\ 0 & 0 & -2 & -1 \end{array}\right]$$

$$\xrightarrow[\ ]{\substack{1°+1/2\cdot 3° \\ 2°+1/2\cdot 3°}} \left[\begin{array}{ccc|c} 0 & 1 & 0 & \frac{3}{2} \\ 1 & 0 & 0 & \frac{5}{2} \\ 0 & 0 & -2 & 1 \end{array}\right]$$

2° equazione: $x=\frac{5}{2}$; 1° equazione: $y=\frac{3}{2}$; 3° equazione: $-2\,z=1$, da cui $z=-\frac{1}{2}$.

Esempio 7.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & 2 & 3 & 5 \end{array}\right] \xrightarrow{2°-1°} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 1 \end{array}\right] \xrightarrow{1°-2°} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 3 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

1° equazione: $x-z=3$; da cui $x=3+z$;
2° equazione: $y+2\,z=1$; da cui $y=1-2\,z$;
z è libera; infinite soluzioni.

Esempio 8.

$$\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & 2 & 5 \\ 1 & 3 & 5 \end{array}\right] \xrightarrow[3^\circ - 1^\circ]{2^\circ - 1^\circ} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{array}\right] \xrightarrow{3^\circ - 2 \cdot 2^\circ} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{array}\right].$$

Il sistema non ha alcuna soluzione.

2 **Teorema.** Per ogni matrice di m righe ed n colonne di numeri reali

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

le seguenti affermazioni sono equivalenti:

(1) ciascun sistema con matrice dei coefficienti A ha un'unica soluzione;

(2) $m = n$ e le colonne di A sono linearmente indipendenti;

(3) $m = n$ e le righe ddi A sono linearmente indipendenti.

Dimostrazione parziale.

(1) equivale a (2).
Ciascun sistema con matrice dei coefficienti A è di m equazioni in n incognite $x_1, \ldots, x_n$ e si scrive come un'equazione

$$x_1\, a_1 + \cdots + x_n\, a_n = p$$

dove $a_1, \ldots, a_n \in \mathbb{R}^m$ sono le colonne di A e $p \in \mathbb{R}^m$.
Ciascun sistema ha un'unica soluzione se e solo se
ciascuna equazione ha un'unica soluzione se e solo se (per definizione di base)
$a_1, \ldots, a_n$ è una base di $\mathbb{R}^m$ se e solo se (per i Th. sulle basi)
$n = m$ e $a_1, \ldots, a_m$ è indipendente.

(1) equivale a (3): lo proveremo più avanti.

2.2 Algebra delle matrici

Prodotto di matrici

Righe, colonne. Una $n-$pla ordinata di numeri reali si scrive solitamente scrivendo le sue componenti nell'ordine, per riga, separate da virgole, tra parentesi tonde, ma si scrive anche scrivendo le componenti per riga, separate da spazi, tra parentesi quadre, oppure per colonna, tra parentesi quadre; ad esempio, la coppia ordinata $(6, 7)$ si scrive anche

$$\begin{bmatrix} 6 & 7 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 7 \end{bmatrix}.$$

D'ora innanzi, consideriamo queste altre scritture, in righe e colonne, e le consideriamo come oggetti diversi. Sotto un simbolo scriviamo $1 \times n$ per indicare che rappresenta una riga di n numeri reali, e indichiamo con $\mathbb{R}^{1 \times n}$ l'insieme delle righe di n numeri reali; analogamente, scriviamo $n \times 1$ per una colonna di n numeri reali, e indichiamo con $\mathbb{R}^{n \times 1}$ l'insieme delle colonne di n numeri reali; nell'esempio,

$$\underset{1\times 2}{a} = \begin{bmatrix} 6 & 7 \end{bmatrix} \in \mathbb{R}^{1\times 2}, \quad \underset{2\times 1}{b} = \begin{bmatrix} 6 \\ 7 \end{bmatrix} \in \mathbb{R}^{2\times 1}.$$

Prodotto righe per colonne. Data una riga ed una colonna di numeri reali, aventi lo stesso numero di componenti, moltiplicando ciascuna componente della riga per la rispettiva componente della colonna e sommando si ottiene un numero reale che si dice prodotto della riga per la colonna e si indica giustapponendo la riga e la colonna (prima la riga e poi la colonna). Il prodotto di una riga per una colonna con diversi numeri di componenti non è definito. Ad esempio,

$$\begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = 2\cdot 4 + 3\cdot 5 = 23$$

$$\begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad \text{non definiti.}$$

In generale,

$$\underset{1\times n}{a} \; \underset{n\times 1}{b} = a_1 b_1 + a_2 b_2 + \cdots = \sum_{i=1}^{n} a_i b_i$$

(qui le a_i sono le componenti di a, analogamente per b_i e b),

$$\underset{1\times n'}{a} \; \underset{n''\times 1}{b} \quad \text{non definito per } n' \neq n''.$$

Questa operazione si può usare per scrivere in modo sintetico le espressioni lineari; ad esempio

$$2\,x_1 + 3\,x_2 = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\sum_{i=1}^{n} a_i x_i = \underset{1\times n}{a} \; \underset{n\times 1}{x}.$$

Matrici, terminologia, notazioni. Al posto di dire che una matrice "ha m righe ed n colonne", diciamo in breve che la matrice ha "tipo" $m \times n$; al posto di "elemento della $i-$ma riga e $j-$ma colonna" spesso diciamo in breve elemento di "posto" (i, j).

Notazioni per una matrice A :

$\underset{m\times n}{A}$ sta per "A ha tipo $m \times n$";

$\mathbb{R}^{m\times n}$ insieme delle matrici di tipo $m \times n$;

$A_{ij} =$ (elemento di posto (i, j) di A);

$A_{i:} =$ ($i-$ma riga di A);

$A_{:j} =$ ($j-$ma colonna di A).

Esempio:

$$\underset{2\times 3}{A} = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \in \mathbb{R}^{2\times 3};$$

$$A_{21} = 5, \quad A_{2:} = \begin{bmatrix} 5 & 6 & 7 \end{bmatrix}, \quad A_{:1} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

Prodotto di matrici. Se il numero delle colonne di una prima matrice è uguale al numero delle righe di una seconda matrice, allora possiamo moltiplicare ciascuna riga della prima per ciascuna colonna della seconda ed organizzare questi prodotti in una tabella; otteniamo così una terza matrice. Abbiamo così un'operazione sulle matrici detta "prodotto" di matrici. Ad esempio

$$\left[\begin{array}{cc} 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline 7 & 8 \end{array}\right] \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] = \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \\ 39 & 54 & 69 \end{bmatrix};$$

il prodotto della 3° riga della prima matrice per la 2° colonna della seconda matrice compare come elemento della 3° riga e 2° colonna della matrice prodotto:

$$\begin{bmatrix} 5 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = 5 \cdot 2 + 6 \cdot 5 = 40,$$

analogamente per gli altri.

In simboli, il prodotto di una matrice A di tipo $m \times n$ per una matrice B di tipo $n \times p$ è la matrice AB di tipo $m \times p$

$$\underset{m\times n}{A}, \quad \underset{n\times p}{B}, \quad \underset{m\times p}{AB}$$

data dai prodotti delle m righe di A per le p colonne di B : l'elemento di posto (i, j) in AB è il prodotto della riga $i-$ma di A per la colonna $j-$ma di B

$$\begin{aligned} (AB)_{ij} &= A_{i:}B_{:j} \\ &= [A_{i1} \ \dots \ A_{in}] \begin{bmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{bmatrix} \\ &= A_{i1}B_{1j} + \cdots + A_{in}B_{nj} \\ &= \sum_{h=1}^{n} A_{ih}B_{hj}. \end{aligned}$$

Il prodotto di una prima matrice per una seconda matrice tali che il numero delle colonne della prima è diverso dal numero delle righe della seconda non è definito:

$$\underset{m\times n'}{A}, \quad \underset{n''\times p}{B}, \quad AB \text{ non definito per } n' \neq n''.$$

Uso: scrittura di combinazioni e sistemi lineari. Questa operazione si può usare per scrivere in modo sintetico le combinazioni lineari. Ad esempio

$$x_1 \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 3 & 6 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Il sistema lineare di 3 equazioni in 2 incognite a coefficienti, incognite, termini noti in $\mathbb{R}$

$$\begin{cases} 2\,x_1 + 5\,x_2 = 7 \\ 3\,x_1 + 6\,x_2 = 8 \\ 4\,x_1 + 7\,x_2 = 9 \end{cases}$$

si può riscrivere come un'unica equazione lineare con una matrice coefficiente 3×2, un'incognita 2×1, un termine noto 3×1

$$\begin{bmatrix} 2 & 5 \\ 3 & 6 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

In generale:

- ogni combinazione lineare di colonne si fattorizza come prodotto della matrice delle colonne per la colonna dei coefficienti

$$\sum_{j=1}^{n} x_j A_{:j} = Ax;$$

- ogni sistema lineare di m equazioni in n incognite con coefficienti, incognite, termini noti in $\mathbb{R}$

$$\begin{cases} a_{11}\, x_1 + \cdots + a_{1n}\, x_n = b_1 \\ \vdots \\ a_{m1}\, x_1 + \cdots + a_{mn}\, x_n = b_m \end{cases}$$

si scrive come un'equazione lineare con matrice coefficiente $m \times n$, incognita $n \times 1$, termine noto $m \times 1$

$$Ax = b.$$

Proprietà

Premessa. Per ogni due matrici A, B, tali che il numero di colonne di A è uguale al numero di righe di B, ciascuna riga di A si può moltiplicare per ciascuna colonna di B, e questi prodotti danno la matrice prodotto AB. In altri termini, per ogni matrice A di tipo $m \times n'$ e B di tipo $n'' \times p$, se $n' = n''$ si definsce la matrice prodotto AB di tipo $m \times p$ ponendo

$$(AB)_{ij} = A_{i:}B_{:j} = \sum_{h=1}^{n} A_{i\,h}B_{h\,j};$$

se invece $n' \neq n''$ non si definisce alcuna matrice prodotto. Gi esempi più semplici si ottengono da matrici riga e matrici colonna:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = [11], \qquad \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \text{non definito}$$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}.$$

I numeri reali possono essere identificati con matrici 1×1 e al prodotto di numeri reali corrisponde il prodotto di matrici $1x1$:

$$a \simeq [a], \; b \simeq [b], \; ab \simeq [a][b];$$

in questo senso, il prodotto di numeri reali ha come estensione il prodotto di matrici. Nel seguito, ripercorriamo gli aspetti principali del prodotto di numeri reali e vagliamo gli analoghi aspetti del prodotto di matrici.

Innanzitutto, sottolineiamo di nuovo che, mentre il prodotto di due numeri reali è sempre definito, il prodotto di due matrici non è sempre definito. Può succedere che il prodotto di una prima matrice per una seconda matrice sia definito, e che il prodotto della seconda per la prima non lo sia. In breve:

il prodotto di matrici è un'operazione parziale, non commutativa.

Numero 1, matrici identità. Fra i numeri reali, il numero 1 è caratterizzato in termini di prodotto dalla proprietà

$$1\,a = a, \quad \forall a \in \mathbb{R}.$$

L'analogo del numero 1 è dato dalle "matrici identità"

$$\mathrm{I}_1 = [1], \qquad \mathrm{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathrm{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \ldots;$$

$$\mathrm{I}_n \text{ è } n \times n \text{ con elementi } (\mathrm{I}_n)_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \qquad \forall i, j = 1, \ldots, n.$$

Queste matrici sono caratterizzate da

$$\mathrm{I}_m A = A = A\mathrm{I}_n, \qquad \text{per ogni } A \; m \times n$$

Un esempio:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + 0 \cdot 3 \\ 0 \cdot 2 + 1 \cdot 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Proviamo una delle proprietà.

I_m ed A sono $m \times m$ e $m \times n$, quindi $\mathrm{I}_m A$ è definita ed è $m \times n$; per ogni i da 1 ad m e j da 1 a n si ha

$$(\mathrm{I}_m A)_{ij} = \sum_{h=1}^{m} (\mathrm{I}_m)_{ih} A_{hj} = (\mathrm{I}_m)_{ii} A_{ij} = A_{ij};$$

quindi $\mathrm{I}_m A = A$.

Associatività. Date tre matrici A, B, C tali che il numero di colonne di A sia uguale al numero di righe di B e il numero di colonne di B sia uguale al numero di righe di C, abbiamo due modi di moltiplicarle, mantenendo l'ordine,

$$(AB)C, \qquad A(BC).$$

Ad esempio, per

$$A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

si ha

$$(AB)C = \left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix}$$

$$A(BC) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix}$$

Quello che abbiamo visto su questo esempio si prova valere in generale.

La moltiplicazione di matrici possiede la proprietà associativa: comunque siano date tre matrici A, B, C, l'espressione $(AB)C$ è definita se e solo se è definita l'espressione $A(BC)$, e in tal caso si ha

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

Potremo così scrivere un prodotto di più matrici senza usare parentesi.

Trasposizione. Data una matrice A, leggendo gli elementi per righe e trascrivendoli per colonne si ottiene una matrice, che si dice "trasposta" di A e si indica con A^T. Ad esempio:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = A^T.$$

In simboli: trasposta di una matrice A $m \times n$ è la matrice A^T $n \times m$ tale che

$$\left(A^T\right)_{ij} = A_{ji} \quad \forall i, j.$$

Ovviamente la trasposta della trasposta di una matrice è la matrice stessa: $(A^T)^T = A$. Si vede che A è moltiplicabile per B se e solo se B^T è moltiplicabile per A^T e si prova che i due prodotti sono uguali:

$$(AB)^T = B^T A^T$$

Un esempio:

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$$

Matrice inversa Ricordiamo che l'inverso di un numero reale a è un numero reale b tale che $ab (= ba) = 1$; a possiede inverso se e solo se $a \neq 0$; in tal caso, a possiede un unico inverso, lo si indica con a^{-1}. Ad esempio: l'inverso di $\frac{2}{3}$ è il numero che moltiplicato per $\frac{2}{3}$ dà 1, dunque $(\frac{2}{3})^{-1} = \frac{3}{2}$; l'inverso di $\pi = 3,1415\ldots$ è un numero reale che moltiplicato per π dà 1, si ha $\pi^{-1} = 0.3183\ldots$ (così come non è banale determinare le cifre decimali di π, allo stesso modo non è banale determinare le cifre decimali di π^{-1}).

Definizione. Si dice che una matrice A $n \times n$ ha una matrice inversa B $n \times n$ se

$$AB = \mathrm{I}_n = BA;$$

in tal caso, si prova che A ha un'unica inversa, la si indica con A^{-1}.

L'unicità dell'inversa segue dalle proprietà del prodotto di matrici: se A ha due inverse B' e B'', allora

$$B' = B'\mathrm{I} = B'(AB'') = (B'A)B'' = \mathrm{I}B'' = B'';$$

si noti che abbiamo usato solo il fatto che A ha inversa sinistra B' e inversa destra B''.

Esempio. $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ ha un'inversa $\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}$ se e solo se

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix};$$

la prima equazione equivale a 2 sistemi lineari in 2 incognite

$$\begin{cases} b_1 + 2\,b_2 = 1 \\ b_1 + 3\,b_2 = 0 \end{cases} \qquad \begin{cases} c_1 + 2\,c_2 = 0 \\ c_1 + 3\,c_2 = 1 \end{cases};$$

risolvendo i sistemi, si trova $\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$;

si verifica che la matrice soddisfa anche la seconda equazione;

dunque

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}.$$

Esempio. $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ ha un'inversa $\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}$ se e solo se

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix};$$

la prima equazione equivale ai 2 sistemi lineari in 2 incognite

$$\begin{cases} b_1 + 3\,b_2 = 1 \\ 2\,b_1 + 6\,b_2 = 0 \end{cases} \qquad \begin{cases} c_1 + 3\,c_2 = 0 \\ 2\,c_1 + 6\,c_2 = 1 \end{cases};$$

uno dei due sistemi (in realtà entrambi) non ha soluzioni;

quindi $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ non ha alcuna inversa.

Equazioni matriciali

Consideriamo l'equazione $2\,x = 3$ nell'incognita x su $\mathbb{R}$; subito ricaviamo la soluzione $x = \frac{3}{2}$. Ricordiamo che questa risoluzione si basa sui seguenti passaggi (che normalmente si omettono)

$$2\,x = 3 \quad \tfrac{1}{2}\,2\,x = \tfrac{1}{2}\,3 \quad 1\,x = \tfrac{3}{2} \quad x = \tfrac{3}{2}.$$

In questa sezione consideriamo più in generale equazioni matriciali, con coefficiente, incognita e termine noto matrici e mostriamo un modo per risolverne alcune.

Una matrice A ha inversa se due certe equazioni che coinvolgono A hanno una soluzione; la soluzione di queste equazioni è l'inversa di A. In tal caso tutte le equazioni che coinvolgono A hanno un'unica soluzione e la soluzione si può ottenere tramite l'inversa di A. Precisamente si ha:

Proposizione. Sia $\underset{n\times n}{A}$ invertibile. Allora

(1) Ogni equazione $AX = \underset{n\times p}{B}$ ha un'unica soluzione: $X = A^{-1}B$.

(2) Ogni equazione $XA = \underset{m\times n}{B}$ ha un'unica soluzione: $X = BA^{-1}$.

In particolare:

(0) Ogni sistema lineare $Ax = b$ ha un'unica soluzione: $x = A^{-1}b$.

Dimostrazione della (1).

Da una parte:

$AX = B$ implica

$A^{-1}(AX) = A^{-1}B$ implica $(A^{-1}A)X = A^{-1}B$ implica $\mathrm{I}_n X = A^{-1}B$ implica

$X = A^{-1}B$.

Dall'altra:

$A(A^{-1}B) = B$, in quanto

$A(A^{-1}B) = (AA^{-1})B = \mathrm{I}_n B = B$.

Esempio. Sappiamo che $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$, dunque

il sistema $\begin{cases} x_1 + 2\,x_2 = 8 \\ x_1 + 3\,x_2 = 9 \end{cases}$, cioè $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \end{bmatrix}$,

ha un'unica soluzione: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$.

Indipendenza di righe, indipendenza di colonne

Abbiamo definito uno spazio vettoriale come un insieme con un'operazione di somma di vettori ed un'operazione di prodotto di numeri per vettori che soddisfano certe proprietà; in senso stretto, abbiamo così definito uno spazio vettoriale sinistro. In modo analogo, si definisce uno spazio vettoriale destro, nel quale si ha un'operazione di prodotto di vettori per numeri. Poichè il prodotto di numeri reali è commutativo, ogni spazio vettoriale sinistro si può anche vedere come spazio vettoriale destro e viceversa.

Osserviamo che la moltiplicazione di matrici fornisce

- un prodotto di vettori colonna $n \times 1$ per (matrici 1×1) $\simeq$ (numeri reali);
- un prodotto di (numeri reali) $\simeq$ (matrici 1×1) per vettori riga $1 \times n$.

L'insieme dei vettori colonna e l'insieme dei vettori riga hanno rispettivamente una struttura di spazio vettoriale destro e una struttura di spazio vettoriale sinistro, compatibili col prodoto di matrici. Non c'è necessità di adottare questo punto di vista, ma questo punto di vista porta a svolgere espressioni e provare identità in modo più naturale.

Combinazioni lineari e prodotti di matrici.

- Il prodotto di una riga per una colonna può essere visto come una combinazione lineare degli elementi della riga (con coefficienti a destra) e come una combinazione lineare (con coefficienti a sinistra) degli elementi della colonna

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + \cdots + a_n b_n.$$

- Il prodotto di una matrice per una colonna di numeri equivale alla combinazione lineare delle colonne della matrice, con coefficienti i numeri.

Un esempio:

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2\,x_1 + 3\,x_2 \\ 4\,x_1 + 5\,x_2 \\ 6\,x_1 + 7\,x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} x_1 + \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} x_2.$$

In generale:

$$\begin{bmatrix} a & b & \dots \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \vdots \end{bmatrix} = a\alpha + b\beta + \cdots$$

($a, b, \dots$ colonne, $\alpha, \beta, \dots$ numeri, soddisfacenti le dovute condizioni). Infatti: i due membri danno due colonne dello stesso tipo, e per ogni indice i,

$$(1^\circ \text{ membro})_i = \begin{bmatrix} a_i & b_i & \dots \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \vdots \end{bmatrix} = a_i\,\alpha + b_i\,\beta + \cdots;$$

$$(2^\circ \text{ membro})_i = (a\,\alpha)_i + (b\,\beta)_i + \cdots = a_i\,\alpha + b_i\,\beta + \cdots.$$

Uso: espressione sintetica dell'indipendenza lineare.

Ogni sequenza $c_1, \ldots, c_n$ di colonne $m \times 1$ dà una matrice $A = [c_1 \ \ldots \ c_n]$ $m \times n$ e viceversa. La sequenza soddisfa la condizione

$c_1 x_1 + \cdots + c_n x_n = \underline{0}$ vale solo per i numeri $x_1 = \cdots = x_n = 0$

se e solo se la matrice soddisfa la condizione

$Ax = \underline{0}$ vale solo per la colonna $x = \underline{0}$.

- Il prodotto di una riga di numeri per una matrice equivale alla combinazione lineare, con coefficienti i numeri, delle righe della matrice.

Un esempio:

$$\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} y_1\, 2 + y_2\, 4 + y_3\, 6 & y_1\, 3 + y_2\, 5 + y_3\, 7 \end{bmatrix}$$
$$= y_1 \begin{bmatrix} 2 & 3 \end{bmatrix} + y_2 \begin{bmatrix} 4 & 5 \end{bmatrix} + y_3 \begin{bmatrix} 6 & 7 \end{bmatrix}$$

In generale:

$$\begin{bmatrix} \alpha & \beta & \ldots \end{bmatrix} \begin{bmatrix} a \\ b \\ \vdots \end{bmatrix} = \alpha a + \beta b + \cdots$$

($\alpha, \beta, \ldots$ numeri, $a, b, \ldots$ righe, soddisfacenti le dovute condizioni).

Uso: espressione sintetica dell'indipendenza lineare.

Ogni sequenza $r_1, \ldots, r_m$ di righe $1 \times n$ dà una matrice $A = \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix}$ $m \times n$ e viceversa.

La sequenza soddisfa la condizione

$y_1 r_1 + \cdots + y_n r_n = \underline{0}$ vale solo per i numeri $y_1 = \cdots = y_n = 0$

se e solo se la matrice soddisfa la condizione

$yA = \underline{0}$ vale solo per la riga $y = \underline{0}$.

Ampliamo, in enunciato e dimostrazione, il Teorema su indipendenza lineare, basi e sistemi lineari (cfr. § 2.1 Sistemi lineari). Di seguito sia A una matrice $m \times n$ con righe $r_1, \ldots, r_m$ e colonne $c_1, \ldots, c_n$

$$\begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix} = A = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}.$$

Prop. Se le righe di A sono linearmente dipendenti (in $\mathbb{R}^n$) allora esiste qualche colonna (in $\mathbb{R}^m$) che non è comnbinazione lineare delle colonne di A; specificamente: se vale una relazione

$\bar{y}_1 r_1 + \cdots \bar{y}_m r_m = \underline{0}$

con qualche $\bar{y}_i \neq 0$, allora l'equazione

$$c_1x_1 + \cdots + c_nx_n = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_m \end{bmatrix}$$

non ha alcuna soluzione. Vale un analogo enunciato scambiando il ruolo delle righe e delle colonne.

Esempio. Per la matrice $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$,

le righe soddisfano la relazione

$$-3\begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix},$$

quindi l'equazione

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 6 \end{bmatrix} x_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

non ha alcuna soluzione (si lascia la verifica al lettore).

Dimostrazione. Per ipotesi si ha $\begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_m \end{bmatrix} A = \underline{0}$, con qualche $\bar{y}_i \neq 0$.

Se valesse $A\bar{x} = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_m \end{bmatrix}$ per qualche $\bar{x} \in \mathbb{R}^n$, allora si avrebbe la contraddizione:

- da una parte $\begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_m \end{bmatrix} (A\bar{x}) = \left(\begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_m \end{bmatrix} A\right) \bar{x} = \underline{0}\,\bar{x} = 0$;

- dall'altra $\begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_m \end{bmatrix} \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_m \end{bmatrix} = \sum_1^m \bar{y}_i^2 > 0.$

Teorema. Le seguenti condizioni su una matrice $m \times n$ sono equivalenti

1: $n = m$ e colonne sono lin. indip.;	1': $m = n$ e righe sono lin. indip.;
2: colonne sono base di $\mathbb{R}^m$;	2': righe sono una base di $\mathbb{R}^n$;

Dimostrazione.

$1 \Rightarrow 2$: per il Th. sulle basi; $2 \Rightarrow 1$: ovvio. Analogamente per $1' \Leftrightarrow 2'$.

$2 \Rightarrow 1'$: se $1'$ non valesse allora, per la Prop., 2 non varrebbe; $2' \Rightarrow 1$ analogo.

Definizione. Una matrice che soddisfa una (e quindi tutte) le condizioni del Teorema si dice "non singolare".

Indipendenza, invertibilità, inversione

Il prodotto matrice-per-matrice è legato al prodotto matrice-per-colonna e al prodotto riga-per-matrice dalla seguente relazione.

Siano A una matrice con righe $r, s, \ldots$ e B una matrice con colonne $c, d, \ldots$, righe e colonne della stessa lunghezza, allora

(1) le colonne di AB sono i prodotti di A per le colonne di B :

$$A \left[\begin{array}{ccc} c & d & \cdots \end{array}\right] = \left[\begin{array}{ccc} Ac & Ad & \cdots \end{array}\right];$$

(2) le righe di AB sono i prodotti delle righe di A per B :

$$\begin{bmatrix} r \\ s \\ \vdots \end{bmatrix} B = \begin{bmatrix} rB \\ sB \\ \vdots \end{bmatrix}.$$

Verifichiamo la (1):

$$A \left[\begin{array}{ccc} c & d & \cdots \end{array}\right] = \begin{bmatrix} r \\ s \\ \vdots \end{bmatrix} \left[\begin{array}{ccc} c & d & \cdots \end{array}\right] = \begin{bmatrix} rc & rd & \cdots \\ sc & sd & \cdots \\ \vdots & \vdots & \end{bmatrix} = \left[\begin{array}{ccc} Ac & Ad & \cdots \end{array}\right].$$

3 **Teorema.** Per ogni matrice A quadrata $n \times n$, le seguenti condizioni sono equivalenti: (1) A è non singolare; (2) A è invertibile.

Dimostrazione.

(1) $\Rightarrow$ (2). Se A è non singolare, allora le colonne di A sono una base di $\mathbb{R}^n$, allora esistono (uniche) colonne $b_1, \ldots b_n$ tali che $Ab_1 = e_1, \ldots, Ab_n = e_n$, allora

$$A[b_1 \ \cdots \ b_n] = [Ab_1 \ \cdots \ Ab_n] = [e_1 \ \cdots \ e_n] = \mathrm{I}_n,$$

e A ha un'inversa destra. In modo simile si trova che A possiede un'inversa sinistra. Quindi A è invertibile.

(2) $\Rightarrow$ (1). Se A è invertibile, allora l'equazione $Ax = \underline{0}$ ha un'unica soluzione, quindi ha l'unica soluzione $x = \underline{0}$, quindi le colonne di A son lin. indip. e A è non singolare.

Esempi. Cosa sarebbe cambiato nello studio dell'invertibilità e ricerca dell'inversa delle due matrici

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \text{ e } \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \text{ (cfr. p. 33) ?}$$

Osservando che la prima ha colonne indipendenti e la seconda dipendenti, avremmo potuto dire subito che la prima è invertibile e la seconda non lo è.

Per il calcolo dell'inversa della prima, avremmo determinato due colonne b, c tali che

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} c = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

e avremmo subito affermato che l'inversa destra $[b \ c]$ è l'inversa della matrice (senza controllare che è anche inversa sinistra).

Definizione. Le "operazioni elementari" su una sequenza $a_1, a_2, \ldots$ di vettori di uno spazio vettoriale V sono:

(1) sommare a un vettore un multiplo reale di un altro vettore

$$a_i := a_i + r\, a_j \qquad i \neq j$$

(2) moltiplicare un vettore per un numero non nullo:

$a_i := s\,a_i \qquad s \neq 0$

(3) scambiare due vettori

$a_i \leftrightarrow a_j.$

Queste operazioni lasciano invariata la indipendenza/dipendenza lineare della sequenza. Inoltre, applicate alla matrice completa (coefficienti e termini noti) di un sistema lineare, lasciano invariate le soluzioni del sistema.

Proposizione. Sia A una matrice non singolare. Allora la matrice

$[A|\mathrm{I}_n]$

si può trasformare, mediante operazioni elementari sulle righe, nella matrice

$[\mathrm{I}_n|A^{-1}.]$

Non diamo la dimostrazione.

Esempio 1.

$$
\begin{aligned}
[A|\mathrm{I}_2] &= \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \\
&\xrightarrow{2^\circ - 1^\circ} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] \\
&\xrightarrow{1^\circ - 2\cdot 2^\circ} \left[\begin{array}{cc|cc} 1 & 0 & 3 & -2 \\ 0 & 1 & -1 & 1 \end{array}\right] = [\mathrm{I}_2|A^{-1}]
\end{aligned}
$$

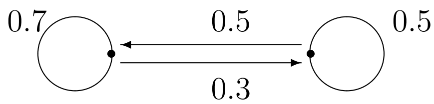
Esempio 2.

$$
\begin{aligned}
[A|\mathrm{I}_3] &= \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 & 0 \\ 3 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \\
&\xrightarrow{3^\circ - 3\cdot 1^\circ} \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 & 0 \\ 3 & 0 & -3 & -3 & 0 & 1 \end{array}\right] \\
&\xrightarrow{\frac{1}{2} 2^\circ} \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & \frac{1}{2} & 0 \\ 3 & 0 & -3 & -3 & 0 & 1 \end{array}\right] \\
&\xrightarrow{3^\circ - 3\cdot 2^\circ} \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & -6 & -3 & -\frac{3}{2} & 1 \end{array}\right] \\
&\xrightarrow{-\frac{1}{6}\cdot 3^\circ} \left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{6} \end{array}\right] \\
&\xrightarrow{2^\circ - 3^\circ;\ 1^\circ - 3^\circ} \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{4} & \frac{1}{6} \\ 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{4} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{6} \end{array}\right] \\
&\xrightarrow{1^\circ \leftrightarrow 2^\circ} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{4} & \frac{1}{6} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{4} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{6} \end{array}\right] = [\mathrm{I}_3|A^{-1}]
\end{aligned}
$$

Applicazione. Supponiamo che fra due città 1 e 2 nel passaggio da un tempo 0 ad un tempo 1 si siano verificati i seguenti spostamenti percentuali:

degli abitanti in 1 al tempo 0, il 70% è rimasto in 1 e il 30% si è trasferito in 2;
degli abitanti in 2 al tempo 0, il 50% si è trasferito in 1 e il 50% è rimasto in 2.

Ci si chiede di determinare i numeri degli abitanti nelle città al tempo 1 in funzione di quelli al tempo 0 e se possibile quelli al tempo 0 in funzione di quelli al tempo 1.



Indicati con x_1 e x_2 i numeri degli abitanti in 1 e 2 al tempo 0, e con y_1 e y_2 quelli al tempo 1, si ha

$$\begin{aligned} y_1 &= 0.7\,x_1 + 0.5\,x_2 \\ y_2 &= 0.3\,x_1 + 0.5\,x_2 \end{aligned}$$

che si può riscrivere

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.5 \\ 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

si ha

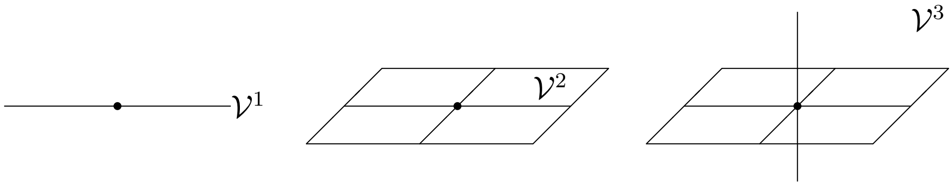
$$\begin{bmatrix} 0.7 & 0.5 \\ 0.3 & 0.5 \end{bmatrix}^{-1} = \begin{bmatrix} 2.5 & -2.5 \\ -1.5 & 3.5 \end{bmatrix};$$

dunque si ricava in modo unico

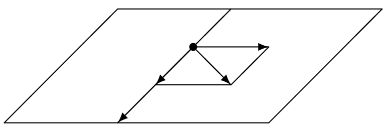
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.5 & -2.5 \\ -1.5 & 3.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Sottospazi

"La retta", "il piano", "lo spazio" euclidei sono entità a sé stanti; nel piano non c'è "la retta", ce ne sono varie; nello spazio ci sono vari piani. Analogamente: fissato nel piano un punto O e identificati i vettori con segmenti orientati uscenti da O, si ha che alla retta vettoriale $\mathcal{V}^1$ corrispondono nel piano vettoriale $\mathcal{V}^2$ varie rette vettoriali, una per ogni retta passante per O; rappresentati i vettori dello spazio con segmenti orientati uscenti da un punto O, al piano vettoriale $\mathcal{V}^2$ corrispondono nello spazio vettoriale $\mathcal{V}^3$ vari piani vettoriali, uno per ogni piano passante per O.



Osserviamo che per ciascun piano passante per O si ha: (1) per ogni due vettori che stanno sul piano, il parallelogramma da essi individuato sta sul piano, dunque il vettore somma sta sul piano; (2) per ogni vettore sul piano, anche la retta da esso individuata sta sul piano, dunque i vettori multipli reali del vettotre stanno sul piano.



Definizione. Un sottinsieme $U \subseteq V$ di uno spazio vettoriale V si dice "sottospazio vettoriale" di V se U è chiuso rispetto alle operazioni definite in V

(1) per ogni $u', u'' \in U$, anche $u' + u'' \in U$;

(2) per ogni $r \in \mathbb{R}$ ed ogni $u \in U$, anche $ru \in U$;

e U non è vuoto

(3) $U \neq \emptyset$;

equivalentemente, U contiene il vettore nullo di V :

(3') $\underline{0} \in U$.

(Infatti, se esiste un $\bar{u} \in U$, allora, per la (2), anche $\underline{0} = 0\,\bar{u} =\in U$.)

Osserviamo che in tal caso U, col vettore nullo $\underline{0}$, coi vettori opposti $-u$ per ogni $u \in U$, e con le operazioni ereditate da V è uno spazio vettoriale.

Infatti: per la (3'), $\underline{0} \in U$, per la (2), $-u = (-1)u \in U$ per ogni $u \in U$; questi elementi e le operazioni soddisfano le identità (1), ..., (8). Ad esempio, per la proprietà commutativa: per ogni $u', u'' \in U$ si ha

$$\begin{aligned} u' + u'' \text{ (calcolato in } U) &= u' + u'' \text{ (calcolato in } V) \\ &= u'' + u' \text{ (calcolato in } V) = u'' + u' \text{ (calcolato in } U). \end{aligned}$$

Esempio 1.

Identificati i vettori dello spazio con i segmenti orientati uscenti da un punto O, i sottospazi di $\mathcal{V}^3$ sono:

0- l'insieme ridotto al vettore nullo (il punto O);

1- per ciascuna retta passante per O, l'insieme i vettori che stanno sulla retta;

2- per ciascun piano passante per O, l'insieme dei vettori che stanno sul piano;

3- l'insieme $\mathcal{V}^3$ di tutti i vettori.

Da una parte, ciascuno di questi insiemi soddisfa le condizioni (1), (2), (3) (lo abbiamo già visto per quelli del tipo 2).

Dall'altra parte, se $\mathcal{U} \subseteq \mathcal{V}^3$ è un sottospazio, allora:

- se $\mathcal{U}$ contiene solo $\underline{0}$, allora è del tipo 0;

- se $\mathcal{U}$ contiene un vettore $i \neq \underline{0}$, allora per la (2) contiene anche tutti i vettori $x\, i$, tutti quelli sulla retta per i; se $\mathcal{U}$ non contiene altri vettori, allora è del tipo 1;

- se $\mathcal{U}$ contiene un vettore j che non sta sulla retta di i, allora per la (1) e (2) contiene anche i vettori combinazioni lineari $x\, i + y\, j$, quindi tutti quelli sul piano per i e j; se $\mathcal{U}$ non contiene altri vettori, allora è del tipo 2;

- se $\mathcal{U}$ contiene un vettore k che non sta sul piano per i, j, allora per la (1) e (2) contiene anche i vettori combinazioni lineari $x\, i + y\, j + z\, k$, quindi tutti quelli dello spazio.

Esempio 2.

Mostriamo algebricamente che l'insieme dei vettori su un piano per O è un sottospazio di $\mathcal{V}^3$. Fissiamo un riferimento i, j, k tale che i e j stiano sul piano e tramite di esso identifichiamo $\mathcal{V}^3$ con $\mathbb{R}^3$.

Un vettore sta sul piano se e solo se è del tipo $x\, i + y\, j \simeq \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$.

La somma di vettori di questo tipo è di questo tipo: $\begin{bmatrix} x' \\ y' \\ 0 \end{bmatrix} + \begin{bmatrix} x'' \\ y'' \\ 0 \end{bmatrix} = \begin{bmatrix} x' + x'' \\ y' + y'' \\ 0 \end{bmatrix}$;

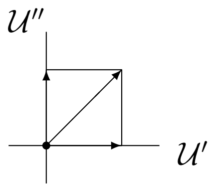
il prodotto di numeri per vettori di questo tipo è di questo tipo: $r \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} r\, x \\ r\, y \\ 0 \end{bmatrix}$.

Esempio 3.

Identificati i vettori del piano con i segmenti orientati uscenti da un punto O, consideriamo due rette rette distinte per O, gli insiemi $\mathcal{U}'$ e $\mathcal{U}''$ dei vettori che stanno sulle rette, e l'insieme unione $\mathcal{U} = \mathcal{U}' \smile \mathcal{U}''$.

$\mathcal{U}$ soddisfa la (2), ma non soddisfa la (1).

Se $u' \in \mathcal{U}'$ e $u'' \in \mathcal{U}''$, con $u' \neq \underline{0} \neq u''$, allora $u', u'' \in \mathcal{U}$ ma $u' + u'' \notin \mathcal{U}$.



Fissato un riferimento i, j con $i \in \mathcal{U}'$ e $j \in \mathcal{U}''$, e identificato $\mathcal{V}^2$ con $\mathbb{R}^2$, si ha:

$\mathcal{U}' = \{\begin{bmatrix} x \\ 0 \end{bmatrix};\ x \in \mathbb{R}\};$

$\mathcal{U}'' = \{\begin{bmatrix} 0 \\ y \end{bmatrix};\ y \in \mathbb{R}\};$

$\mathcal{U} = \{\begin{bmatrix} x \\ y \end{bmatrix};\ x, y \in \mathbb{R},\ x = 0 \text{ o } y = 0\}.$

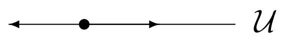
$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathcal{U}, \text{ ma } \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin \mathcal{U}.$

Esempio 4.

Identificati i vettori della retta con i segmenti orientati uscenti da un punto O, consideriamo una delle due semirette uscenti da O, e l'insieme $\mathcal{U}$ dei vettori che stanno sulla semiretta.

$\mathcal{U}$ soddisfa la (1), ma non soddisfa la (2).

Se $u \in \mathcal{U}$ con $u \neq \underline{0}$, allora $(-1)u = -u \notin \mathcal{U}$.



Fissato un riferimento i con $i \in \mathcal{U}$, e identificato $\mathcal{V}^1$ con $\mathbb{R}$, si ha:

$\mathcal{U} = \{x;\ x \geq 0\};$

$1 \in \mathcal{U}$, ma $(-1)1 = -1 \notin \mathcal{U}$.

Sottospazi generati da insiemi di vettori.

In analogia con la descrizione dei sottospazi di $\mathcal{V}^3$, iniziamo a descrivere i sottospazi di uno spazio vettoriale V. Sia $U \subseteq V$ un sottospazio di V.

$U \ni \underline{0}$. L'insieme $\{\underline{0}\}$ è un sottospazio di V.

Se $U \ni a \neq \underline{0}$, allora per la (2) U contiene anche tutti i multipli reali αa di a. L'insieme

$$\{\alpha a;\ \alpha \in \mathbb{R}\}$$

è un sottospazio di V. Infatti: per ogni due multipli $\alpha' a, \alpha'' a$ di a, la somma $\alpha' a + \alpha'' a = (\alpha' + \alpha'')a$ è ancora un multiplo di a; per ogni $r \in \mathbb{R}$ ed ogni multiplo αa di a, anche $r(\alpha a) = (r\alpha)a$ è un multiplo di a.
Osserviamo inoltre che, essendo $a \neq 0$, a è una base di questo spazio vettoriale; infatti ogni multiplo di a si scrive in un unico modo come multiplo di a : se $\alpha' a = \alpha'' a$, allora $(\alpha' - \alpha'')a = \underline{0}$ ed essendo $a \neq \underline{0}$ si ha $\alpha' - \alpha'' = 0$, da cui $\alpha' = \alpha''$.

Proposizione. Sia $a, b, \ldots$ una sequenza finita di vettori di V. L'insieme delle combinazioni lineari di $a, b, \ldots$ è un sottospazio di V, il più piccolo sottospazio di V

che contiene $a, b, \ldots$ Lo si dice "sottospazio generato da $a, b, \ldots$" e lo si indica con $\operatorname{span}\{a, b, \ldots\}$:

$$\operatorname{span}\{a, b, \ldots\} = \{\alpha a + \beta b + \cdots \,;\ \alpha, \beta, \ldots \in \mathbb{R}\}$$

Dimostrazione parziale.

(1) per ogni due vettori comb. lin. di $a, b, \ldots$, il vettore somma è comb. lin. di $a, b, \ldots$:

$$(\alpha' a + \beta' b + \cdots) + (\alpha'' a + \beta'' b + \cdots) = (\alpha' + \alpha'')a + (\beta' + \beta'')b + \cdots$$

(2) per ogni $r \in \mathbb{R}$ ed ogni vettore comb. lin. di $a, b, \ldots$, il vettore prodotto è comb. lin. di $a, b, \ldots$:

$$r(\alpha a + \beta b + \cdots) = (r\alpha)a + (r\beta)b + \cdots$$

(3) $\underline{0} = 0\, a + 0\, b + \cdots$ è comb. lin. di $a, b, \ldots$.

Ogni vettore della sequenza $a, b, \ldots$ è uguale a una comb. lin. dei vettori della sequenza, quella con coefficienti 1 per sé stesso e 0 per gli altri.

Esempio.

In $\mathcal{V}^3$, identificato con $\mathbb{R}^3$ mediante un riferimento i, j, k, siano

$$a = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

$\operatorname{Span}\{a, b, c\}$ è l'insieme delle combinazioni lineari di a, b, c

$$\alpha \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \alpha + \beta \\ -\alpha + \gamma \\ -\beta - \gamma \end{bmatrix}, \text{ con } \alpha, \beta, \gamma \in \mathbb{R}.$$

Essendo a, b non multipli uno dell'altro, $\operatorname{Span}\{a, b, c\}$ non è una retta vettoriale, ci chiediamo se è un piano vettoriale o l'intero spazio $\mathcal{V}^3$. Rispondiamo fra un po'.

Proprietà.

Diamo alcune proprietà della costruzione che associa ad ogni sequenza di vettori lo spazio da essa generato.

Proposizione. Siano $a, \ldots, c$ e d una sequenza finita di vettori e un vettore di V.

(1) $\operatorname{Span}\{a, \ldots, c\} \subseteq \operatorname{Span}\{a, \ldots, c, d\}$;

(2) $\operatorname{Span}\{a, \ldots, c\} = \operatorname{Span}\{a, \ldots, c, d\}$ se e solo se d è comb. lin. di $a, \ldots, c$.

1 *Dimostrazione.*

(1) ogni comb. lin. di $a, \ldots, c$ è anche una comb. lin. di $a, \ldots, c, d$:

$$\alpha\, a + \cdots + \gamma\, c = \alpha\, a + \cdots + \gamma\, c + 0\, d;$$

(2') Se $\operatorname{Span}\{a, \ldots, c\} = \operatorname{Span}\{a, \ldots, c, d\}$, allora da $d \in \operatorname{Span}\{a, \ldots, c, d\}$ segue $d \in \operatorname{Span}\{a, \ldots, c\}$, così d è comb. lin. di $a, \ldots, c$.

(2") Proviamo che, se d comb. lin. di $a, \ldots, c$, allora ogni comb. lin. di $a, \ldots, c, d$ si può scrivere come comb. lin. di $a, \ldots, c$. Infatti, se per certi $\bar{\alpha}, \ldots, \bar{\gamma} \in \mathbb{R}$

$$d = \bar{\alpha}\, a + \cdots + \bar{\gamma}\, c$$

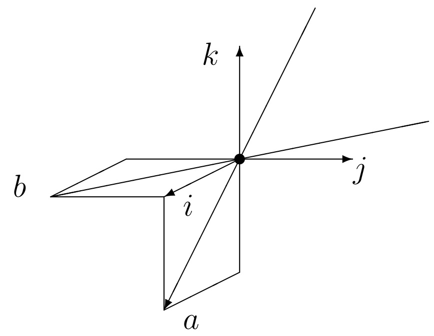
allora per ogni $\alpha, \ldots, \gamma \in \mathbb{R}$,

$$\alpha\, a + \cdots + \gamma\, c + \delta\, d = \alpha\, a + \cdots + \gamma\, c + \delta\, (\bar{\alpha}\, a + \cdots + \bar{\gamma}\, c) = (\alpha + \delta\bar{\alpha})\, a + \cdots + (\gamma + \delta\bar{\gamma})\, c.$$

Esempio, ripresa.

c è comb. lin. di a, b : $\quad \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = - \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix};$

quindi $\mathrm{Span}\{a, b, c\} = \mathrm{Span}\{a, b\}$ è un piano vettoriale: l'insieme dei vettori che stanno sul piano di a, b. Osserviamo che a, b è un riferimento, cioè una base di $\mathrm{Span}\{a, b\}$.



Questo esempio suggerisce una procedura per determinare una base e dunque la dimensione di un sottospazio generato da una sequenza (finita) di vettori: se c'è un vettore che è combinazione lineare dei precedenti, allora lo si può cancellare dalla sequenza lasciando invariato il sottospazio, si può proseguire fino a quando si ottine una sottosequenza nella quale nessun vettore è combinazione lineare dei precedenti; la sequenza finale è linearmente indipendente, ed è una base del sottospazio.

Svilupperemo questo tema prossimamente.

Sottospazi di uno spazio vettoriale finito-dimensionale

Ripresa.

Abbiamo definito sottospazio di uno spazio vettoriale V un sottinsieme $U \subseteq V$ chiuso rispetto all'operazione di somma e all'operazione di prodotto per numeri,

(1) $\qquad \forall u', u'' \in U, \quad u' + u'' \in U;$

(2) $\qquad \forall r \in \mathbb{R}, \forall u \in U, \quad r\, u \in U;$

e non vuoto $U \neq \emptyset$, equivalentemente, tale che $\underline{0} \in U$.

Le condizioni (1), (2) sono equivalenti all'unica condizione

$$\forall r', r'' \in \mathbb{R}, \forall u', u'' \in U, \quad r'u' + r''u'' \in U;$$

che a sua volta equivale alla condizione

$$\forall r_1, \ldots, r_m \in \mathbb{R}, \forall u_1, \ldots, u_m \in U, \quad \sum_{i=1}^{m} r_i u_i \in U.$$

Abbiamo osservato che un sottospazio U di uno spazio vettoriale V, con gli stessi elementi proivilegiati di V (vettore nullo ed opposti) e le stesse operazioni di V (operazione di somma e operazione di prodotto per numeri) è uno spazio vettoriale. Dunque, tutte le proposizioni provate per gli spazi vettoriali valgono per i sottospazi degli spazi vettoriali, in particolare:

ogni sottospazio di dimensione finita di uno spazio vettoriale possiede una base.

Per ogni sequenza finita $a_1, \ldots, a_m$ di vettori di uno spazio vettoriale V, abbiamo provato che l'insieme delle combinazioni lineari della sequenza è un sottospazio di V, il più piccolo, che contiene la sequenza; lo abbiamo detto "sottospazio generato" da $a_1, \ldots, a_m$ e lo abbiamo indicato con $\mathrm{Span}\{a_1, \ldots, a_m\}$.

Fatto. Per definizione, ogni $v \in \mathrm{Span}\{a_1, \ldots, a_m\}$ si scrive come combinazione lineare di $a_1, \ldots, a_m$; se $a_1, \ldots, a_m$ è linearmente indipendente, allora la scrittura è unica. Quindi:

Se $a_1, \ldots, a_m$ è lin. indip. allora è una base di $\mathrm{Span}\{a_1, \ldots, a_m\}$, e

$\dim \mathrm{Span}\{a_1, \ldots, a_m\} = m.$

Esempi.

I sottospazi di $\mathcal{V}^3$ sono

$\{\underline{0}\}$ (di dimensione 0);

le rette vettoriali $\mathrm{Span}\{a\}$, con $a \neq \underline{0}$ (quelli di dimensione 1);

i piani vettoriali $\mathrm{Span}\{a, b\}$, con a, b lin. indip. (quelli di dimensione 2);

$\mathcal{V}^3$ (di dimensione 3).

In $\mathbb{R}^n$, esistono sottospazi di dimensione $0, 1, 2, \ldots, n$

$\{\underline{0}\}$ (di dimensione 0);

$\mathrm{Span}\{e_1\}$ (di dimensione 1);

Span$\{e_1, e_2\}$ (di dimensione 2);

...

Span$\{e_1, e_2, \ldots, e_n\} = \mathbb{R}^n$ (di dimensione n).

In generale, si ha:

Proposizione.

(1) Ogni sottospazio U di uno spazio vettoriale V di dimensione finita ha dimensione (finita) $\dim(U) \leq \dim(V)$; solo $U = V$ ha $\dim(U) = \dim(V)$.

(2) Per ogni $0 \leq d \leq \dim(V)$ esiste un sottospazio $U \leq V$ con $\dim(U) = d$.

2 *Dimostrazione.*

(1) Ogni sequenza lin. indip. in U è anche sequenza lin. indip. in V (perchè U è un sottospazio di V), quindi ha una numero di vettori $\leq \dim(V)$, quindi U ha dimensione finita e $\dim(U) \leq$ dim(V). Se $\dim(U) = \dim(V)$, allora in U c'è una sequenza lin. indip. $u_1, \ldots, u_n$ con $n = \dim(V)$, allora per ogni $v \in V$ la sequenza $u_1, \ldots, u_n, v$ è lin. dip., allora v è combin. lin. di $u_1, \ldots, u_n$, allora $v \in U$ (perchè U è un sottospazio di V); allora $U = V$.

(2) Così come per ogni $0 \leq d \leq n$ si è costruito un sottospazio $U \leq \mathbb{R}^n$ con $\dim(U) = d$ usando la base canonica di $\mathbb{R}^n$, analogamente per ogni $0 \leq d \leq \dim(V)$ si costruisce un sottospazio $U \leq V$ con $\dim(U) = d$ usando una base di V.

Uso di basi di sottospazi. Esempio

Siano $a = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $c = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Non è immediato riconoscere se un vettore appartiene a Span$\{a, b, c\}$.

Ad esempio $d = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in$ Span$\{a, b, c\}$ se e solo se

esistono $\alpha, \beta, \gamma \in \mathbb{R}$ tali che $\alpha\, a + \beta\, b + \gamma\, c = d$ se e solo se

il sistema $\left[\begin{array}{rrr|r} 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 \\ 0 & -1 & -1 & 1 \end{array}\right]$ ha soluzioni.

Abbiamo visto che, essendo c combinazione lineare di a, b, si ha

$$\text{Span}\{a, b, c\} = \text{Span}\{a, b\};$$

i vettori a, b, essendo lin. indip., sono una base del sottospazio.

$d \in$ Span$\{a, b\}$ se e solo se

esistono $\alpha, \beta \in \mathbb{R}$ tali che $\alpha\, a + \beta\, b = d$, se e solo se

il sistema $\left[\begin{array}{rr|r} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{array}\right]$ ha soluzioni;

la 2° e 3° equazione danno $\alpha = \beta = -1$, inserendo nella 1° si ha $-2 = 1$, impossibile; quindi $d \notin \mathrm{Span}\{a, b\}$.

Determiniamo, in due modi, le condizioni sotto le quali $d = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \in \mathrm{Span}\{a, b\}$.

I° modo. $d \in\in \mathrm{Span}\{a, b\}$ se e solo se

il sistema $\left[\begin{array}{rr|l} 1 & 1 & d_1 \\ -1 & 0 & d_2 \\ 0 & -1 & d_3 \end{array}\right]$ ha soluzioni, se e solo se

il sistema $\left[\begin{array}{rr|l} 1 & 1 & d_1 \\ 0 & 1 & d_1 + d_2 \\ 0 & 0 & d_1 + d_2 + d_3 \end{array}\right]$ ha soluzioni, se e solo se

$d_1 + d_2 + d_3 = 0$.

(abbiamo applicato alle equazioni/righe le operazioni: $2° + 1°$; $3° + 2°$.)

2° modo. Essendo a, b lin. indip., $d \in \mathrm{Span}\{a, b\}$ se e solo se

$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$ sono lin. dip., se e solo se

$\begin{bmatrix} 0 \\ d_1 + d_2 + d_3 \\ 0 \end{bmatrix} = \underline{0}$, se e solo se

$d_1 + d_2 + d_3 = 0$.

(abbiamo applicato alle colonne le operazioni: $2° - 1°$; $3° - d_1 \cdot 1°$; $3° + d_3 \cdot 2°$).

Operazioni elementari e sottospazi.

Le operazioni elementari su vettori $a, b, c, \ldots$ di uno spazio vettoriale V

(1) sommare a un vettore un multiplo reale di un altro vettore

(2) moltiplicare un vettore per un numero non nullo:

(3) scambiare due vettori

lasciano invariato il sottospazio generato $\mathrm{Span}\{a, b, c, \ldots\}$.

E' piuttosto semplice provarlo per le operazioni (3) e (2).
Lo proviamo per l'operazione (1) di sostituire b con $b' = b +$ (multiplo di a).
E' chiaro che ogni combinazione lineare di $a, b', c, \ldots$ si può scrivere come combinazione lineare di $a, b, c, \ldots$ D'altro canto, essendo $b = b' +$ (multiplo di a), ogni combinazione lineare di $a, b, c, \ldots$ si può scrivere come combinazione lineare di $a, b', c, \ldots$ Quindi $\mathrm{Span}\{a, b', c \ldots\} = \mathrm{Span}\{a, b, c \ldots\}$.

Una procedura per determinare una base di un sottospazio generato in un $\mathbb{R}^n$ (descrizione informale).

Input: una sequenza di vettori di uno spazio vettoriale $\mathbb{R}^n$;
Output: una base del sottospazio generato dalla sequenza.

Casi
Se 1° vettore $= \underline{0}$, allora passa al successivo;
altrimenti, scegli una componente $\neq 0$ di 1° vettore e annullala [1] nei vettori successivi.
Fine casi.

Ripeti. Al termine, scrivi i vettori $\neq \underline{0}$ incontrati nel processo.

Commenti. E' chiaro che:

(1) il sottospazio generato dalla sequenza rimane invariato durante il processo;
(2) i vettori $\neq \underline{0}$ incontrati nel processo generano il sottospazio.

Si può intuire che questi vettori sono linearmente indipendenti. Non entriamo nei dettagli.

Spazio colonna, spazio riga, rango di una matrice

Spazio colonna e spazio riga di una matrice.

Una matrice $m \times n$ si può guardare come una sequenza di n colonne $c_1, \dots, c_n$ in $\mathbb{R}^m$ oppure come una sequenza di m righe $r_1, \dots, r_m$ in $\mathbb{R}^n$

$$\begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} = A = \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix}.$$

Gli spazi generati dalle colonne e dalle righe di A si dicono rispettivamente "spazio colonna" e "spazio riga" di A e si indicano con $\mathcal{C}(A)$ e $\mathcal{R}(A)$:

$$\mathcal{C}(A) = \mathrm{Span}\{c_1, \dots, c_n\} \subseteq \mathbb{R}^m;$$

$$\mathcal{R}(A) = \mathrm{Span}\{r_1, \dots, r_m\} \subseteq \mathbb{R}^n.$$

Esempio. $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix};$

$$\mathcal{C}(A) = \mathrm{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ (un piano vettoriale in } \mathbb{R}^3\text{);}$$

$$\mathcal{R}(A) = \mathrm{Span}\{ \begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \} = \mathbb{R}^2.$$

Lo spazio colonna e lo spazio riga, in generale, stanno in spazi vettoriali diversi; si possono comunque confrontare le loro dimensioni.

Esempio.

Indichiamo con $c_1, \dots, c_5 \in \mathbb{R}^4$ le colonne della matrice

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 1 & 2 & 5 & 4 & 1 \\ 2 & 1 & 4 & 5 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

[1] mediante un'opportruna operazione elementare del tipo (1)

Cerchiamo una base per $\mathcal{C}(A) = \text{Span}\{c_1, \ldots, c_5\}$. Per vedere meglio quali colonne scartare, usiamo operazioni elementari sulle colonne.

Usando la 1° colonna, annulliamo la 1° componente delle altre

$$\begin{array}{l} 3° - 1° \\ 4° - 2 \cdot 1° \\ 5° - 1° \\ \longrightarrow \end{array} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 4 & 2 & 0 \\ 2 & 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

Usando la 2° colonna, annulliamo la 2° componente delle altre

$$\begin{array}{l} 3° - 2 \cdot 2° \\ 4° - 2° \\ \longrightarrow \\ \end{array} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Indichiamo con $c'_1, \ldots, c'_5$ le nuove colonne. Si ha

$\mathcal{C}(A) = \text{Span}\{c'_1, c'_2, c'_3, c'_4, c'_5\} = \text{Span}\{c'_1, c'_2, c'_5\}$;

c'_1, c'_2, c'_5. lin. indip.

Dunque c'_1, c'_2, c'_5 è una base di $\mathcal{C}(A)$, e $\dim \mathcal{C}(A) = 3$.

Indichiamo con $r_1, r_2, r_3, r_4 \in \mathbb{R}^5$ le righe della matrice

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 1 & 2 & 5 & 4 & 1 \\ 2 & 1 & 4 & 5 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

Cerchiamo una una base per $\mathcal{R}(A) = \text{Span}\{r_1, \ldots, r_4\}$. Per vedere meglio quali righe scartare, usiamo operazioni elementari sulle righe.

Usando la 1° riga, annulliamo la 1° componente delle altre righe

$$\begin{array}{l} 2° - 1° \\ 3° - 2 \cdots 1° \\ \longrightarrow \\ \end{array} \qquad \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

Usiamo la 2° riga per annullare la 2° componente nelle righe successive

$$\begin{array}{l} 3° - \frac{1}{2} \cdot 2° \\ 4° - \frac{1}{2} \cdot 2° \\ \longrightarrow \\ \end{array} \qquad \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Usiamo la 3° riga per annullare la 5° componente nelle riga successiva

$$\begin{array}{l} 5° + 4° \\ \longrightarrow \end{array} \qquad \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Indichiamo con $r'_1, \ldots, r'_4$ le nuove righe. Si ha

$\mathcal{R}(A) = \mathrm{Span}\{r_1', r_2', r_3', r_4'\} = \mathrm{Span}\{r_1', r_2', r_3'\}$;

r_1', r_2', r_5'. lin. indip.

Dunque r_1', r_2', r_3' è una base di $\mathcal{R}(A)$, e $\dim \mathcal{R}(A) = 3$.

Rango di una matrice.

Quello che abbiamo visto nell'esempio non è un caso. In generale, si prova

Teorema. Per ogni matrice A, le dimensioni degli spazi colonna e riga di A coincidono, il loro valore comune si dice "rango" di A e si indica con $\mathrm{r}(A)$:

$$\dim \mathcal{C}(A) = \mathrm{r}(A) = \dim \mathcal{R}(A).$$

Osservazione.

$0 \leq \mathrm{r}(A) \leq \mathrm{Min}(m, n)$;

$\mathrm{r}(A) = 0$ se e solo se A è la matrice nulla;

per A quadrata $n \times n$, si ha $\mathrm{r}(A) = n$ se e solo se A è non singolare, cioè invertibile.

Prodotto righe per colonne e operazioni sulle righe/colonne

L'operazione di prodotto di righe per colonne è compatibile con le operazioni vettoriali sulle righe e sulle colonne, specificamente: per ogni tre righe a, a', a'' $(1 \times n)$, tre colonne b, b', b'' $(n \times 1)$, e numero reale r, si ha

(1) $\quad (a' + a'')b = a'b + a''b$

(2) $\quad a(b' + b'') = ab' + ab''$

(3) $\quad (r\,a)b = r\,(ab) = a(r\,b)$

(La (1) si prova come segue. Il 1° mebro e il 2° membro, per definizione, sono

$$\sum_{i=1}^{n}(a' + a'')_i b_i = \sum_{i=1}^{n}(a'_i + a''_i)b_i,$$

$$\sum_{i=1}^{n} a'_i b_i + \sum_{i=1}^{n} a''_i b_i;$$

sono uguali per la proprietà commutativa della somma e la proprietà distributiva. Analogamente per la (2). Si lascia al lettore di provare la (3).)

Da queste proprietà segue che più in generale per ogni tre righe a, a', a'' $(1 \times m)$, matrice B $(m \times n)$, tre colonne c, c', c'' $(n \times 1)$, e numero reale r, si ha

(1) $\quad (a' + a'')B = a'B + a''B$

(2) $\quad B(c' + c'') = Bc' + Bc''$

(3) $\quad B(r\,c) = r\,(Bc), \quad (r\,a)B = r\,(aB).$

Vederemo un poco più avanti come queste proprietà giochino nello studio dei sottospazi.

Insieme delle soluzioni di un sistema lineare - Esempi

Alcune equazioni lineari in 2 incognite

Consideriamo alcune equazioni lineari in 2 incognite x, y

$$ax + by = c$$

con a, b, c costanti in $\mathbb{R}$; ricordiamo che una soluzione dell'equazione è una coppia di numeri reali che sostituiti ordinatamente ad x e y rende vero l'uguale. Per ciascuna equazione, descriviamo l'insieme delle soluzioni in termini vettoriali. Fissato un riferimento nel piano, identifichiamo coppie ordinate con vettori rappresentati da segmnenti orientati uscenti dall'origine.

(2.1) $\quad x + y = 0;$

per ogni valore di una delle incognite, ad esempio la x, esiste uno ed un solo valore dell'altra, nell'esempio la y, che rende vero l'uguale;

le soluzioni dell'equazione sono date dalla regola

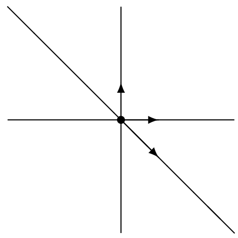
$$y = -x \text{ e } x \text{ libera;}$$

sono le coppie

$$\begin{bmatrix} x \\ -x \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad (x \in \mathbb{R});$$

formano il sottospazio Span$\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\}$, con base $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$,

rappresentato da una retta vettoriale per l'origine, con un suo riferimento



(2.1′) $\qquad x = 0;$

le soluzioni dell'equazione sono date dalla regola

$x = 0$ e y libera;

sono le coppie

$$\begin{bmatrix} 0 \\ y \end{bmatrix} = y \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad (y \in \mathbb{R});$$

formano ...

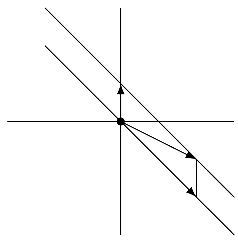
(2.1″) $\qquad x + y = 1;$

le soluzioni sono le coppie

$$\begin{bmatrix} x \\ -x+1 \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad (x \in \mathbb{R});$$

formano l'insieme Span$\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, che non è un sottospazio,

rappresentato dalla somma di una retta vettoriale e di un vettore fisso.



Un'equazione e un sistema lineari in 3 Incognite

Consideriamo un'equazione lineare e un sistema lineare in 3 incognite x, y, z e descriviamo l'insieme delle soluzioni in termini vettoriali. Fissato un riferimento nello spazio, identifichiamo terne ordinate con vettori rappresentati da segmenti orientati uscenti dall'origine.

(3.1) $\qquad x + y + z = 0;$

per ogni valore di due delle incognite, ad esempio la y e la z, esiste uno ed un solo valore della terza, nell'esempio la x, che rende vero l'uguale; le soluzioni dell'equazione sono date dalla regola

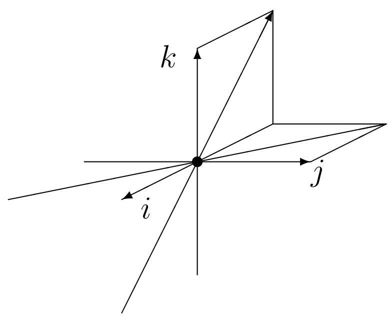
$x = -y - z$ e y, z libere;

sono le terne

$$\begin{bmatrix} -y-z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \qquad (y, z \in \mathbb{R});$$

formano il sottospazio Span$\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\}$, con base $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$,

rappresentato da un piano vettoriale per l'origine, con un suo riferimento.



Cambiando la scelta delle due incognite libere, si dà un'altra descrizione dello spazio delle soluzioni, che porta ad un'altra base.

(3, 2) $$\begin{cases} x + \ y + \ z = 0 \\ x + 2y + 3z = 0 \end{cases}$$

applichiamo la procedura di eliminazione

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

otteniamo il sistema equivalente

$$\begin{cases} x - \ z = 0 \\ y + 2z = 0 \end{cases}$$

le soluzioni del sistema sono date dalla regola

$x = z$, $y = -2z$, e z libera;

sono le terne

$$\begin{bmatrix} z \\ -2z \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \qquad (z \in \mathbb{R}),$$

formano il sottospazio Span$\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\}$, con base $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, rappresentato da una retta vettoriale per l'origine, con un suo riferimento.

Un sistema lineare in 4 incognite.

Consideriamo li sistema lineare nelle incognite x_1, x_2, x_3, x_4

$$(4.1) \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \end{cases}$$

applichiamo la procedura di eliminazione

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 & 0 \end{array}\right]$$

otteniamo il sistema equivalente

$$\begin{cases} x_1 - x_3 - 2x_4 = 0 \\ x_2 + 2x_3 + 3x_4 = 0 \end{cases}$$

le soluzioni del sistema sono date dalla regola

$$\begin{array}{ll} x_1 = & x_3 + 2x_4 \\ x_1 = & -2x_3 - 3x_4 \\ x_3 & \text{libera} \\ x_3 & \text{libera} \end{array}$$

sono le quaterne

$$\begin{bmatrix} x_3 + 2x_4 \\ -2x_3 - 3x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} = x_3\bar{s}_3 + x_4\bar{s}_4 \qquad (x_3, x_4 \in \mathbb{R}),$$

formano il sottospazio $\mathrm{Span}\{\bar{s}_3, \bar{s}_4\}$, con base $\bar{s}_3, \bar{s}_4$,

Sistemi lineari omogenei e sottospazi

Abbiamo visto che l'insieme delle soluzioni dell'equazione $x+y=1$ nelle incognite x, y non è un sottospazio di $\mathbb{R}^2$, ma si può dedurre dall'insieme delle soluzioni di $x+y=0$, che è un sottospazio. Questo fatto particolare è un'istanza di un fatto generale.

1 *Definizione.* Un sistema lineare si dice "omogeneo" se tutte le equazioni del sistema hanno termine noto nullo, cioè è del tipo

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{cases}, \quad \text{in breve} \quad Ax = \underline{0}.$$

Proposizione. L'insieme delle soluzioni di un sistema lineare omogeneo $Ax = \underline{0}$ è un sottospazio di $\mathbb{R}^n$ (A matrice $m \times n$).

2 *Dimostrazione.* (1) Se $s', s'' \in \mathbb{R}^n$ sono due soluzioni, allora anche $s' + s''$ è una soluzione. Infatti, da $As' = \underline{0}$ e $As'' = \underline{0}$, sommando termine a termine, si ottiene $As' + As'' = \underline{0} + \underline{0}$, da cui $A(s' + s'') = \underline{0}$. (2) Se $r \in \mathbb{R}$ e $s \in \mathbb{R}^n$ è una soluzione, allora

anche rs è una soluzione. Infatti, da $As = \underline{0}$, moltiplicando entrambi i membri per r, si ottiene $r(As) = r\underline{0}$, da cui $A(rs) = \underline{0}$. (3) $\underline{0}$ è una soluzione. Infatti $A\underline{0} = \underline{0}$.

L'insieme delle soluzioni di un sistema lineare non omogeneo non è mai un sottospazio ma, tranne nel caso in cui sia vuoto, si può sempre identificare, come insieme, con un sottospazio.

Proposizione. Se un sistema lineare $Ax = b$ ha una soluzione s^*, allora le soluzioni di $Ax = b$ sono tutti e soli i vettori del tipo

$$s^* + v,$$

dove v varia fra le soluzioni el sistema lineare omogeneo $Ax = \underline{0}$. In breve, indicati con $\mathcal{S}$ e con $\mathcal{S}_0$ gli insieme delle soluzioni di $Ax = b$ e di $Ax = \underline{0}$:

$$\mathcal{S} = s^* + \mathcal{S}_0.$$

Dimostrazione. (1) Da una parte, se v una soluzione di $Ax = \underline{0}$, allora $s^* + v$ è una soluzione di $Ax = b$. Infatti, da $As^* = b$ e $Av = \underline{0}$, sommando membro a membro, si ha $As^* + Av = b + \underline{0}$, da cui $A(s^* + v) = b$. (2) Dall'altra, se s è una soluzione di $Ax = b$, allora s si può scrivere $s = s^* + (s - s^*)$ e $s - s^*$ è una soluzione di $Ax = \underline{0}$.

Definzione. Diciamo "spazio nullo" di una matrice A, ed indichiamo con $\mathcal{N}(A)$, lo spazio delle soluzioni del sistema lineare omogeneo $Ax = \underline{0}$; in simboli, indicato con n il numero delle colonne di A,

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : \; Ax = \underline{0}\}.$$

Dimensione dello spazio nullo di una matrice

Abbiamo visto nei vari esempi che applicando una procedura di eliminazione a un sistema lineare omgeneo si trova una descrizione delle soluzioni che porta ad identificare una base dello spazio delle soluzioni. Un'analisi attenta della procedura porta a stabilire una relazione generale fra le dimensioni dello spazio riga e colonna di una matrice, cioè il rango di una matrice, e la dimensione dello spazio nullo della matrice.

Teorema 1. Data una matrice A con n colonne $c_1, \ldots, c_n$, si consideri il sistema lineare omogeneo $Ax = \underline{0}$ nelle n incognite $x_1, \ldots, x_n$. Se le r colonne $c_{i_1}, \ldots, c_{i_r}$ sono una base di $\mathcal{C}(A)$, allora

(1) ciascuna delle r incognite $x_{i_1}, \ldots, x_{i_r}$ è funzione delle altre $n-r$ incognite, che sono libere;

(2) le $n - r$ soluzioni ottenute assegnando a una incognita libera 1 e alle altre 0 sono una base dello spazio delle soluzioni.

(3) $\quad \dim(\mathcal{N}(A)) = n - \mathrm{r}(A)$.

3 *Dimostrazione, qualche aspetto.*

(1) Il sistema lineare omogeneo si scrive

$$x_1c_1 + \cdots + x_nc_n = \underline{0}.$$

equivalentemente, posto $I = \{i_1, \ldots, i_r\}$,

$$\sum_{i\in I} x_i c_i = -\sum_{j\notin I} x_j c_j;$$

essendo c_i $(i \in I)$ una base di $\mathcal{C}(A)$, per ogni sequenza di valori delle x_j $(j \notin I)$, esiste un'unica sequenza di valori delle x_i $(i \in I)$ che assieme alla prima sequenza dà una soluzione dell'equazione.

(3) Per la (2) si ha

$$\dim(\mathcal{N}(A)) = n - r = n - \dim(\mathcal{C}(A)) = n - \mathrm{r}(A).$$

Il teorema descrive teoricamente come la procedura di eliminazione, applicata a una sequenza di colonne base di $\mathcal{C}(A)$, porta ad identificare una base di $\mathcal{N}(A)$ e trae una conseguenza dulla dimensione. L'utilità pratica del Teorema consiste nel permettere di calcolare la dimensione di $\mathcal{N}(A)$ senza doverne determinare una base. (Esistono forme più fini del Teorema che hanno anche altre utilità).

Esempio (1).

$$A = \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

le 2 righe di A sono indipendenti, $\mathrm{r}(A) = \dim(\mathcal{R}(A)) = 2$, quindi

$$\dim(\mathcal{N}(A)) = 5 - 2 = 3;$$

Esempio (2).

$$A = \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 1 & 2 & 5 & 4 & 1 \\ 2 & 1 & 4 & 5 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

Abbiamo trovato (cfr. es. p.49) che $\mathrm{r}(A) = 3$, quindi

$$\dim(\mathcal{N}(A)) = 5 - 3 = 2.$$

Per la matrice trasposta

$$A^T = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 1 \\ 1 & 5 & 4 & 2 \\ 2 & 4 & 5 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

si ha $\mathrm{r}(A^T) = \mathrm{r}(A) = 3$, quindi

$$\dim(\mathcal{N}(A^T)) = 4 - 3 = 1.$$

3 Spazi vettoriali Euclidei

3.1 Esempi, Definizioni

$\mathcal{V}^n$ $(n \leq 3)$, lunghezza, ortogonalità

Fino ad avviso contrario, l'ambito del nostro discorso è uno qulasiasi fra retta piano, spazio, coi suoi vettori.

In geometria Euclidea è data una relazione primitiva di congruenza fra segmenti, che soddisfa certi assiomi che permettono di associare a ciascun segmento un numero reale, la lunghezza del segmento rispetto ad un segmento unità fissato.

Definiamo la lunghezza di un vettore come la lunghezza del segmento associato a un segmento orientato che dà il vettore,

$$(\text{lunghezza del vettore } AB) = (\text{lunghezza del segmento } AB);$$

questa definzione ha senso: un stesso vettore è dato da vari segmenti orientati, ma tutti questi segmenti orientati come segmenti hanno la stessa lunghezza.

Per ogni vettore v, poniamo

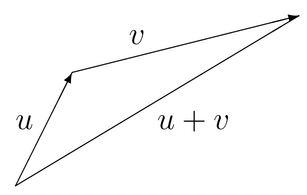
$$\|v\| = (\text{lunghezza di } v).$$

La funzione "lunghezza" da vettori a numeri reali è legate alle operazioni sui vettori dalle seguenti proprietà

$$\|v\| \geq 0; \quad \|v\| = 0 \text{ se e solo se } u = \underline{0}; \tag{1}$$
$$\|u + v\| \leq \|u\| + \|v\|; \tag{2}$$
$$\|rv\| = |r|\|v\| \tag{3}$$

per ogni $u, v \in V^n$ e $r \in \mathbb{R}$.

La seconda proprietà si può visualizzare come



E' equivalente all'affermazione "la lunghezza di un lato di un triangolo è minore o uguale alla somma delle lunghezze degli altri due lati"; per questo motivo si dice "disuguaglianza triangolare".

In gemetria Euclidea del piano si hanno una nozione primitiva di angolo fra due semirette ed una relazione primitiva di congruenza di angoli che soddisfano certi assiomi che permettono in particolare di definire la relazione di ortogonalità fra rette. Per indicare che due rette r ed s sono ortogonali, scriviamo $r \perp s$ e/o $s \perp r$.

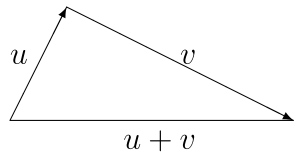
Queste nozioni e relazioni si trasferiscono allo spazio; in particolare, si dice che due rette r ed s dello spazio (eventualmente sghembe) sono ortogonali se e solo se, le rette r' ed s' ad esse parallele passanti per un punto T sono ortogonali nel piano che le contiene: $r \perp s$ se e solo se $r' \perp s'$.

Diciamo che due vettori non nulli u, v sono fra loro "ortogonali" e scriviamo $u \perp v$ se e solo se sono dati da segmenti orientati che stanno su rette fra loro ortogonali; per convenzione, diciamo inoltre che il vettore nullo è ortogonale ad ogni vettore:

$\underline{0} \perp v$ per ogni v.

La funzione lunghezza, la relazione di ortogonalità e l'operazione di somma di vettori sono legate dal teorema di Pitagora e dal suo inverso

$$u \perp v \quad \text{se e solo se} \quad \|u+v\|^2 = \|u\|^2 + \|v\|^2$$



$\mathcal{V}^n$ $(n \leq 3)$, prodotto scalare

Fatto. Esiste uno ed un solo prodotto, detto "prodotto scalare" ed indicato con $\cdot$, che a coppie di vettori associa numeri reali, che è compatibile con le operazioni sui vettori, commutativo, e tale che il prodotto di un vettore con sé stesso è il quadrato della sua lunghezza

$$\begin{aligned}
&(u' + u'') \cdot v = u' \cdot v + u'' \cdot v, \qquad \text{e analoga sul secondo fattore}\\
&(r\,u) \cdot v = r\,(u \cdot v), \qquad \text{e analoga sul secondo fattore}\\
&u \cdot v = v \cdot u\\
&v \cdot v = \|v\|^2
\end{aligned}$$

per ogni $u, v, u', u'', \ldots \in V^n$ e $r \in \mathbb{R}$.

Conseguenze. La lunghezza di un vettore si può ottenere come la radice quadrata del prodotto scalare del vettore con sè stesso, e due vettori sono ortogonali se e solo se il loro prodotto scalare è 0:

$$\begin{aligned}
&\|v\| = \sqrt{v \cdot v};\\
&u \perp v \quad \text{se e solo se} \quad u \cdot v = 0
\end{aligned}$$

per ogni $u, v \in V^n$. La prima affermazione segue direttamente dalla defінzione di prodotto scalare; la seconda affermazione si può provare come segue. Vale l'identità

$$\begin{aligned}
\|u+v\|^2 &= (u+v) \cdot (u+v)\\
&= u \cdot u + u \cdot v + v \cdot u + v \cdot v\\
&= \|u\|^2 + 2\,u \cdot v + \|v\|^2
\end{aligned}$$

in breve
$$\|u+v\|^2 = \|u\|^2 + 2\,u\cdot v + \|v\|^2;$$
le seguenti affermazioni sono equivalenti

$u \cdot v = 0$;

$\|u+v\|^2 = \|u\|^2 + \|v\|^2$;

$u \perp v$.

(la 1° equivalenza segue dall'identità e la seconda è il th di Pitagora e suo inverso); dunque $u \cdot v = 0$ se e solo se $u \perp v$.

Un vettore di lunghezza 1 si dice "versore":
$$u \text{ versore} \quad \text{se e solo se} \quad \|u\| = 1 \quad \text{se e solo se} \quad u\cdot u = 1.$$

$\mathcal{V}^2$, formule

Siano i, j due versori ortogonali in $\mathcal{V}^2$, in termini di prodotto scalare tali che
$$i\cdot i = j\cdot j = 1, \quad i\cdot j = 0.$$

Le proprietà del prodotto scalare permettono di esprimere il prodotto scalare di due combinazioni lineari di i, j in funzione dei prodotti scalari di i e j e quindi di calcolarlo. Ad esempio
$$\begin{aligned}(2\,i+3\,j)\cdot(4\,i+5\,j) &= (2\,i)\cdot(4\,i) + (2\,i)\cdot(5\,j) + (3\,j)\cdot(4\,i) + (3\,j)\cdot(5\,j)\\ &= (2\cdot 4)\,1 + (2\cdot 5)\,0 + (3\cdot 4)\,0 + (3\cdot 5)\,1\\ &= (2\cdot 4) + (3\cdot 5) = 23\end{aligned}$$

In generale, il prodotto scalare di due vettori è la somma dei prodotti delle coordinate dell'uno per le corrispondenti coordinate dell'altro: per ogni due vettori $u, v \in \mathcal{V}^2$, posto
$$u = u_1\,i + u_2\,j, \quad v = v_1\,i + v_2\,j,$$
si ha
$$u\cdot v = u_1v_1 + u_2v_2.$$
Dunque, la lunghezza di un vettore e la condizione di ortogonalità in coordinate divengono
$$\begin{aligned}&\|u\| = \sqrt{u_1^2 + u_2^2}\\ &u \perp v \quad \text{se e solo se} \quad u_1v_1 + u_2v_2 = 0.\end{aligned}$$

$\mathcal{V}^3$, formule

Siano i, j, k versori a due a due ortogonali in $\mathcal{V}^3$, in termini di prodotto scalare tali che
$$i\cdot i = j\cdot j = k\cdot k = 1, \quad i\cdot j = i\cdot k = j\cdot k = 0.$$

Le proprietà del prodotto scalare permettono di esprimere il prodotto scalare di due combinazioni lineari di i, j, k in funzione dei prodotti scalari di i, j, k e quindi di calcolarlo. Il prodotto scalare di due vettori risulta essere la somma dei prodotti delle coordinate dell'uno per le corrispondenti coordinate dell'altro: per ogni due vettori $u, v \in V^3$, posto

$$u = u_1\, i + u_2\, j + u_3\, k, \quad v = v_1\, i + v_2\, j + v_3\, k,$$

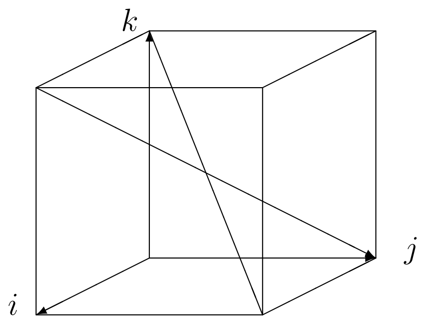
si ha

$$u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Dunque, la lunghezza di un vettore e la condizione di ortogonalità in coordinate divengono

$$\begin{aligned} &\|u\| = \sqrt{u_1^2 + u_2^2 + u_3^2} \\ &u \perp v \quad \text{se e solo se} \quad u_1 v_1 + u_2 v_2 + u_3 v_3 = 0. \end{aligned}$$

Esempio. In un cubo unitario, ogni diagonale lunga ha lunghezza $\sqrt{3}$ e ogni due diagonali lunghe non sono ortogonali.



Ad esempio, per le diagonali lunghe uscenti dai punti finali di k e di j,

$$d_1 = i + j - k, \quad d_2 = i - j + k,$$

si ha

$$\|d_1\| = \sqrt{1^2 + 1^2 + (-1)^2} = \sqrt{3}$$

e analogamente per d_2; e

$$d_1 \cdot d_2 = 1 \cdot 1 + 1 \cdot (-1) + (-1) \cdot 1 = -1 \neq 0, \quad \text{cioè} \quad d_1 \not\perp d_2.$$

Spazi vettoriali Euclidei

Definizione. Uno "spazio vettoriale Euclideo" è uno spazio vettoriale V con un'operazione, detta "prodotto scalare" e denotata con $\cdot$, che ad ogni $u, v \in V$ associa un numero

$u \cdot v \in \mathbb{R}$ che è compatibile con le operazioni vettoriali, commutativa, e tale che il quadrato scalare di un vettore $\neq \underline{0}$ sia positivo:

$$
\begin{array}{lll}
(1) & (u' + u'') \cdot v = u' \cdot v + u'' \cdot v, & \text{analoga sul secondo fattore} \\
(2) & (r\, u) \cdot v = r\, (u \cdot v), & \text{analoga sul secondo fattore} \\
(3) & u \cdot v = v \cdot u & \\
(4) & v \cdot v \geq 0, \quad \text{con } v \cdot v = 0 \ \text{ solo per } v = \underline{0} &
\end{array}
$$

per ogni $u, v, u', u'', \ldots \in V$ e $r \in \mathbb{R}$.
Si definisce "lunghezza" di un vettore la radice quadrata del quadrato scalare del vettore, e due vettori si dicono "ortogonali" se e solo se il loro prodotto scalare è 0:

$$
\begin{array}{l}
\|v\| = \sqrt{v \cdot v}; \\
u \perp v \quad \text{se e solo se} \quad u \cdot v = 0
\end{array}
$$

Spazio vettoriale Euclideo $\mathbb{R}^n$

Sia n un intero positivo fissato. Nello spazio vettoriale $\mathbb{R}^n$ consideriamo il prodotto che associa a due $n-$ple un numero reale dato da

$$u \cdot v = u^T v = \sum_{i=1}^{n} u_i v_i$$

(u e v sono identificati con vettori colonna $n \times 1$). Questo prodotto soddisfa le condizioni (1), (2) per le proprietà del prodotto di righe per colonne rispetto alle operazioni vettoriali e soddisfa la (3) per la proprietà commutativa del prodotto di numeri reali. Inoltre,

$$v \cdot v = \sum_{i=1}^{n} v_i^2 \geq 0, \qquad \text{ed è} = 0 \ \text{ solo per } v = \underline{0},$$

in quanto in $\mathbb{R}$ i quadrati sono ≥ 0, solo 0 ha quadrato 0, le somme di sequenze di numeri ≥ 0 sono ≥ 0, e fra queste solo le somme di sequenze di 0 sono 0.

Lo spazio vettoriale $\mathbb{R}^n$, con questo prodotto scalare, si dice "spazio vettoriale Euclideo $n-$dimensionale standard".

Per definizione, la lunghezza di un vettore e la relazione di ortogonalità fra due vettori sono date da

$$\|v\| = \sqrt{\sum_{i=1,\cdots,n} v_i^2}$$

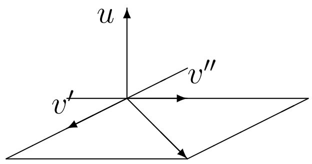
$$u \perp v \quad \text{se e solo se} \quad \sum_{i=1,\cdots,n} u_i v_i = 0.$$

3.2 Ortogonalità

Di regola, nel seguito identifichiamo i vettori geometrici con segmenti orientati uscenti da un punto fissato. Mostriamo come alcune proprietà e costruzioni sui vettori geometrici si estendono a proprietà e costruzioni in spazi vettoriali euclidei qualsiasi.

Proprietà.

Negli spazi vettoriale Euclidei geometrici, la relazione di ortogonalità possiede le seguenti proprietà: se un vettore è ortogonale a un secondo vettore, allora il secondo è ortogonale al primo; se un vettore è ortogonale a due vettori in direzioni diverse, allora il vettore è ortogonale a tutti i vettori sul piano dei due vettori.



Queste proprietà valgno in generale, specificamente:

Proposizione. Sia V uno spazio vettoriale Euclideo. Allora

se $u \perp v$ allora $v \perp u$

se $u \perp v', v''$ allora $u \perp (\alpha' v' + \alpha'' v'')$

per ogni $u, v, v', v'' \in V$ e $\alpha', \alpha'' \in \mathbb{R}$.

Infatti:

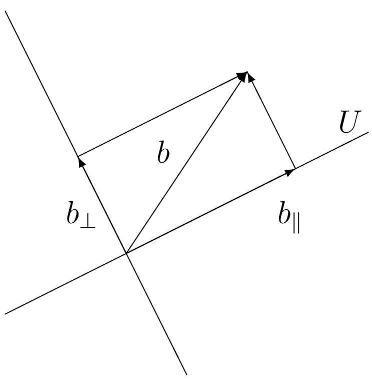
$u \perp v$ significa $u \cdot v = 0$ implica (per la (3)) $v \cdot u = 0$ significa $v \perp u$.

$u \perp v', v''$ significa $u \cdot v' = u \cdot v'' = 0$ implica (per le (1),(2)) $u \cdot (\alpha' v' + \alpha'' v'') = \alpha'(u \cdot v') + \alpha''(u \cdot v'') = \alpha' 0 + \alpha'' 0 = 0$, significa $u \perp (\alpha' v' + \alpha'' v'')$.

Proiezioni ortogonali.

Negli spazi vettoriali geometrici, si può effettuare la proiezione ortogonale di un vettore su una retta vettoriale.

Fatto. Siano U una retta vettoriale in uno spazio vettoriale Euclideo $\mathcal{V}^n$ $(n = 2, 3)$. Ogni vettore b si scompone in un unico modo come somma di un vettore $b_{\|}$ in U ed un vettore $b_{\perp}$ ortogonale a U; la componente $b_{\|}$ si dice "proiezione ortogonale" di b su U.

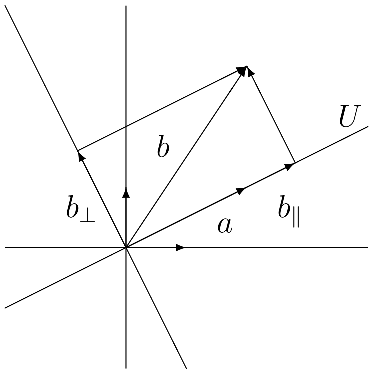


Di seguito mostriamo come questo fatto possa essere dedotto, e una formula esplicita possa essere ricavata, usando solo il prodotto scalare.

Per fissare le idee, identificato $\mathcal{V}^2$ con $\mathbb{R}^2$ mediante un riferimento di versori ortogonali, condideriamo

il sottospazio U delle soluzioni di $x - 2y = 0$, che ha una base $a = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$;

il vettore $\quad b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.



Consideriamo le condizioni

$$\begin{cases} b = b_\| + b_\perp \\ b_\| \in U \\ b_\perp \perp U \end{cases};$$

essendo a una base di U, le condizioni si possono riscrivere

$$\begin{cases} b = b_\| + b_\perp \\ b_\| = r\,a \\ b_\perp \cdot a = 0 \end{cases}, \text{ dove } r \text{ è un'incognita in } \mathbb{R};$$

inserendo la 2° uguaglianza nella 1° si ha

$$b = r\,a + b_\perp;$$

moltiplicando a per entrambi i membri ed usando la 3° condizione si ha

$$a \cdot b = a \cdot (r\,a + b_\perp)$$

$$a \cdot b = r\,(a \cdot a) + a \cdot b_\perp$$

$$a \cdot b = r\,(a \cdot a);$$

essendo $a \neq \underline{0}$, l'equazione ha l'unica soluzione

$$r = \frac{a \cdot b}{a \cdot a}.$$

Abbiamo che il sistema di condizioni ha un'unica soluzione, data da

$$b_\| = r\,a, \text{ con } r = \frac{a \cdot b}{a \cdot a},$$

$$b_\perp = b - b_\|,$$

dove a è una base di U.

Si lascia al lettore di verificare che il valore dell'espressione trovata per $b_{\parallel}$ non dipende dalla base a e che $b_{\perp} = b - b_{\parallel}$ è ortogonale ad U.

Nell'esempio, $a \cdot b = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 7$ e $a \cdot a = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5$, quindi

$$b_{\parallel} = \tfrac{7}{5} \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Poiché abbiamo usato solo le proprietà del prodotto scalare, abbiamo provato una proposizione valida in ogni spazio vettoriale Euclideo, che permette di dare una definzione di proiezione ortogonnale e una relativa formula. Precisamente:

Proposizione. Siano V uno spazio vettoriale Euclideo, U un sottospazio 1-dimensionale di V e $b \in V$. Allora:

(1) b si scompone in un unico modo come somma di un vettore $b_{\parallel} \in U$ ed un vettore $b_{\perp}$ ortogonale a U;

(2) se $a \in U$ è una base di U, allora $b_{\parallel} = r\,a$, con $r = \dfrac{a \cdot b}{a \cdot a}$.

La componente $b_{\parallel}$ di b in U si dice "proiezione ortogonale" di b su U.

Indipendenza lineare, Basi, coordinate

Negli spazi vettoriali geometrici, due vettori non nulli fra loro ortogonali hanno direzioni diverse e così sono linearmente indipendenti e tre vettori non nulli a due a due ortogonali non sono complanari e così sono linearmente indipendenti. In generale, si ha

Proposizione. In uno spazio vettoriale Euclideo, ogni sequenza di vettori $v_1, v_2, \ldots, v_n$ a due a due ortogonali, non nulli, è linearmente indipendente.

Dimostrazione. Consideriamo un'uguaglianza

$$(*) \qquad x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = \underline{0}$$

con $x_1, \ldots, x_n \in \mathbb{R}$; moltiplicando v_1 per entrambi i membri e usando l'ipotesi che v_1 sia ortogonale a tutti gli altri si ha

$$v_1 \cdot (x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = v_1 \cdot \underline{0}$$
$$x_1(v_1 \cdot v_1) + x_2(v_1 \cdot v_2) + \cdots + x_n(v_1 \cdot v_n) = 0$$
$$x_1(v_1 \cdot v_1) = 0;$$

usando l'ipotesi che $v_1 \neq \underline{0}$, si trova $x_1 = 0$. Allo stesso modo, moltiplicando i vari v_i per entrambi i membri della $(*)$ si trova che i vari x_i sono $= 0$.

In generale, calcolare le coordinate rispetto ad una base è complicato, richiede di risolvere un sistema lineare (con tante equazioni quante incognite). Invece, calcolare le coordinate rispetto ad una base di vettori fra loro ortrogonali è semplice:

Proposizione. Sia $v_1, v_2, \ldots, v_n$ una sequenza di n vettori a due a due ortogonali, non nulli, in uno spazio vettoriale Euclideo V di dimensione n. Allora $v_1, v_2, \ldots, v_n$ è una base di V e le coordinate di un $b \in V$ rispetto ad essa sono date da

$$\frac{b \cdot v_i}{v_i \cdot v_i}, \quad (i = 1, 2, \ldots, n).$$

Dimostrazione. Per la Proposizione precedente, gli n vettori sono linearmente indipendenti e, essendo in uno spazio vettoriale V di dimensione n, sono una base di V. Condideriamo l'uguaglianza

$$x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = b$$

moltiplicando v_i per entrambi i membri e usando l'ipotesi di mutua ortogonalità, si ha

$$v_i \cdot (x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = v_i \cdot b$$

$$x_1(v_i \cdot v_1) + x_2(v_i \cdot v_2) + \cdots + x_n(v_i \cdot v_n) = v_i \cdot b$$

$$x_i(v_i \cdot v_i) = v_i \cdot b;$$

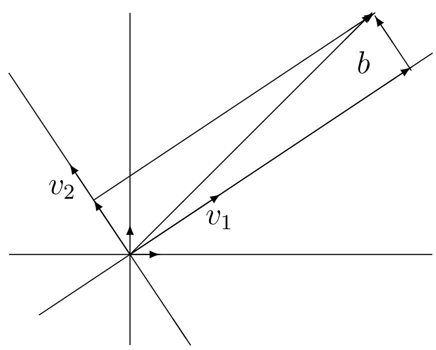
essendo $v_i \neq \underline{0}$, si ricava $x_i = \dfrac{v_i \cdot b}{v_i \cdot v_i}$.

Esempio. In $\mathcal{V}^2$, identificato con $\mathbb{R}^2$ mediante un riferimento di due versori ortogonali,

$v_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $v_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, sono fra loro ortogonali, quindi una base di $\mathcal{V}^2$;

le coordinate di $b = \begin{bmatrix} 8 \\ 8 \end{bmatrix}$ rispetto alla base sono $\dfrac{v_1 \cdot b}{v_1 \cdot v_1} = \dfrac{40}{13}$ e $\dfrac{v_2 \cdot b}{v_2 \cdot v_2} = \dfrac{8}{13}$;

quindi $\frac{40}{13} \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \frac{8}{13} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \end{bmatrix}$.



Lunghezza, proprietà.

Sia V uno spazio vettoriale Euclideo. Ricordiamo che si definisce "lunghezza" di un vettore la radice quadrata del prodotto scalare del vettore con sè stesso

$$\|v\| = \sqrt{v \cdot v}.$$

Le proprietà della lunghezza di vettori che abbiamo evidenziato negli spazi vettoriali geometrici valgono in generale:

Teorema. In ogni spazio vettoriale Euclideo V,

$$\begin{aligned}&\|v\| \geq 0; \quad \|v\| = 0 \text{ se e solo se } u = \underline{0};\\ &\|u+v\| \leq \|u\| + \|v\|;\\ &\|rv\| = |r|\|v\|\end{aligned}$$

per ogni $u, v \in V$ e $r \in \mathbb{R}$.

La 1° segue direttamente dalle proprietà del prodotto scalare. La 2°, disuguaglianza triangolare, è la proprietà più profonda. A noi interessa la 3°; segue dalla definizione e dalle proprietà del prodotto scalare:

$$\|rv\| = \sqrt{(rv) \cdot (rv)} = \sqrt{r^2(v \cdot v)} = |r|\,\sqrt{v \cdot v} = |r|\,\|v\|.$$

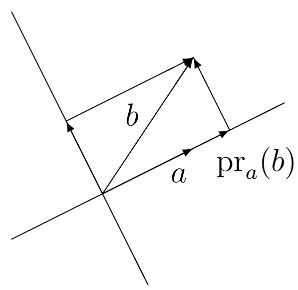
Ricordiamo che un vettore si dice "versore" se ha lunghezza 1. Ogni vettore non nullo $v \neq \underline{0}$, diviso per la sua lunghezza, diviene un versore, infatti

$$\|\frac{1}{\|v\|}v\| = \frac{1}{\|v\|}\|v\| = 1.$$

Terminologia, notazioni.

Abbiamo visto che, per ogni sottospazio $1-$ dimensionale $U \subseteq V$, ogni vettore b si scrive in un unico modo come somma di un vettore $\in U$, che si dice "proiezione ortogonale" di b su U, e un vettore $\perp U$. In altri termini, per ogni vettore $a \neq \underline{0}$, ogni vettore b si scrive in un unico modo come somma di un vettore multiplo di a, che si dice "proiezione ortogonale" di b su a, e un vettore $\perp a$. La prima formulazione è in linea di principio migliore della seconda, in quanto l'operazione di proiezione ortogonale dipende solo dal sottospazio, ma la seconda formulazione risulta spesso più comoda. Indichiamo la proiezione ortogonale di b su a con $\mathrm{pr}_a(b)$; dunque si ha

$\mathrm{pr}_a(b) = r\,a$, con $r = \frac{a \cdot b}{a \cdot a}$.



Nel seguito, al posto di dire “i vettori ... sono a due a due ortogonali” diremo un po’ più in breve “la sequenza dei vettori ... è ortogonale”. Inoltre, al posto di dire “i vettori ... sono a due a due ortogonali e di lunghezza 1” diremo un po’ più in breve “la sequenza dei vettori ... è ortonormale”. Poichè ogni vettore non nullo si può normalizzare, da ogni base ortogonale si può ricavare una base orrtonormale.

Abbiamo anche visto che le coordinate di un vettore v rispetto a una base ortogonale $a_1, \ldots, a_n$ di V, sono i prodotti scalari di v con gli a_i sui quadrati scalari degli a_i :

$$v = \sum_{i=1}^{n} \frac{a_i \cdot v}{a_i \cdot a_i} a_i;$$

possiamo anche dire che ogni vettore v si scompone come somma delle sue proiezioni ortogonali sui vettori della base:

$$v = \sum_{i=1}^{n} \mathrm{pr}_{a_i}(b).$$

Basi ortogonali.

Fatti.

Ogni spazio vettoriale Euclideo geometrico $\mathcal{V}^n$ $(n = 1, 2, 3)$ possiede qualche base ortogonale, e quindi qualche base ortonormale. Al solito, identifichiamo i vettori con segmenti orientati uscenti da un punto fissato. Più in dettaglio:

$\mathcal{V}^1$: ogni vettore $b \neq \underline{0}$ è una base ortogonale; ci sono esattamente due versori, uno opposto dell’altro; ciscuno dei due versori è una base ortonormale; non ci son altre basi ortonormali.
$\mathbb{R}$: ogni $b \neq 0$ è una base ortogonale; 1 è una base ortonormale, -1 è una base ortonormale, non ce ne sono altre.

$\mathcal{V}^2$: per ogni retta passante per O, esiste un’unica retta per O ad essa ortogonale; comunque scelti due vettori b_1, b_2 diversi da $\underline{0}$ sulle due rette, si ha una base ortogonale; tutte le basi ortogonali si ottengono così; su ciascuna retta si hanno esattamente due versori, dunque a partire da due rette ortogonali si ottengono esattamente 4 basi ortonormali.
$\mathbb{R}^2$: un esempio di base ortogonale:

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

normalizzando, si ha una base ortonormale:

$$\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}, \begin{bmatrix} -\frac{4}{\sqrt{20}} \\ \frac{2}{\sqrt{20}} \end{bmatrix}$$

(si può semplificare).

$\mathcal{V}^3$: per ogni retta passante per O, e per ciascuna delle infinite rette per O ad essa ortogonali, esiste un’unica retta per O ortogonale ad esse; comunque scelti tre vettori b_1, b_2, b_3 diversi da $\underline{0}$ sulle tre rette, si ha una base ortogonale; tutte le basi ortogonali si ottengono così; su ciascuna retta si hanno esattamente due versori, dunque a partire da tre rette a due a due ortogonali si ottengono esattamente 8 basi ortonormali.
$\mathbb{R}^3$: un esempio di base ortogonale:

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}$$

normalizzando si ha una base ortonormale ...

Per ogni n intero positivo fissato, nello spazio vettoriale Euclideo $\mathbb{R}^n$ si ha che la sequenza dei vettori unità $e_1, e_2, \ldots, e_n$ è una base ortonormale. Infatti,

$$e_i \cdot e_i = \sum_{h=1}^{n} (e_i)_h^2 = (e_i)_i^2 = 1^2 = 1$$

$$e_i \cdot e_j = \sum_{h=1}^{n} (e_i)_h (e_j)_h = (e_i)_i (e_j)_i + (e_i)_j (e_j)_j = 1 \cdot 0 + 0 \cdot 1 = 0$$

Problemi:

Sia U il sottospazio di $\mathbb{R}^3$ costituito dalle soluzioni dell'equazione $x - y - 2z = 0$. Geometricamente, U è un piano vettoriale, quindi posside basi ortogonali (e ortonormali). Come se ne può costruire una?

Sia U il sottospazio di $\mathbb{R}^4$ costituito dalle soluzioni dell'equazione $x_1 - x_2 - 2\,x_3 - 3\,x_4 = 0$. U posside basi ortogonali?

Un qualsiasi spazio vettoriale Euclideo possiede basi ortogonali? come si possono costruire?

Teorema (Gram-Schmidt). Sia $v_1, v_2, \ldots, v_p$ una sequenza lin. indip. in uno spazio vett. Euclideo. Allora: esiste un'unica sequenza $u_1, u_2, \ldots, u_p$ ortogonale tale che per ogni $i = 1, 2, \ldots, p$

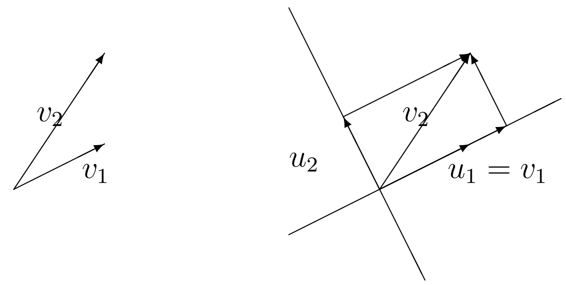
3 $u_1, \ldots, u_i$ è una base di $\mathrm{Span}\{v_1, \ldots, v_i\}$ in cui la $i-$ma coord. di v_i è 1;

in particolare:

$$\begin{aligned} v_1 &= u_1; \\ v_2 &= \mathrm{pr}_{u_1}(v_2) + u_2; \\ v_3 &= \mathrm{pr}_{u_1}(v_3) + \mathrm{pr}_{u_2}(v_3) + u_3; \\ &\vdots \\ v_p &= \sum_{j=1}^{p-1} \mathrm{pr}_{u_j}(v_p) + u_p. \end{aligned}$$

Esplicitamente, la sequenza $u_1, u_2, \ldots, u_p$ è

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \mathrm{pr}_{u_1}(v_2) \\ u_3 &= v_3 - \mathrm{pr}_{u_1}(v_3) + \mathrm{pr}_{u_2}(v_3). \\ &\vdots \\ u_p &= v_p - \sum_{j=1}^{p-1} \mathrm{pr}_{u_j}(v_p). \end{aligned}$$



In particolare, dal Teorema segue che

ogni spazio vett. Euclideo di dim. finita possiede qualche base ortogonale.

Commenti. E' chiaro che una sequenza $u_1, u_2, \ldots, u_p$ soddisfacente le date condizioni, se esiste, deve soddisfare la prima serie di uguaglianze, quindi deve essere data esplicitamente dalle seconda serie di uguaglianze, quindi è unica. In sostanza, bisogna mostrare che queste espressioni sono ben definite (cioè $u_i \neq \underline{0}$ per ogni i) e definiscono una sequenza ortogonale. Non lo facciamo.

Esempio Sia U il sottospazio di $\mathbb{R}^3$ costituito dalle soluzioni dell'equazione

$x - y - 2\,z = 0.$

Una base di U è

$(1, 1, 0),\ (2, 0, 1).$

A questa sequenza corrisponde la base ortogonale di U

$$\begin{aligned}
u_1 &= (1, 1, 0) \\
u_2 &= (2, 0, 1) - \mathrm{pr}_{(1,1,0)}(2, 0, 1) \\
&= (2, 0, 1) - \tfrac{2}{2}(1, 1, 0) \\
&= (1, -1, 1)
\end{aligned}$$

Esempio Sia U il sottospazio di $\mathbb{R}^4$ costituito dalle soluzioni dell'equazione

$x_1 - x_2 - 2\,x_3 - 3\,x_4 = 0$

Una base di U è

$(1, 1, 0, 0),\ (2, 0, 1, 0),\ (3, 0, 0, 1).$

A questa sequenza corrisponde la base ortogonale di U

$$\begin{aligned}
u_1 &= (1, 1, 0, 0) \\
u_2 &= (2, 0, 1, 0) - \mathrm{pr}_{(1,1,0,0)}(2, 0, 1, 0) \\
&= (2, 0, 1, 0) - \tfrac{2}{2}\,(1, 1, 0, 0) \\
&= (1, -1, 1, 0) \\
u_3 &= (3, 0, 0, 1) - \mathrm{pr}_{(1,1,0,0)}(3, 0, 0, 1) - \mathrm{pr}_{(1,-1,1,0)}(3, 0, 0, 1) \\
&= (3, 0, 0, 1) - \tfrac{3}{2}\,(1, 1, 0, 0) - \tfrac{3}{3}\,(1, -1, 1, 0) \\
&= (\tfrac{1}{2}, -\tfrac{1}{2}, -1, 1).
\end{aligned}$$

Matrici ortogonali

Il Teorema secondo il quale le colonne di una matrice quadrata sono lin. indip. se e solo se le righe della matrice sono lin. indip. se e solo se la matrice è invertibile ha il seguente analogo ortogonale:

Teorema. Per ogni matrice quadrata A le seguenti affermazioni sono equivalenti

(1) la sequenza $c_1, \ldots, c_n$ delle n colonne di A è ortonormale;

3 (2) la sequenza $r_1, \ldots, r_n$ delle n righe di A è ortonormale;

(3) la matrice A è invertibile con inversa la sua trasposta:

$$A^T A = \mathrm{I}_n = AA^T.$$

Esempio.

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{4}{\sqrt{20}} \\ \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{20}} \end{bmatrix}$$

Le colonne sono una base ortonormale di $\mathbb{R}^2$, le righe sono una base ortonormale di $\mathbb{R}^2$, e la matrice è invertibile con inversa la sua trasposta

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{4}{\sqrt{20}} & \frac{2}{\sqrt{20}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{4}{\sqrt{20}} \\ \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{20}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{4}{\sqrt{20}} \\ \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{20}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{4}{\sqrt{20}} & \frac{2}{\sqrt{20}} \end{bmatrix}.$$

Dimostrazione.

La (1) equivale al sistema di uguaglianze

$$c_i^T \cdot c_j = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases} \quad \text{per ogni } i, j = 1, \ldots, n$$

che equivale all'uguaglianza

$$\begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_n^T \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{bmatrix}$$

cioè

$$A^T A = \mathrm{I}_n;$$

la (1) inoltre, implica che (le righe sono lin. indip. e quindi che) A è invertibile; quindi la (1) equivale alla (3).

Analogamente si prova che la (2) equivale alla (3).

Definizione. Una matrice quadrata che soddisfa una (quindi ciascuna) delle condizioni del teorema si dice "matrice ortogonale".

4 Determinanti

4.1 Introduzione

Consideriamo una matrice quadrata $n \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Abbiamo visto che le righe di A sono linearmente indipendenti se e solo se le colonne di A sono linearmente indipendenti se e solo se A è invertibile; abbiamo detto “non singolare” una matrice che ha una (quindi ha ciascuna) di queste proprietà, dciamo “singolare” una matrice che non ha una (quindi non ha alcuna) di queste proprietà, Ci chiediamo

sotto quali condizioni sugli elementi a_{ij} succede che A è singolare?

Il caso $n = 1$ è ovvio: $A = [a_{11}]$ è singolare se e solo se $a_{11} = 0$. Di seguito discutiamo prima il caso $n = 2$ e poi quello generale.

4.2 Determinante di matrici 2×2

Definizione per proprietà, formula.

Proposizione. Esiste un'unica funzione “determinante”

$$\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad [a, b] \to \det[a, b]$$

che è compatibile con le operazioni su una colonna fissata l'altra, che sulle coppie di colonne uguali si annulla e che sulla coppia dei vettori unità vale 1:

1) $\det[a' + a'', b] = \det[a', b] + \det[a'', b]$

   analoga su 2° colonna

2) $\det[ra', b] = r \det[a', b]$

   analoga su 2° colonna

3) $\det[a, a] = 0$

4) $\det[e_1, e_2] = 1$

per ogni $a', a'', b, \ldots \in \mathbb{R}^2$ ed ogni $r \in \mathbb{R}$.

Osservazioni.

- Combinando la 1) e la 2) si ha che per ogni $a', a'', b \in \mathbb{R}^2$ ed ofgni $r', r'' \in \mathbb{R}$,

  $\det[r'a' + r''a'', b] = r' \det[a', b] + r'' \det[a'', b]$.

- Per ogni $a, b \in \mathbb{R}^2$ si ha

  $0 = \det[a + b, a + b]$

$$= \det[a, a] + \det[a, b] + \det[b, a] + \det[b, b]$$
$$= \det[a, b] + \det[b, a];$$

quindi, scambiando le colonne il determinante cambia segno:

$$\det[b, a] = -\det[a, b].$$

Le proprietà 1) ... 4) permettano di calcolare il determinate di ogni matrice 2×2.

Un esempio:

$$\det \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix} = 2 \det \begin{bmatrix} 1 & 4 \\ 0 & 5 \end{bmatrix} + 3 \det \begin{bmatrix} 0 & 4 \\ 1 & 5 \end{bmatrix}$$
$$= 2 \left(4 \det \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + 5 \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) + 3 \left(4 \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 5 \det \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right)$$
$$= 2 \cdot 5 - 3 \cdot \cdot 4 = -2.$$

In generale, il determinate di una matrice 2×2 è il prodotto degli elementi sulla diagonale discendente meno il prodotto degli elementi sulla diagonale ascendente:

$$\det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = a_1 b_2 - a_2 b_1.$$

Si ha

$$\det(A) = \det(A^T),$$

quindi

il determinante è equivalentemente definito da proprietà per righe.

Singolarità di matrici e annullamento del determinante

L'operazione elementare "sommare alla 2° colonna un multiplo della 1° colonna" lascia invariato il determinante. Infatti:

$$\det[a, b + ra] = \det[a, b] + r \det[a, a] = \det[a, b]$$

(lo stesso vale per l'operazione di sommare alla 1° colonna un multiplo della 2°).

Proposizione. Per ogni matrice A,

se A è singolare allora $\det A = 0$, e

se A è non singolare allora $\det A \neq 0$.

In altri termini,

A è singolare se e solo se $\det(A) = 0$.

Dimostrazione. Sia $A = [a, b]$.

Se A è singolare, allora a, b sono linearmente dipendenti; allora $a = \underline{0}$ oppure $b = ra$ per qualche $r \in \mathbb{R}$; allora $\det(A) = \det[\underline{0}, b] = 0$ oppure $\det(A) = \det[a, ra] = r \det[a, a] = 0$; in ogni caso, $\det(A) = 0$.

Se A è non singolare, allora a, b sono linearmente indipendenti; allora $a \neq \underline{0}$, supponiamo per semplicità che $a_1 \neq 0$; applicando un'opportuna operazione elementare, si ha

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \longrightarrow \begin{bmatrix} a_1 & 0 \\ a_2 & b_2' \end{bmatrix};$$

inoltre, deve esssere $b_2' \neq 0$. Dunque,

$$\det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \det \begin{bmatrix} a_1 & 0 \\ a_2 & b_2' \end{bmatrix} = a_2 b_2' \neq 0.$$

Applicazione.

$$A = \begin{bmatrix} k & k+1 \\ k+2 & 1 \end{bmatrix} \qquad (k \text{ parametro} \in \mathbb{R}).$$

$\det(A) = -k^2 - 2k - 2$, $\Delta = 4 - 8 = -4 < 0$, $\det(A) \neq 0$ per ogni $k \in \mathbb{R}$;
quindi A è non singolare per ogni $k \in \mathbb{R}$.

(A è singolare esattamente per i due valori $k_1 = -1 + i$ e $k_2 = -1 - i$ di k in $\mathbb{C}$.)

Sistemi lineari e determinanti

Proposizione. E' dato il sistema lineare di due equazioni in 2 incognite x_1, x_2,

$$x_1 a + x_2 b = p,$$

con a, b, p costanti in $\mathbb{R}^2$.

1) Se $\det[a, b] = 0$, allora il sistema non ha soluzioni oppure ne ha infinite;

2) Se $\det[a, b] \neq 0$, allora il sistema ha una ed una sola soluzione; una formula per la soluzione è

$$x_1 = \frac{\det[p\ b]}{\det[a\ b]}$$

$$x_2 = \frac{\det[a\ p]}{\det[a\ b]}.$$

Esempio.

$$\begin{cases} 2x + 5y = 1 \\ 3x + 7y = 8 \end{cases} \qquad \left[\begin{array}{cc|c} 2 & 5 & 1 \\ 3 & 7 & 8 \end{array}\right]$$

$$\det \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix} = -1 \neq 0, \text{ dunque il sistema ha un'unica soluzione,}$$

$$x = \frac{\det \begin{bmatrix} 1 & 5 \\ 8 & 7 \end{bmatrix}}{-1} = \frac{-33}{-1} = 33$$

$$y = \frac{\det \begin{bmatrix} 2 & 1 \\ 3 & 8 \end{bmatrix}}{-1} = \frac{13}{-1} = -13.$$

Commento. La formula deriva direttamente dalle proprietà del determinante. Infatti, ciascuna delle seguenti uguaglianze implica la successiva

$$x_1 a + x_2 b = p,$$
$$\det[x_1 a + x_2 b, b] = \det[p, b],$$
$$x_1 \det[a, b] = \det[p, b],$$
$$x_1 = \frac{\det[p, b]}{\det[a, b]}.$$

In modo analogo si ricava la formula per x_2.

Area e determinante

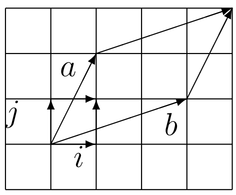
Fatto. Il determinante di una matrice 2×2 ha il seguente significato geometrico:

Fissato un punto O del piano, ad ogni ogni coppia vettori $a, b \in \mathcal{V}^2$ associamo il parallelogramma con un vertice in O e due lati i segmenti orientati uscenti da O che danno a, b (gli altri due lati sono allora determinati). Identificato $\mathcal{V}^2$ con $\mathbb{R}^2$ mediante un riferimento i, j, allora per ogni $a = a_1 i + a_2 j$ e ogni $b = b_i i + b_2 j$,

$$\frac{\text{area del parallelogramma su } a, b}{\text{area del parallelogramma su } i, j} = |\det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}|,$$

dove a destra compare il valore assoluto del determinante.

Esempio.



$$\frac{\text{area del parallelogramma } a, b}{\text{area del quadrato } i, j} = |\det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}| = |-5| = 5.$$

(Applicando ad a, b l'algoritmo di Gram-Schmidt si ottengono due vettori $a' = a$ e b' che danno un rettangolo che ha la stessa area del parallelogramma su a, b, dunque deve risultare $\|a\|\|b'\| = 5$.)

4.3 Determinante di matrici $n \times n$

Definizione per proprietà, formula.

Teorema. Esiste un'unica funzione "determinante"

$$\det : \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}, \quad [a_1, a_2, \cdots, a_n] \to \det[a_1, a_2, \cdots, a_n]$$

che è compatibile con le operazioni su una colonna fissate le altre, che sulle sequenze con due colonne uguali si annulla e che sulla base canonica vale 1:

1) $\det[a_1' + a_2'', a_2, \cdots] = \det[a_1', a_2, \cdots] + \det[a_2'', a_2, \cdots]$

analoghe sulle altre colonne

2) $\det[ra_1', a_2, \cdots] = r \det[a_1', a_2, \cdots]$

analoghe sulle altre colonne

3) $\det[a, a, a_3, \cdots] = 0$

analoghe sulle altre coppie di colonne

4) $\det[e_1, e_2, \cdots, e_n] = 1$

per ogni $a_j, a_j', a_j'' \in \mathbb{R}^n$ ed ogni $r \in \mathbb{R}$.

Osservazioni.

- Combinando la 1) e la 2) si ha che per ogni $a_1^{(1)}, \ldots, a_1^{(p)} \in \mathbb{R}^n$ ed ogni $r^{(1)}, \ldots, r^{(p)} \in \mathbb{R}$,

$$\det[\sum_{i=1}^m r^{(i)} a_1^{(i)}, a_2, \cdots] = \sum_{i=1}^m r^{(i)} \det[a_1^{(i)}, a_2, \cdots]$$

e analogamente sulle altre colonne.

- Dalle 1, 2, 3 segue che

$$\det[a_2, a_1, a_3, \cdots] = -\det[a_2, a_1, a_3, \cdots];$$

in generale: scambiando le colonne il determinante cambia segno.

Le proprietà 1) ... 4) permettano di calcolare immediatamente il determinante di certi tipi di matrici.

$$\det \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = 2 \cdot 3 \cdot 5 \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 2 \cdot 3 \cdot 5 = 30.$$

Una matrice diagonale è una matrice nella quale tutti gli elementi al di fuori della diagonale discendente sono nulli. Il determinante di una matrice diagonale $n \times n$ è il prodotto degli n elementi diagonali

$$\det \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix} = d_{11} d_{22} \cdots d_{nn}.$$

Di seguito mostriamo delle formule, dette "sviluppi di Laplace", che esprimono il determinate di una matrice $n \times n$ con $n > 1$ in funzione dei determinanti delle sottomatrici $(n-1) \times (n-1)$. Per completezza, definiamo il determinante di una matrice 1×1 come il suo unico elemento

$$\det[a] = a.$$

Sia A una matrice $n \times n$ con $n > 1$. Per ogni elemento di A, cancellando la riga e la colonna dell'elemento si ottiiene una matrice $(n-1) \times (n-1)$, il cui determinante si dice "cofattore" dell'elemento. Indichiamo con a_{ij} e con C_{ij} l'elemento di posto (i, j) e il rispettivo cofattore.

Lo sviluppo di Laplace del determinante di A rispetto ad una colonna è

$$\sum \pm(\text{elementi della colonna per rispettivi cofattori})$$

dove il segno è + o − secondo che la somma degli indici dell'elemento è pari o dispari. In simboli, lo sviluppo rispetto alla j−ma colonna è

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

Esempio Due sviluppi per colonna del determinante della matrice $\begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 9 \end{bmatrix}$

per 1° colonna:

$$\begin{aligned} &2 \det \begin{bmatrix} 5 & 7 \\ 6 & 9 \end{bmatrix} - 3 \det \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} + 4 \det \begin{bmatrix} 4 & 6 \\ 5 & 7 \end{bmatrix} \\ &= 2(45 - 42) - 3(36 - 36) + 4(28 - 30) \\ &= -2 \end{aligned}$$

per 2° colonna:

$$\begin{aligned} &- 4 \det \begin{bmatrix} 3 & 7 \\ 4 & 9 \end{bmatrix} + 5 \det \begin{bmatrix} 2 & 6 \\ 4 & 9 \end{bmatrix} - 6 \det \begin{bmatrix} 2 & 6 \\ 3 & 7 \end{bmatrix} \\ &= -4(27 - 28) + 5(18 - 24) - 6(14 - 18) \\ &= -2 \end{aligned}$$

Lo sviluppo di Laplace del determinante di A rispetto ad una riga è

$$\sum \pm(\text{elementi della riga per rispettivi cofattori})$$

dove il segno è + o − secondo che la somma degli indici dell'elemento è pari o dispari. In simboli, lo sviluppo rispetto alla i−ma riga è

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

Esempio Uno sviluppo per riga del determinante della matrice $\begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 9 \end{bmatrix}$

per 3° riga:

$$\begin{aligned} &4 \det \begin{bmatrix} 4 & 6 \\ 5 & 7 \end{bmatrix} - 6 \det \begin{bmatrix} 2 & 6 \\ 3 & 7 \end{bmatrix} + 9 \det \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix} \\ &= 4(28 - 30) - 6(14 - 18) + 9(10 - 12) \\ &= -2 \end{aligned}$$

Teorema. Tutti gli sviluppi di Laplace del determinante di una matrice A ($n \times n$, $n > 1$) hanno lo stesso valore, il determinante di A.

Trasposizione. Come conseguenza del Teorema, si ha che il determinante di una matrice è uguale a quello della sua trasposta:

$$\det(A) = \det(A^T),$$

quindi

il determinante è equivalentemente definito da proprietà per righe.

3 Singolarità di matrici e annullamento del determinante

Matrici triangolari

Una matrice quadrata si dice "triangolare superiore" se è del tipo

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}, \quad \ldots$$

in altri termini, indicati con a_{ij} $(i, j = 1, \ldots, n)$ gli elementi della generica matrice $n \times n$, una matrice è triangolare superiore se e solo se soddisfa le condizioni

$$\begin{aligned} 0 &= a_{2,1} \\ 0 &= a_{3,1} = a_{3,2} \\ \vdots &= \\ 0 &= a_{n,1} = a_{n,2} = \cdots = a_{n,n-1} = 0; \end{aligned}$$

in sintesi:

$$a_{ij} = 0 \text{ per ogni } n \geq i > j \geq 1.$$

Il determinante di una matrice triangolare superiore è semplice:

$$\det \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = ac, \quad \det \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} = adf, \quad \ldots$$

in altri termini, il determinate di una matrice A triangolare superiore $n \times n$ è il prodotto degli elementi diagonali di A :

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

Analogamente, si definiscono le matrici triangolari inferiori; il determinante di una matrice triangolare inferiore è il prodotto degli elementi diagonali.

Operazione elementare

L'operazione elementare "sommare a una colonna un multiplo di un'altra colonna" lascia invariato il determinante. Infatti:

$$\begin{aligned} \det[\cdots, a_i, \cdots, a_j + r a_i, \cdots] &= \det[\cdots, a_i, \cdots, a_j, \cdots] + r \det[\cdots, a_i, \cdots, a_i, \cdots] \\ &= \det[\cdots, a_i, \cdots, a_j, \cdots]. \end{aligned}$$

Questa proprietà permette di ricondurre il calcolo del determinante di una matrice al determinante di una matrice triangolare. Ad esempio:

$$\det \begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 9 \end{bmatrix} = \det \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & -2 \\ 4 & -2 & -3 \end{bmatrix} = \det \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 0 \\ 4 & -2 & 1 \end{bmatrix} = 2(-1)1 = -2$$

(operazioni elementari: $2^\circ - 2 \cdot 1^\circ$, $3^\circ - 2 \cdot 2^\circ$)

Teorema principale

Teorema. Per ogni matrice A quadrata $n \times n$,

(1) se A è singolare allora $\det A = 0$, e

(2) se A è non singolare allora $\det A \neq 0$.

In altri termini,

A è singolare se e solo se $\det(A) = 0$.

Dimostrazione. Sia $A = [a_1, \cdots, a_n]$ $(a_j \in \mathbb{R}^n)$. Per ogni colonna a_j, indichiamo la sua 1°, 2°, ... componente con $(a_j)_1, (a_j)_2, \ldots$

(1) Se A è singolare, allora $a_1, \ldots, a_n$ sono linearmente dipendenti; allora $a_1 = \underline{0}$ oppure esiste una colonna a_i (con $i \geq 2$) che è combinazione lineare delle precedenti,

$$a_i = \sum_{1 \leq j < i} r_j a_j$$

allora nel primo caso $\det(A) = \det[\underline{0}, a_2, \cdots] = 0$ e nel secondo caso

$$\begin{aligned}
\det(A) &= \det[a_1, \cdots, a_{i-1}, \sum_{1 \leq j < i} r_j a_j, a_{i+1}, \cdots] \\
&= \sum_{1 \leq j < i} r_j \det[a_1, \cdots, a_{i-1}, a_j, a_{i+1}, \cdots] \\
&= \sum_{1 \leq j < i} r_j 0 \\
&= 0;
\end{aligned}$$

quindi in ogni caso $\det(A) = 0$.

(2) Indichiamo con a_{ij} $(i, j = 1, \ldots, n)$ gli elementi della generica matrice $n \times n$. Sia data una matrice A' non singolare.

1° passo; le colonne di A' sono linearmente indipendenti, quindi la 1° è $\neq \underline{0}$; supponiamo per semplicità che A' soddisfi la condizione

$$a_{11} \neq 0;$$

sommando alle colonne dalla 2° in poi opportuni multipli della 1°, trasformiamo A' in una A'' che soddisfa (la condizione precedente e) la condizione

$$a_{12} = a_{13} = \cdots = a_{1n} = 0;$$

2° passo; le colonne di A'' sono linearmente indipendenti, quindi la 2° è $\neq \underline{0}$; supponiamo per semplicità che A'' soddisfi la condizione

$$a_{22} \neq 0;$$

sommando alle colonne dalla 3° in poi opportuni multipli della 2°, trasformiamo A'' in una A''' che soddisfa (le condizioni precedenti e) la condizione

$$a_{23} = a_{24} = \cdots = a_{2n} = 0;$$

3° passo; le colonne di A''' sono linearmente indipendenti, quindi la 3° è $\neq \underline{0}$; supponiamo per semplicità che A''' soddisfi la condizione

$$a_{33} \neq 0;$$

...

procedendo in questo modo si giunge ad una matrice $A^{(n)}$ triangolare inferiore con elementi diagonali tutti $\neq 0$.

Durante il processo il determinante della matrice rimane invariato, quindi

$$\det(A') = \det(A^{(n)}) \neq 0.$$

Esempio. $\det \begin{bmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 9 \end{bmatrix} = -2 \neq 0$, quindi la matrice è non singolare.

Esempio. Sia p un parametro $\in \mathbb{R}$.

$$\det \begin{bmatrix} p & -1 & 1 \\ 1 & p & -1 \\ -1 & 1 & p \end{bmatrix} = p^3 + 3p$$

quindi la matrice è singolare se e solo se $p^3 + 3p = 0$, essendo $p^3 + 3p = p(p^2 + 3)$, se e solo se $p = 0$.

Sistemi lineari e determinanti

Teorema. E' dato il sistema lineare di n equazioni in n incognite $x_1, x_2, \ldots, x_n$

$$x_1 a_1 + \cdots + x_n a_n = p,$$

con $a_1, \ldots, a_n, p$ costanti in $\mathbb{R}^n$.

1) Se $\det[a_1, \cdots, a_n] = 0$, allora il sistema non ha soluzioni oppure ne ha infinite;

2) Se $\det[a_1, \cdots, a_n] \neq 0$, allora il sistema ha una ed una sola soluzione; una formula per la soluzione è

$$x_i = \frac{\det[a_1, \cdots, \overset{i}{p}, \cdots, a_n]}{\det[a_1, \cdots, a_i, \cdots, a_n]} \qquad \text{per ogni } i = 1, \ldots, n.$$

dove il numeratore si ottiene dal denominatore sostituendo alla $i-$ma colonna a_i la colonna p.

Esempio. Consideriamo il sistema lineare di 3 equazioni in 3 incognite x_1, x_2, x_3 associato alla matrice

$$\left[\begin{array}{ccc|c} p & -1 & 1 & 1 \\ 1 & p & -1 & 1 \\ -1 & 1 & p & 1 \end{array}\right]$$

$$\det \begin{bmatrix} p & -1 & 1 \\ 1 & p & -1 \\ -1 & 1 & p \end{bmatrix} = p^3 + 3p = p(p^2 + 3) = 0 \text{ se e solo se } p = 0.$$

Caso $p \neq 0$. Il sistema ha un'unica soluzione;

$$x_1 = \frac{\det \begin{bmatrix} 1 & -1 & 1 \\ 1 & p & -1 \\ 1 & 1 & p \end{bmatrix}}{p^2 + 3p} = \frac{p^{+}3}{p^3 + 3p} = \frac{1}{p}$$

x_2, x_3 lasciate al lettore;

Caso $p = 0$ lasciato al lettore.

Commento. La formula deriva direttamente dalle proprietà del determinante. Proviamo quella per x_1. Ciascuna delle seguenti uguaglianze implica la successiva:

$$\sum_{j=1}^{n} x_j a_j = p,$$

$$\det[\sum_{j=1}^{n} x_j a_j, a_2, \cdots, a_n] = \det[p, a_2, \cdots, a_n],$$

$$\sum_{j=1}^{n} x_j \det[a_j, a_2, \cdots, a_n] = \det[p, a_2, \cdots, a_n],$$

$$x_1 \det[a_1, a_2, \cdots, a_n] = \det[p, a_2, \cdots, a_n]$$

$$x_1 = \frac{\det[p, a_2, \cdots, a_n]}{\det[a_1, a_2, \cdots, a_n]}.$$

Volume e determinante

Informalmente, un parallelepipedo è una configurazione che si ottiene da un cubo modificando le lunghezze dei lati e gli angoli fra i lati e conservando parallelismo e congruenza dei lati opposti.

Fatto. Il determinante di una matrice 3×3 ha il seguente significato geometrico:

Fissato un punto O dello spazio, ad ogni terna di vettori $a, b, c \in \mathcal{V}^3$ associamo il parallelepipedo con un vertice in O e tre lati i segmenti orientati uscenti da O che danno a, b, c (gli altri 9 lati sono allora determinati). Identificato $\mathcal{V}^3$ con $\mathbb{R}^3$ mediante un riferimento i, j, k per ogni $a = a_1 i + a_2 j + a_3 k$, $b = b_i i + b_2 j + b_3 k$, $c = c_i i + c_2 j + c_3 k$, si ha

$$\frac{\text{area del parallepipedo su } a, b, c}{\text{area del parallepipedo su } i, j, k} = |\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}|,$$

dove a destra compare il valore assoluto del determinante.

Prodotto di matrici, singolarità, determinante

Proposizione. Due matrici A, B quadrate $n \times n$ sono non singolari se e solo se il prodotto AB è non singolare.

Dimostrazione parziale. Supponiamo che A e B siano non singolari. Ciascuna delle seguenti uguaglianze implica la successiva

$$(AB)x = \underline{0} \qquad (x \text{ e } \underline{0} \text{ colonne } n \times 1)$$

$$A(Bx) = \underline{0};$$

$$Bx = \underline{0};$$

$$x = \underline{0};$$

ciò prova che le colonne di AB sono linearmente indipendenti, quindi AB è non singolare. (nel 2° e 3° passaggio abbiamo usato le ipotesi che le colonne di A e le colonne di B sono linearmente indipendenti).

Non proviamo che se AB è non singolare allora A e B sono non singolari.

La proposizione si può esprimere nella forma

$$\forall A, B \text{ quadrate moltiplicabili, } \det(AB) \neq 0 \text{ se e solo se } \det(A) \neq 0 \text{ e } \det(B) \neq 0.$$

Vale un risultato più forte:

Teorema Per ogni A, B quadrate moltiplicabili,

$$\det(AB) = \det(A)\det(B).$$

Non diamo dimostrazione.

Commento. Una conseguenza:

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

Infatti, ciascuna delle seguenti uguaglianze implica la successiva

$$AA^{-1} = \mathrm{I}_n, \;\; \det(AA^{-1}) = \det(\mathrm{I}_n), \;\; \det(A)\det(A^{-1}) = 1, \;\; \det(A^{-1}) = \frac{1}{\det A}.$$

5 Applicazioni lineari

5.1 Definizione di applicazione lineare, applicazioni $\mathbb{R}^n \to \mathbb{R}^m$

Introduzione

Applicazioni

La parola “applicazione” è un sinonimo della parola “funzione”. Una “applicazione” è data da: un insieme “dominio”, un insieme “codominio”, per ciscun elemento del dominio un'unico elemento del codominio; si dice che l'applicazione “va” dal dominio al codominio e che assume su ciascun elemento del dominio un “valore” nel codominio. Indicate con F, D, C, d applicazione, dominio, codominio e tipico elemento del dominio, si scrive

$$\mathrm{F} : D \to C, \quad d \mapsto \mathrm{F}(d).$$

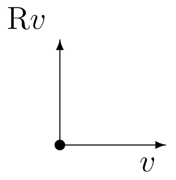
Due applicazioni si dicono uguali se e solo se il dominio e il codominio dell'una sono uguali al dominio e codominio dell'altra e su ciascun elemento del dominio l'una e l'altra assumono lo stesso valore; in simboli: $\mathrm{F} : D \to C$, e $\mathrm{G} : B \to A$ sono uguali se e solo se $D = B,\ C = A$, e

$$\mathrm{F}(d) = \mathrm{G}(d) \quad \text{per ogni } d \in D.$$

Commento. Spesso un'applicazione sarà data da un'espressione in una variabile nel dominio, espressione che per ogni valore della variabile assume un'unico valore nel codominio; il simbolo usato per indicare la variabile è irrilevante: le applicazioni $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ e $\mathrm{G} : \mathbb{R} \to \mathbb{R}$ date da $\mathrm{F}(x) = 2\,x$ e $\mathrm{G}(y) = 2\,y$ sono uguali, $\mathrm{F} = \mathrm{G}$.

Una rotazione

Fissiamo nel piano un punto O e rappresentiamo ogni vettore del piano con un segmento orientato uscente dal punto O. Per ogni vettore, consideriamo il corrispondente segmento orientato uscente da O, applichiamo una rotazione di $\frac{\pi}{2}$ in senso antiorario attorno ad O, otteniamo un nuovo segmento orientato uscente da O, e consideriamo il corrispondente vettore (diamo per nota la nozione di rotazione di un dato angolo attorno a dato un punto in un dato senso; la diamo per nota da un punto di vista operativo, non di definizione formale, che è piuttosto delicata). Abbiamo così un'applicazione di $\mathcal{V}^2$ in sè che diciamo “rotazione di $\frac{\pi}{2}$ in senso antiorario” ed indichiamo con R.



Fatto. La rotazione è compatibile con le operazioni sui vettori: per ogni $v_1, v_2 \in \mathcal{V}^2$ e ogni $\alpha \in \mathbb{R}$,

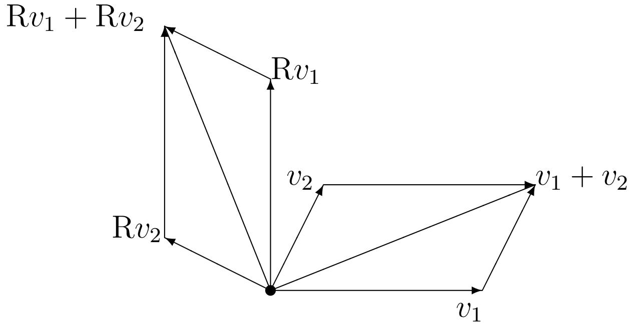
$$\mathrm{R}(v_1 + v_2) = \mathrm{R}(v_1) + \mathrm{R}(v_2);$$

$\mathrm{R}(\alpha\, v_1) = \alpha\, \mathrm{R}(v_1);$

in altri termini: per ogni $v_1, v_2 \in \mathcal{V}^2$ e ogni $\alpha, \beta \in \mathbb{R}$,

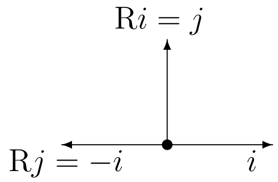
$\mathrm{R}(\alpha\, v_1 + \beta\, v_2) = \alpha\, \mathrm{R}(v_1) + \beta\, \mathrm{R}(v_2).$

Illustrazione della 1° proprietà:



Queste proprietà permettono di descrivere la rotazione in un dato riferimento. Per semplicità, fissiamo un riferimento ortonormale i, j, con j ottenuto da i mediante la rotazione R. Allora

$\mathrm{R}(i) = j, \quad \mathrm{R}(j) = -i$



Ogni vettore di $\mathcal{V}^2$ si scrive in un unico modo come combinazione lineare $x\, i + y\, j$ $(x, y \in \mathbb{R})$ e il suo rotato è

$\mathrm{R}(x\, i + y\, j) = x\, \mathrm{R}(i) + y\, \mathrm{R}(j) = x\, j - y\, i.$

Rappresentando $\mathcal{V}^2$ con $\mathbb{R}^2$ mediante il riferimento i, j, l'applicazione $\mathrm{R} : \mathcal{V}^2 \to \mathcal{V}^2$ viene rappresentata dall'applicazione $\overline{\mathrm{R}} : \mathbb{R}^2 \to \mathbb{R}^2$,

$\overline{\mathrm{R}}(x, y) = (-y, x).$

Definizione. Un'applicazione $\mathrm{F} : V \to W$ fra due spazi vettoriali V, W si dice "lineare" se e solo se è compatibile con le operazioni nei due spazi vettoriali:

$\mathrm{F}(v_1 + v_2) = \mathrm{F}(v_1) + \mathrm{F}(v_2), \quad \forall v_1, v_2 \in V;$

$\mathrm{F}(\alpha\, v_1) = \alpha\, \mathrm{F}(v_1), \quad \forall v_1 \in V, \alpha \in \mathbb{R};$

equivalentemente:

$\mathrm{F}(\alpha\, v_1 + \beta\, v_2) = \alpha\, \mathrm{F}(v_1) + \beta\, \mathrm{F}(v_2). \quad \forall v_1, v_2 \in V, \alpha, \beta \in \mathbb{R}.$

Tutte le rotazioni del piano danno applicazioni lineari $\mathcal{V}^2 \to \mathcal{V}^2$, che, fissato un riferimento, sono rappresentate da applicazioni lineari $\mathbb{R}^2 \to \mathbb{R}^2$. Svilupperemo questo punto in seguito.

Applicazioni lineari $\mathbb{R}^n \to \mathbb{R}^m$, qualche caso

Applicazioni lineari $\mathbb{R} \to \mathbb{R}$.

1- Esempio. $\mathrm{F} : \mathbb{R} \to \mathbb{R}$, $\mathrm{F}(x) = 2\,x$ è lineare. Infatti:

$$\mathrm{F}(x_1 + x_2) = \mathrm{F}(x_1) + \mathrm{F}(x_2), \quad \forall x_1, x_2 \in \mathbb{R}$$

significa

$$2\,(x_1 + x_2) = 2\,x_1 + 2\,x_2, \quad \forall x_1, x_2 \in \mathbb{R}$$

vera (per la proprietà distributiva del prodotto rispetto alla somma in $\mathbb{R}$);

$$\mathrm{F}(\alpha\,x_1) = \alpha\,\mathrm{F}(x_1) \quad \forall x_1, \alpha \in \mathbb{R}$$

significa

$$2\,(\alpha\,x_1) = \alpha\,(2\,x_1), \quad \forall x_1, \alpha \in \mathbb{R}$$

vera (per le proprietà associativa e commutativa del prodotto in $\mathbb{R}$).

2- Fatto. Ogni $\mathrm{F} : \mathbb{R} \to \mathbb{R}$, $\mathrm{F}(x) = (\text{costante})\,x$ è lineare. Si prova come sopra.

3- Esempio. $\mathrm{F} : \mathbb{R} \to \mathbb{R}$, $\mathrm{F}(x) = x + 1$ non è lineare. Infatti:

$$\mathrm{F}(x_1 + x_2) = \mathrm{F}(x_1) + \mathrm{F}(x_2), \quad \forall x_1, x_2 \in \mathbb{R}$$

significa

$$x_1 + x_2 + 1 = x_1 + 1 + x_2 + 1, \quad \forall x_1, x_2 \in \mathbb{R}$$

falsa.

4- Fatto. Ogni $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ lineare è del tipo $\mathrm{F}(x) = (\text{costante})\,x$. Infatti, se F è lineare, allora per ogni $x \in \mathbb{R}$

$$\mathrm{F}(x) = \mathrm{F}(x \cdot 1) = x\,\mathrm{F}(1) = (\text{costante})\,x.$$

Riassumendo, abbiamo visto che:

Le $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ *lineari sono tutte e sole le* $\mathrm{F}(x) = m\,x$ *(*m *costante).*

Applicazioni lineari $\mathbb{R}^2 \to \mathbb{R}^3$.

Se $\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3$ è lineare, allora

$$\begin{aligned}
\mathrm{F}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \mathrm{F}\left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \\
&= x_1\,\mathrm{F}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\,\mathrm{F}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= x_1 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + x_2 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \\
&= \begin{bmatrix} a_1x_1 + b_1x_2 \\ a_2x_1 + b_2x_2 \\ a_3x_1 + b_3x_2 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},
\end{aligned}$$

per ogni x_1, x_2 variabili $\in \mathbb{R}$, con a_i, b_j costanti $\in \mathbb{R}$ definite implicitamente nel 3° passaggio.

Quindi, ogni applicazione lineare $\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3$ può essere descritta:

- usando solo le nozioni di coppia e terna ordinata, come

$$\mathrm{F}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_1x_1 + b_1x_2 \\ a_2x_1 + b_2x_2 \\ a_3x_1 + b_3x_2 \end{bmatrix}$$

per ogni $x_1, x_2 \in \mathbb{R}$, con a_i, b_j costanti $\in \mathbb{R}$;

- usando la struttura vettoriale:

$$\mathrm{F}(x_1e_1 + x_2e_2) = x_1\, f_1 + x_2\, f_2,$$

per ogni $x_1, x_2 \in \mathbb{R}$, con f_1, f_2 costanti $\in \mathbb{R}^3$;

- usando l'algebra delle matrici:

$$\mathrm{F}(x) = Ax$$

per ogni x colonna 2×1, con A matrice costante 3×2.

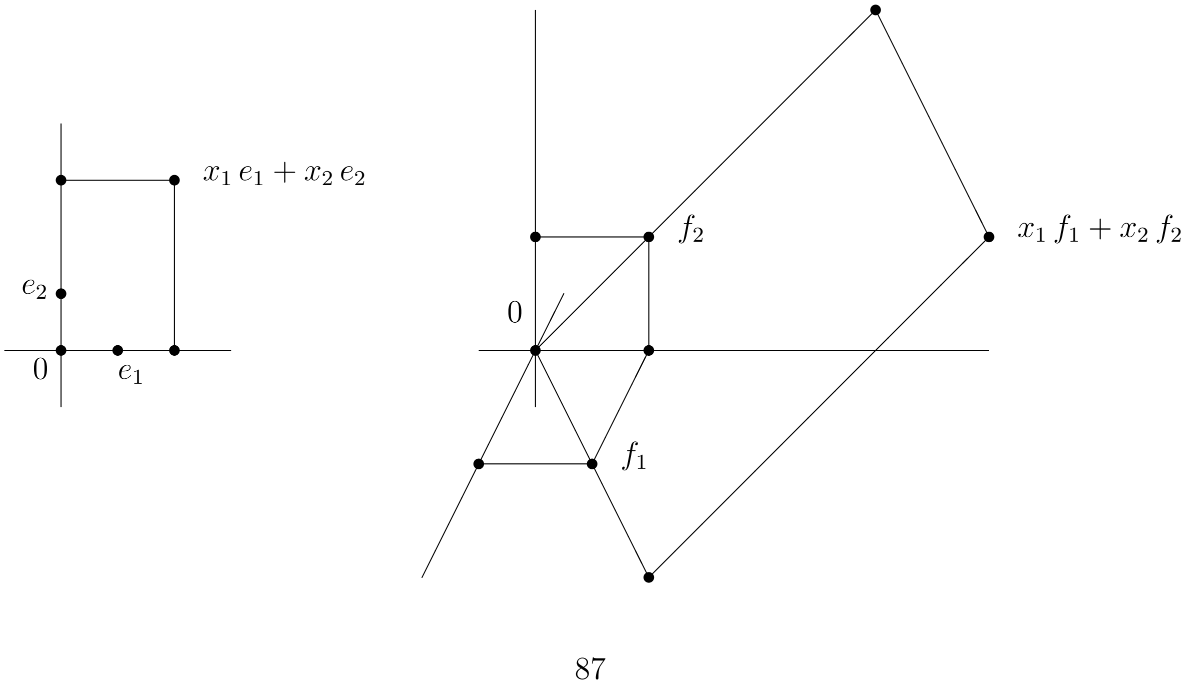
Viceversa, si prova che ogni applicazione descritta in uno di queste tre modi è lineare e risulta

$$f_1 = \mathrm{F}(e_1),\ \ f_2 = \mathrm{F}(e_2)$$
$$A = [f_1, f_2] = [\mathrm{F}(e_1), \mathrm{F}(e_2)].$$

Esempio. L'applicazione

$$\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3, \qquad \mathrm{F}\left(x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

è lineare; F trasforma il piano vettoriale $\mathbb{R}^2$ in un piano vettoriale immerso nello spazio vettoriale $\mathbb{R}^3$



Per semplicità, nella figura si sono identificati i vettori con i segmenti orientati uscenti da un punto fissato e i segmenti orientati col loro punto finale.

In generale, un'applicazione

$$\mathrm{F}(x_1e_1 + x_2e_2) = x_1 f_1 + x_2 f_2,$$

trasforma il piano vettoriale $\mathbb{R}^2$ in un piano vettoriale, una retta vettoriale, lo spazio nullo contenuti nello spazio vettoriale $\mathbb{R}^3$ secondo che f_1, f_2 siano rispettivamente linearmente indipendenti, linearmente dipendenti ma non entrambi nulli, entrambi nulli.

Applicazioni lineari $\mathbb{R}^n \to \mathbb{R}^m$

Proposizione. Ogni applicazione lineare $\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^m$ può essere descritta in ciascuna delle seguenti forme:

$$\mathrm{F}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{bmatrix}$$

per ogni $x_1, \ldots, x_n \in \mathbb{R}$, con a_{ij} costanti $\in \mathbb{R}$;

$$\mathrm{F}(\sum_{j=1}^n x_je_j) = \sum_{j=1}^n x_jf_j,$$

per ogni $x_1, \ldots, x_n \in \mathbb{R}$, con $f_1, \ldots, f_n$ costanti $\in \mathbb{R}^m$;

$$\mathrm{F}(x) = Ax$$

per ogni x colonna $n \times 1$, con A matrice costante $m \times n$.

Viceversa, ogni applicazione descritta in uno di queste tre modi è lineare e risulta

$$f_j = \mathrm{F}(e_j), \text{ per ogni } j = 1, \ldots, n;$$
$$A = [f_1 \cdots f_n] = [\mathrm{F}(e_1) \cdots \mathrm{F}(e_n)].$$

2 *Idea della dimostrazione.*

1° parte: se F è lineare, allora si scrive in ciascuno dei tre modi. La dimostrazione nel caso generale si sviluppa secondo la stessa linea già vista nel caso $\mathbb{R}^2 \to \mathbb{R}^3$.

2° parte: se F si scrive in uno dei tre modi, allora è lineare. Le tre scritture sono equivalenti. Supponiamo che

$$\mathrm{F}(x) = Ax$$

per ogni x colonna $n \times 1$, con A matrice costante $m \times n$. Allora

- per ogni $v_1, v_2 \in \mathbb{R}^n$ l'uguaglianza $\mathrm{F}(v_1+v_2) = \mathrm{F}(v_1)+\mathrm{F}(v_2)$ equivale alla $A(v_1+v_2) = Av_1 + Av_2$, che è vera per la proprietà delle operazioni fra matrici e colonne (cfr. p.52);

- per ogni $\alpha \in \mathbb{R}$ e $v_1 \in \mathbb{R}^n$ l'uguaglianza $\mathrm{F}(\alpha\, v_1) = \alpha\, \mathrm{F}(v_1)$ equivale alla $A(\alpha\, v_1) = \alpha\,(Av_1)$, che è vera per la proprietà delle operazioni fra matrici e colonne (cfr. p.52).

Intermezzo

Abbiamo visto principalmente i seguenti esempi di spazi vettoriali

$\mathcal{V}^1 \quad \mathcal{V}^2 \quad \mathcal{V}^3$

$\mathbb{R}^1 \quad \mathbb{R}^2 \quad \mathbb{R}^3 \quad \mathbb{R}^4 \quad \dots \quad \mathbb{R}^n \quad \dots$

una volta fissato un riferimento, precisamente una base, ciascun spazio vettoriale geometrico $\mathcal{V}^n$ si può identificare, come spazio vettoriale col rispettivo spazio vettoriale numerico $\mathbb{R}^n$ $(n = 1, 2, 3)$. Per ogni intero positivo n si può considerare come analogo $n-$dimensionale degli spazi vettoriali geometrici uno spazio vettoriale $n-$dimensionale astratto; una volta fissata una base, tale spazio vettoriale si può identificare con il corrispondente spazio vettoriale $\mathbb{R}^n$. Di regola, le definizioni, le proposizioni, i teoremi e le dimostrazioni sono più naturali per gli spazi vettoriali astratti.

Di seguito vedremo in particolare come le descrizioni della applicazioni lineari fra pazi vettoriali $\mathbb{R}^n$ si possano estendere a descrizioni di applicazioni lineari fra spazi vettoriali astratti.

D'ora innanzi, tranne avviso contrario, ogni spazio vettoriale considerato sarà tacitamente supposto di dimensione finita.

Applicazioni lineari

Siano V, W spazi vettoriali. Cosa possiamo dire delle applicazioni lineari da V a W?

Innanzitutto, l'unico elemento che sicuramente esiste in uno spazio vettoriale è il vettore nullo. Dunque possiamo definire un'applicazione $\mathrm{F} : V \to W$ ponendo $\mathrm{F}(v) = \underline{0} \in W$, per ogni $v \in V$. Questa applicazione è lineare, viene detta "applicazione nulla".

Teorema. Siano dati: uno spazio vettoriale V, una sequenza di vettori $a_1, \dots, a_n$ base di V, uno spazio vettoriale W, una sequenza di vettori $w_1, \dots, w_n$ di W. Allora esiste un'unica applicazione lineare $\mathrm{F} : V \to W$ tale che

$$\mathrm{F}(a_1) = w_1, \dots, \mathrm{F}(a_n) = w_n;$$

esplicitamente, per ogni $v = x_1 a_1 + \cdots + x_n a_n \in V$,

$$\mathrm{F}(v) = x_1 w_1 + \cdots + x_n w_n.$$

Commenti:

- Il senso del teorema è che, dato uno spazio vettoriale $n-$dimensionale V, una volta fissata una base in V, si possono identificare le applicazioni lineari da V verso uno spazio vettoriale W con le sequenze di n vettori di W : ogni applicazione lineare $V \to W$ si può rappresentare con un sequenza di n vettori di W, le immagini dei vettori base di V, e ogni sequenza di n vettori di W rappresenta una ed una sola applicazione lineare $V \to W$.

- L'ipotesi che $a_1, \dots, a_n$ sia una base di V è fondamentale, in quanto assicura per ciascun vettore $v \in V$ l'esistenza e l'unicità della sequenza dei coefficienti $x_1, \dots, x_n$ che servono per costruire $\mathrm{F}(v)$.

- L'unicità dell'applicazione F, con la sua descrizione esplicita, è ovvia. La parte principale della dimostrazione consiste nel provare che una tale applicazione F è lineare.

Esempio. In $\mathcal{V}^2$, identificato con $\mathbb{R}^2$ mediante un riferimento, consideriamo

una base, ad esempio $a_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $a_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

due vettori, ad esempio $w_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $w_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

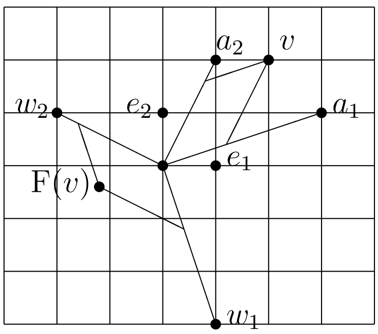
Esiste un'unica applicazione lineare $\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^2$ tale che $\mathrm{F}(a_1) = w_1$, $\mathrm{F}(a_2) = w_2$.

Calcoliamo il valore di F su un vettore, ad esempio $v = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = x_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ se esolo se } x_1 = \tfrac{2}{5},\ x_2 = \tfrac{4}{5};$$

quindi

$$\mathrm{F} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \tfrac{2}{5} \begin{bmatrix} 1 \\ -3 \end{bmatrix} + \tfrac{4}{5} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{6}{5} \\ -\frac{2}{5} \end{bmatrix}.$$



Applicazioni lineari iniettive, suriettive, biiettive

Applicazioni fra insiemi, iniettive, suriettive, biiettive.

Ricordiamo che un'applicazione $\mathrm{F} : D \to C$ fra insiemi si dice

- iniettiva se soddisfa una delle tre condizioni equivalenti

  per ogni $d_1, d_2 \in D$, da $d_1 \neq d_2$ segue $\mathrm{F}(d_1) \neq \mathrm{F}(d_2)$;

  per ogni $d_1, d_2 \in D$, da $\mathrm{F}(d_1) = \mathrm{F}(d_2)$ segue $d_1 = d_2$;

  per ogni $c \in C$, esiste al più un $d \in D$ tale che $\mathrm{F}(d) = c$;

- suriettiva se

  per ogni $c \in C$, esiste almeno un $d \in D$ tale che $\mathrm{F}(d) = c$;

- biiettiva se

  per ogni $c \in C$, esiste un'unico $d \in D$ tale che $\mathrm{F}(d) = c$.

In altri termini, l'applicazione è biiettiva se e solo se è sia iniettiva che suriettiva.

Cosa si può dire delle applicazioni lineari iniettive, suriettive, biiettive? Iniziamo con le applicazioni fra spazi vettoriali numerici.

Sia data un'applicazione lineare fra spazi vettoriali numerici

$$\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^m,\ \mathrm{F}(x) = Ax \qquad (A \text{ costante } m \times n).$$

Osserviamo che per ogni $b \in \mathbb{R}^m$, l'equazione nell'incognita x su $\mathbb{R}^n$

$$\mathrm{F}(x) = b$$

equivale all'equazione

$$Ax = b,$$

che è un sistema lineare di m equazioni in n incognite.

Proposizione.

- F è inettiva se e solo se le colonne di A sono linearmente indipendenti, cioè $\mathrm{r}(A) = n$;
- F è suriettiva se e solo se le righe di A sono linearmente indipendenti; cioè $\mathrm{r}(A) = m$;
- F è biiettiva se e solo se $n = \mathrm{r}(A) = m$, cioè A è non singolare.

2 *Dimostrazione parziale.* Proviamo solo la 1° e la 3° affermazione. Indicate con $f_1, \ldots, f_n$ le colonne di A, l'equazione $Ax = b$ si scrive anche come

$$x_1 f_1 + \cdots + x_n f_n = b.$$

- Se F è inettiva, allora per ogni $b \in \mathbb{R}^m$ l'equazione $x_1 f_1 + \cdots + x_n f_n = b$ ha al più una soluzione, allora l'equazione $x_1 f_1 + \cdots + x_n f_n = \underline{0}$ ha solo la soluzione $x_1 = \ldots = x_n = 0$, allora $f_1, \ldots, f_n$ sono linearmente indipendenti. Viceversa, si prova che se $f_1, \ldots, f_n$ sono linearmente indipendenti, allora per ogni $b \in \mathbb{R}^m$ l'equazione $x_1 f_1 + \cdots + x_n f_n = b$ ha al più una soluzione, quindi F è iiettiva.

- F è biiettiva se e solo se per ogni $b \in \mathbb{R}^m$ l'equazione $x_1 f_1 + \cdots + x_n f_n = b$ ha aun'unica soluzione se e solo se $f_1, \ldots, f_n$ è una base di $\mathbb{R}^m$ se e solo se A è non singolare se e solo se $m = \mathrm{r}(A) = n$.

Esempi.

$$\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3, \quad \mathrm{F}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ è iniettiva ma non suriettiva.}$$

$$\mathrm{G} : \mathbb{R}^2 \to \mathbb{R}^3, \quad \mathrm{G}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ non è iniettiva né suriettiva.}$$

$$\mathrm{H} : \mathbb{R}^3 \to \mathbb{R}^2, \quad \mathrm{H}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ è suriettiva ma non iniettiva.}$$

$$\mathrm{L} : \mathbb{R}^2 \to \mathbb{R}^2, \quad \mathrm{L}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ è biiettiva.}$$

Queste considerazioni e proposizione si possono riformulare in vari modi più o meno forti in termini di spazi vettoriali astratti, ad esempio come segue. Sia data un'applicazione lineare da uno spazio vettoriale V con una base $a_1, \ldots, a_n$ verso uno spazio vettoriale W

$$\mathrm{F}(x_1 a_1 + \cdots + x_n a_n) = x_1 f_1 + \cdots + x_n f_n \qquad (f_i \text{ costanti} \in W).$$

Proposizione.

- F è iniettiva se e solo se $f_1, \ldots, f_n$ è linearmente indipendente;
- F è suriettiva se e solo se $\mathrm{Span}\{f_1, \ldots, f_n\} = W$;
- F è biiettiva se e solo se $f_1, \ldots, f_n$ è una base di W.

Commento. La dimostrazione di questa proposizione segue quasi direttamente dalle definizioni. Viene lasciata al lettore.

Spazi nullo e colonna, spazi nucleo e immagine

Ricordiamo che a ciascuna matrice A di tipo $m \times n$ abbiamo associato alcuni spazi vettoriali, in particolare:

- lo spazio nullo di A

$\mathcal{N}(A) =$ insieme delle soluzioni del sistema lineare omogeneo $Ax = \underline{0}$;

- lo spazio colonna di A

$\mathcal{C}(A) = \mathrm{Span}\{f_1, \ldots, f_n\}$ ($f_j \in \mathbb{R}^m$ colonne di A);

e abbiamo visto la relazione fra le loro dimensioni

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = n.$$

Indicata con $\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare data da

$$\mathrm{F}(x) = Ax$$

possiamo descrivere questi spazi come segue:

- $\mathcal{N}(A)$ è l'insieme dei vettori di $\mathbb{R}^n$ che vengono mandati nel vettore nullo di $\mathbb{R}^m$:

$$\mathcal{N}(A) = \{x : \ \mathrm{F}(x) = \underline{0}\} \subseteq \mathbb{R}^n;$$

- $\mathcal{C}(A)$ è l'insieme delle immagini in $\mathbb{R}^m$ dei vettori di $\mathbb{R}^n$:

$$\mathcal{C}(A) = \{\mathrm{F}(x); \ x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

Lo spazio nullo e lo spazio colonna della matrice si dicono rispettivamente anche "spazio nucleo" e "spazio immagine" dell'applicazione F. Più in generale, si ha la

Definizione. Sia $\mathrm{F} : V \to W$ un'applicazione lineare fra due spazi vettoriali V, W;

- il nucleo di F è l'insieme dei vettori che F manda nel vettore nullo di W :

$$\mathrm{Ker}(F) = \{x : \ \mathrm{F}(x) = \underline{0}\} \subseteq V;$$

- l'immagine di F è l'insieme dei vettori immagine che F crea in W :

$$\mathrm{Im}(F) = \{\mathrm{F}(x); \ x \in V\} \subseteq W.$$

Si verifica che Ker(F) e Im(F) sono sottospazi, rispettivamente di V e W.

Esempio.

$$\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3, \quad \mathrm{F}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

$$\begin{aligned}
\mathrm{Ker}(F) &= \mathcal{N}\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= (\text{ soluzioni di } \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}) \\
&= (\text{ insieme dei } \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ con } x_2 \text{ libera }) \\
&= \mathrm{Span}\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}.\}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Im}(F) &= \mathcal{C}\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= \mathrm{Span}\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}\} \\
&= \mathrm{Span}\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.\}
\end{aligned}$$

Esempio. Siano date: una base a_1, a_2 di uno spazio vettoriale V, due vettori w_1, w_2 linearmente indipendenti in uno spazio vettoriale W, l'applicazione lineare $\mathrm{F} : V \to W$ tale che

$$\mathrm{F}(a_1) = w_1, \ \mathrm{F}(a_2) = w_2$$

esplicitamente,

$$\mathrm{F}(xa_1 + ya_2) = xw_1 + yw_2 \qquad \text{per ogni } x, y \in \mathbb{R}.$$

Allora

$$\mathrm{Ker(F)} = \{xa_1 + ya_2 : xw_1 + yw_2 = \underline{0}\} = \{\underline{0}\};$$
$$\mathrm{Im(F)} = \{xw_1 + yw_2;\ x, y \in \mathbb{R}\} = \mathrm{Span}\{w_1, w_2\}.$$

Al teorema sulla relazione fra le dimensioni dello spazio nullo e spazio colonna di una matrice corrisponde il seguente teorema sulla relazione fra le dimensioni dello spazio nucleo e spazio immagine di un'applicazione lineare

Teorema. Siano: $\mathrm{F} : V \to W$ un'applicazione lineare, $z_1, \ldots, z_p$ una bse di Ker(F), $\mathrm{F}(v_1), \ldots, \mathrm{F}(v_q)$ una base di Im(F). Allora $z_1, \ldots, z_p, v_1, \ldots, v_q$ è una base di V, quindi

$$\dim(\mathrm{KerF}) + \dim(\mathrm{ImF}) = \dim(V).$$

Dimostrazione.

- Consideriamo l'uguaglianza

$$x_1z_1 + \cdots + x_pz_p + y_1v_1 + \cdots + y_qv_q = \underline{0};$$

applicando F ad entranmbi i membri, essendo F lineare, si ha

$$x_1\mathrm{F}(z_1) + \cdots + x_p\mathrm{F}(z_p) + y_1\mathrm{F}(v_1) + \cdots + y_q\mathrm{F}(v_q) = \underline{0};$$

essendo $z_1, \ldots, z_n \in \mathrm{Ker(F)}$, si ha

$$y_1\mathrm{F}(v_1) + \cdots + y_q\mathrm{F}(v_q) = \underline{0};$$

essendo $\mathrm{F}(v_1), \ldots, \mathrm{F}(v_q)$ linearemente indipendenti, si ha $y_1 = \cdots = y_q = 0$; quindi

$$x_1z_1 + \cdots + x_pz_p = \underline{0};$$

essendo $z_1, \ldots, z_p$ linearemente indipendenti, si ha $x_1 = \cdots = x_p = 0$.

Quindi $z_1, \ldots, z_p, v_1, \ldots, v_q$ è linearmente indipendente.

- Sia $v \in V$. Poichè $\mathrm{F}(v_1), \ldots, \mathrm{F}(v_q)$ genera $\mathrm{Im(F)}$, esistono dei numeri $y_1, \ldots, y_q$ tali che

$$\mathrm{F}(v) = y_1\mathrm{F}(v_1) + \cdots + y_q\mathrm{F}(v_q);$$

essendo F lineare, si ha

$$\mathrm{F}(v - y_1v_1 - \cdots - y_qv_q) = \underline{0};$$

essendo $z_1, \ldots, z_p$ una base di $\mathrm{Ker(F)}$, esistono dei numeri $x_1, \ldots, x_p$ tali che

$$v - y_1v_1 - \cdots - y_qv_q = x_1z_1 + \cdots + x_pz_p;$$

quindi

$$v = x_1z_1 + \cdots + x_pz_p + y_1v_1 + \cdots + y_qv_q.$$

Quindi $z_1, \ldots, z_p, v_1, \ldots, v_q$ genera V.

5.2 Matrici di un'applicazione lineare

Coordinate, notazione

Ogni vettore di uno spazio vettoriale si scrive in un unico modo come combinazione lineare dei vettori di una base; ricordiamo che i coefficienti nella scrittura del vettore sono detti "le coordinate" del vettore rispetto ai vettori della base; diciamo che la sequenza delle coordinate di un vettore v rispetto a una sequenza base $\mathcal{A} : a_1, \dots, a_n$ è "la coordinata" (al singolare) di v rispetto ad $\mathcal{A}$, la indichiamo con $[v]_\mathcal{A}$ e la scriviamo come colonna:

$$\text{se } v = r_1 a_1 + \cdots + r_n a_n \ (r_i \in \mathbb{R}), \text{ allora } [v]_\mathcal{A} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}.$$

Esempio. $\mathcal{A} : a_1, a_2$; essendo $a_1 = 1\, a_1 + 0\, a_2$, si ha $[a_1]_\mathcal{A} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Osservazione. Ogni vettore di $\mathbb{R}^n$ si scrive come combinazione lineare, con coefficienti le sue componenti, dei vettori unità di $\mathbb{R}^n$; quindi la coordinata del vettore rispetto alla base canonica $\mathcal{E} : e_1, \dots, e_n$ di $\mathbb{R}^n$ è il vettore stesso:

$$[x]_\mathcal{E} = x.$$

Esempi.

$$v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}; \ \mathcal{A} : \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}; \quad \begin{bmatrix} 2 \\ 3 \end{bmatrix} = -4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix}; \quad [v]_\mathcal{A} = \begin{bmatrix} -4 \\ 3 \end{bmatrix};$$

$$v = \begin{bmatrix} 2 \\ 3 \end{bmatrix}; \ \mathcal{E} : \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \quad \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \quad [v]_\mathcal{E} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Matrici di un'applicazione lineare. Caso introduttivo

Siano
V uno spazio vettoriale con una base $\mathcal{A} : a_1, a_2$;
W uno spazio vettoriale con una base $\mathcal{B} : b_1, b_2, b_3$;
$\mathrm{F} : V \to W$ l'applicazione lineare che sui vettori della base $\mathcal{A}$ assume i valori

$$\mathrm{F}(a_1) = \ell\, b_1 + m\, b_2 + n\, b_3$$

$$\mathrm{F}(a_2) = p\, b_1 + q\, b_2 + r\, b_3$$

$(\ell, \dots, r$ costanti $\in \mathbb{R})$. Sul vettore tipico di V

$$v = x_1 a_1 + x_2 a_2$$

F assume il valore

$$\begin{aligned} \mathrm{F}(v) &= x_1 \mathrm{F}(a_1) + x_2 \mathrm{F}(a_2) \\ &= x_1(\ell\, b_1 + m\, b_2 + n\, b_3) + x_2(p\, b_1 + q\, b_2 + r\, b_3) \\ &= (\ell\, x_1 + p\, x_2) b_1 + (m\, x_1 + q\, x_2) b_2 + (n\, x_1 + r\, x_2) b_3. \end{aligned}$$

Osserviamo che

$$[\mathrm{F}(a_1)]_{\mathcal{B}} = \begin{bmatrix} \ell \\ m \\ n \end{bmatrix}, \ [\mathrm{F}(a_2)]_{\mathcal{B}} = \begin{bmatrix} p \\ q \\ r \end{bmatrix};$$

$$[v]_{\mathcal{A}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

$$[\mathrm{F}(v)]_{\mathcal{B}} = \begin{bmatrix} \ell\, x_1 + p\, x_2 \\ m\, x_1 + q\, x_2 \\ n\, x_1 + r\, x_2 \end{bmatrix} = \begin{bmatrix} \ell & p \\ m & q \\ n & r \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [(\mathrm{F}a_1)_{\mathcal{B}}\ (\mathrm{F}a_2)_{\mathcal{B}}]\, [v]_{\mathcal{A}}\,.$$

(per semplicità, abbiamo scritto $(\mathrm{F}a_i)_{\mathcal{B}}$ al posto di $[\mathrm{F}(a_i)]_{\mathcal{B}}$).

Diciamo che $[(\mathrm{F}a_1)_{\mathcal{B}}\ (\mathrm{F}a_2)_{\mathcal{B}}]$ è la matrice di F rispetto a $\mathcal{B}$ e $\mathcal{A}$ (in questo ordine) e la indichiamo con $[\mathrm{F}]_{\mathcal{B}\mathcal{A}}$:

$$[\mathrm{F}]_{\mathcal{B}\mathcal{A}} = [(\mathrm{F}a_1)_{\mathcal{B}}\ (\mathrm{F}a_2)_{\mathcal{B}}]\,.$$

Allora possiamo dire che la coordinata di $\mathrm{F}(v)$ rispetto a $\mathcal{B}$ è il prodotto della matrice di F rispetto a $\mathcal{B}$ e $\mathcal{A}$ per la coordinata di v rispetto ad $\mathcal{A}$:

$$[\mathrm{F}v]_{\mathcal{B}} = [\mathrm{F}]_{\mathcal{B}\mathcal{A}}\, [v]_{\mathcal{A}}\,.$$

Esempio. Sia $\mathrm{F} : \mathcal{V}^2 \to \mathcal{V}^3$ un'applicazione lineare iniettiva. Se $\mathcal{A} : a_1, a_2$ è una base di $\mathcal{V}^2$, allora $\mathrm{F}(a_1), \mathrm{F}(a_2)$ è linearmente indipendente ed esiste una base $\mathcal{B} : b_1, b_2, b_3$ di $\mathcal{V}^3$ tale che $b_1 = \mathrm{F}(a_1), b_2 = \mathrm{F}(a_2)$. Così

$$\mathrm{F}(a_1) = 1\, b_1 + 0\, b_2 + 0\, b_3,$$

$$\mathrm{F}(a_2) = 0\, b_1 + 1\, b_2 + 0\, b_3,$$

e la matrice di F rispetto a queste basi è

$$[\mathrm{F}]_{\mathcal{B}\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Osservazione. La matrice di un'applicazione lineare $\mathrm{F} : \mathbb{R}^2 \to \mathbb{R}^3$ rispetto alle basi canoniche $\mathcal{E}', \mathcal{E}$ di $\mathbb{R}^3, \mathbb{R}^2$ è la matrice che ha per colonne le immagini dei vettori unità di $\mathbb{R}^2$:

$$[\mathrm{F}]_{\mathcal{E}',\mathcal{E}} = [\mathrm{F}(e_1)\ \mathrm{F}(e_2)].$$

Matrici di un'applicazione lineare

In generale, siano
V uno spazio vettoriale con una base $\mathcal{A} : a_1, \ldots, a_n$;
W uno spazio vettoriale con una base $\mathcal{B} : b_1, \ldots, b_m$;
$\mathrm{F} : V \to W$ un'applicazione lineare.

Diciamo che la matrice $m \times n$ che ha per colonne le coordinate rispetto a $\mathcal{B}$ delle immagini dei vettori di $\mathcal{A}$ è la matrice di F rispetto a $\mathcal{B}$ e $\mathcal{A}$ (in questo ordine) e la indichiamo con $[\mathrm{F}]_{\mathcal{B}\mathcal{A}}$:

$$[\mathrm{F}]_{\mathcal{B}\mathcal{A}} = [(\mathrm{F}a_1)_{\mathcal{B}}\ \cdots\ (\mathrm{F}a_n)_{\mathcal{B}}]$$

Così come nel caso $n = 2$ ed $m = 3$ si trova che per ogni $v \in V$ la coordinata di $\mathrm{F}(v)$ rispetto a $\mathcal{B}$ è il prodotto della matrice di F rispetto a $\mathcal{B}$ e $\mathcal{A}$ per la coordinata di v rispetto ad $\mathcal{A}$:

$$[\mathrm{F}(v)]_{\mathcal{B}} = [\mathrm{F}]_{\mathcal{B}\mathcal{A}}\,[v]_{\mathcal{A}}\,.$$

Osservazione. La matrice di un'applicazione lineare $\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^m$ rispetto alle basi canoniche $\mathcal{E}', \mathcal{E}$ di $\mathbb{R}^m, \mathbb{R}^n$ è la matrice che ha per colonne le immagini dei vettori unità di $\mathbb{R}^n$:

$$[\mathrm{F}]_{\mathcal{E}',\mathcal{E}} = [\mathrm{F}(e_1)\ \cdots \mathrm{F}(e_n)].$$

Le matrici più semplici

Una stessa applicazione lineare è si può dunque rappresentare rispetto tutte le possibili basi del codominio e del dominio. A certe basi corrisponderanno matrici più complicate e a certe altre matrici più semplici.

Esempio. Sia $\mathrm{F} : \mathbb{R}^3 \to \mathbb{R}^2$,

$$\mathrm{F}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 5\,x_3 \\ x_1 + 2\,x_2 + 8\,x_3 \end{bmatrix}.$$

La matrice di F rispetto alle basi $\mathcal{E}' : \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ e $\mathcal{E} : \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ è

$$[\mathrm{F}]_{\mathcal{E}'\mathcal{E}} = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 8 \end{bmatrix}$$

La matrice di F rispetto alle basi $\mathcal{B} : \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ e $\mathcal{A} : \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}$ è

$$[\mathrm{F}]_{\mathcal{B}\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

(vale a dire:

$$\mathrm{F}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 2 \end{bmatrix};$$

$$\mathrm{F}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 2 \end{bmatrix};$$

$$\mathrm{F}\begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 2 \end{bmatrix};$$

si lascia la verifica al lettore).

Si possono sempre scegliere le basi in modo che la matrice sia particolarmente semplice.

Proposizione Siano:

V, W spazi vettoriali di dimensioni n ed m;
$\mathrm{F} : V \to W$ un'applicazione lineare;
$v_1, \ldots, v_p \in V$ tali che $\mathrm{F}(v_1), \ldots, \mathrm{F}(v_p)$ sia una base di $\mathrm{Im}(\mathrm{F})$;
$z_1, \ldots, z_q \in V$ una base di Ker(F;

Allora:

$\mathcal{A} : v_1, \ldots, v_p, z_1, \ldots, z_q$ è una base di V;
esiste una base $\mathcal{B} : b_1, \ldots, b_p, \ldots$ di W tale che $\mathrm{F}(v_1) = b_1, \ldots, \mathrm{F}(v_p) = b_p$;
la matrice di F rispetto a queste basi è la matrice $m \times n$ con i primi elementi diagonali uguali a 1 e tutti gli altri elementi uguali a 0:

$$[\mathrm{F}]_{\mathcal{B}\mathcal{A}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 1 & & 0 & 0 & \cdots \\ \vdots & & \ddots & \vdots & \vdots & \\ 0 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \end{bmatrix}.$$

Commenti.

- Le due basi $\mathcal{B}, \mathcal{A}$ che rendono semplice la matrice dell'applicazione nell'esempio precedente sono state costruite secondo questa proposizione (se ne possono costruire altre).
- Che la sequenza $\mathcal{A}$ sia una base di V segue dal Teorema sulle dimensioni di Ker e Im a p.93.

- Che esista una base $\mathcal{B}$ di W del tipo indicato segue dal fatto che ogni sequenza linearmente indipendente di uno spazio vettoriale si può completare a una base.

Composizione e inversione di applicazioni lineari

Applicazioni fra insiemi.

Siano date due applicazioni F, G tali che il codominio di F sia uguale al dominio di G. Allora a ciascun elemento a del suo dominio, l'applicazione F associa un elemento $\mathrm{F}(a)$ e a $\mathrm{F}(a)$ l'applicazione G associa un elemento $\mathrm{G}(\mathrm{F}(a))$ del suo codominio. Si ha così un'applicazione dal dominio di F al codominio di G che si dice "G composta dopo F", in breve "G dopo F", e si indica con $\mathrm{G} \circ \mathrm{F}$. In simboli:

$$\begin{array}{lll} A \xrightarrow{\mathrm{F}} B & \xrightarrow{\mathrm{G}} C & \qquad A \xrightarrow{\mathrm{G}\circ\mathrm{F}} C \\ a \longmapsto \mathrm{F}(a) & \longmapsto \mathrm{G}(\mathrm{F}(a)) & \qquad a \longmapsto (\mathrm{G}\circ\mathrm{F})(a) \end{array}$$

Se il codominio di F è diverso dal dominio di G, allora $\mathrm{G}\circ\mathrm{F}$ non viene definita. Questa è la convenzione adottata in Algebra lineare; altre convenzioni, più lasche, sono adottate in altre parti della matematica, ad esempio in Analisi.

Si ha così l'operazione di "composizione" fra applicazioni. Le sue prime proprietà sono:

(1) La composizione non è commutativa, in quanto può succedere che di due composizioni $\mathrm{G}\circ\mathrm{F}$ e $\mathrm{F}\circ\mathrm{G}$ una sia definita e l'altra no; anche quando entrambe le composizioni sono definite, spesso sono diverse.

(2) La composizione è associativa: per ogni tre applicazioni $\mathrm{F}, \mathrm{G}, \mathrm{H}$, l'applicazione $(\mathrm{H}\circ\mathrm{G})\circ\mathrm{F}$ è definita se e solo se l'applicazione $\mathrm{H}\circ(\mathrm{G}\circ\mathrm{F})$ è definita, e in tal caso le due applicazioni sono uguali:

$$(\mathrm{H}\circ\mathrm{G})\circ\mathrm{F} = \mathrm{H}\circ(\mathrm{G}\circ\mathrm{F}).$$

(3) Per ogni insieme A, l'applicazione che ad ogni elemento di A associa sé stesso si dice "identità su A" e si indica con id_A; quindi

$$\underset{\circlearrowleft \mathrm{id}_A}{A} \quad \text{è data da } \mathrm{id}_A(a) = a.$$

Le applicazioni identità sono caratterizzate dalla seguente proprietà

$$\text{per ogni } A \xrightarrow{\mathrm{F}} B, \text{ si ha } \mathrm{F}\circ\mathrm{id}_A = \mathrm{F} = \mathrm{id}_B \circ \mathrm{F}.$$

Applicazione inversa. Sia data un'applicazione biiettiva

$$\mathrm{F} : A \to B.$$

Per ogni $b \in B$, l'equazione

$$\mathrm{F}(a) = b$$

ha una ed una sola soluzione, che indichiamo con

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

Si ha così un'applicazione

$$\mathrm{F}^{-1} : B \to A, \quad b \mapsto \mathrm{F}^{-1}(b),$$

che si dice "applicazione inversa" di F.

Proposizione. Per ogni $\mathrm{F} : A \to B$ le seguenti affermazioni sono equivalenti:

(1) F è biiettiva;

(2) esiste una $\mathrm{G} : B \to A$ tale che

$$\mathrm{G} \circ \mathrm{F} = \mathrm{id}_A, \quad \mathrm{F} \circ \mathrm{G} = \mathrm{id}_B.$$

In tal caso, G è unica, e

$$\mathrm{G} = \mathrm{F}^{-1}.$$

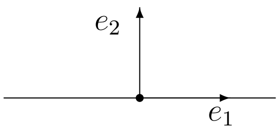
Una rotazione, una proiezione.

Fissato un punto O nel piano, identifichiamo i vettori del piano con segmenti orientati uscenti da O. Consideriamo le seguenti applicazioni di $\mathcal{V}^2$ in sé:

Rot_+, rotazione di $\frac{\pi}{2}$ in senso antiorario;

Pr, proiezione su una retta fissata passante per O.

Fissiamo un riferimento di due versori ortogonali, il primo sulla retta della proiezione, il secondo rotato del primo di $\frac{\pi}{2}$ in senso antiorario; tramite questo riferimento identifichiamo $\mathcal{V}^2$ con $\mathbb{R}^2$. Quindi i due vettori del riferimento vengono identificati coi rispettivi vettori unità e_1, e_2 di $\mathbb{R}^2$.



Rappresentiamo ciascuna applicazione lineare F con una matrice rispetto alla base canonica $\mathcal{E} : e_1, e_2$, e la indichiamo con [F] (sottintendendo la scelta della base).

R_+ è l'unica applicazione lineare tale che

$$\begin{array}{l} \mathrm{R}_+(e_1) = e_2 \\ \mathrm{R}_+(e_2) = -e_1 \end{array} \quad \text{cioè } [\mathrm{R}_+] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix};$$

quindi, esplicitamente

$$\mathrm{R}_+(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix};$$

altrimenti scritto

$$\mathrm{R}_+(x_1, x_2) = (-x_2, x_1).$$

Pr è l'unica applicazione lineare tale che

$$\begin{array}{r} \mathrm{Pr}(e_1) = e_1 \\ \mathrm{Pr}(e_2) = \underline{0} \end{array} \quad \text{cioè } [\mathrm{Pr}] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};$$

quindi, esplicitamente

$$\mathrm{Pr}(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix};$$

altrimenti scritto

$$\mathrm{Pr}(x_1, x_2) = (x_1, 0).$$

Di seguito, calcoliamo l'applicazione composta $\mathrm{Pr} \circ \mathrm{R}_+$, in tre modi diversi:

- calcolando le immagini dei vettori base:

$(\mathrm{Pr} \circ \mathrm{R}_+)(e_1) = \mathrm{Pr}(\mathrm{R}_+(e_1)) = \mathrm{Pr}(e_2) = \underline{0}$
$(\mathrm{Pr} \circ \mathrm{R}_+)(e_2) = \mathrm{Pr}(\mathrm{R}_+(e_2)) = \mathrm{Pr}(-e_1) = -e_1$

- usando le matrici:

$$\begin{aligned}(\mathrm{Pr} \circ \mathrm{R}_+)(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) &= \mathrm{Pr}(\mathrm{R}_+(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix})) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} -x_2 \\ 0 \end{bmatrix}\end{aligned}$$

- usando solo la definizione di applicazione composta:

$$(\mathrm{Pr} \circ \mathrm{R}_+)(x_1, x_2) = \mathrm{Pr}(\mathrm{R}_+(x_1, x_2)) = \mathrm{Pr}(-x_2, x_1) = (-x_2, 0).$$

Per quanto riguarda biiettività e inversa:

R_+ ha come inversa l'applicazione R_- rotazione di $\frac{\pi}{2}$ in senso orario, descritta da

$$\begin{aligned}\mathrm{R}_-(e_1) &= -e_2 \\ \mathrm{R}_-(e_2) &= \quad e_1\end{aligned} \quad \text{cioè } [\mathrm{R}_-] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

esplicitamente:

$$\mathrm{R}_-(x_1, x_2) = (x_2, -x_1);$$

una verifica:

$$(\mathrm{R}_- \circ \mathrm{R}_+)(x_1, x_2) = \mathrm{R}_-(\mathrm{R}_+(x_1, x_2)) = \mathrm{R}_-(-x_2, x_1) = (x_1, x_2),$$

quindi $\mathrm{R}_- \circ \mathrm{R}_+ = \mathrm{Id}$;

In definitiva, in simboli:

$$(\mathrm{R}_+)^{-1} = \mathrm{R}_-.$$

Pr non è biiettiva; motiviamo l'affermazione in due modi:

- l'affermazione

“ per ogni y, l'equazione $\mathrm{Pr}(x) = y$ ha una ed una sola soluzione”

è falsa, in quanto l'equazione $\mathrm{Pr}(x) = e_2$ non ha alcuna soluzione (oppure, in quanto l'equazione $\mathrm{Pr}(x) = e_1$ ha infinite soluzioni);

- la matrice $[\mathrm{Pr}] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ è singolare.

Quindi, Pr non possiede inversa.

Applicazioni fra spazi $\mathbb{R}^n$

Proposizione. Sono date due applicazioni lineari fra spazi $\mathbb{R}^*$

$\mathrm{F}(x) = Ax$ (A matrice costante; x colonna variabile);

$\mathrm{G}(y) = By$ (B matrice costante; y colonna variabile);

Allora l'applicazione composta $\mathrm{G}\circ\mathrm{F}$ è definita se e solo se è definita la matrice prodotto BA; in tal caso

$$(\mathrm{G}\circ\mathrm{F})(x) = (BA)x.$$

In altri termini: date due applicazioni lineari F, G fra spazi $\mathbb{R}^*$ con matrici $[\mathrm{F}], [\mathrm{G}]$ rispetto alle basi canoniche, la composizione $\mathrm{G}\circ\mathrm{F}$ è definita se e solo se il prodotto $[\mathrm{G}][\mathrm{F}]$ è definito, e in tal caso

$$[\mathrm{G}\circ\mathrm{F}] = [\mathrm{G}][\mathrm{F}].$$

Dimostrazione. Le seguenti affermazioni sono equivalenti:

$\mathrm{G}\circ\mathrm{F}$ è definita;
per opportuni m, n, p si ha $\mathrm{F}: \mathbb{R}^n \to \mathbb{R}^m$ e $\mathrm{G}: \mathbb{R}^m \to \mathbb{R}^p$;
per opportuni m, n, p si ha A, B hanno tipi $m \times n$ e $p \times m$;
BA è definito.

Inoltre, in tal caso

$$(\mathrm{G}\circ\mathrm{F})(x) = \mathrm{G}(\mathrm{F}(x)) = B(Ax) = (BA)x$$

(per l'associatività del prodotto di matrici).

Proposizione. Un'applicazione lineare

$$\mathrm{F}: \mathbb{R}^n \to \mathbb{R}^m,\ \mathrm{F}(x) = Ax.$$

è invertibile se e solo se la matrice A è invertibile (quindi in particolare $n = m$) e in tal caso l'applicazione inversa è

$$\mathrm{F}: \mathbb{R}^n \to \mathbb{R}^n,\ \mathrm{F}^{-1}(x) = A^{-1}x.$$

In altri termini: un' applicazione lineare F fra spazi $\mathbb{R}^*$ con matrice $[\mathrm{F}]$ rispetto alle basi canoniche è invertibile se e solo se la matrice $[\mathrm{F}]$ è invertibile, e in tal caso

$$[\mathrm{F}^{-1}] = [\mathrm{F}]^{-1}.$$

Dimostrazione parziale. Supponiamo che la matrice A sia invertibile. Per ogni $y \in \mathbb{R}^n$, l'equazione $\mathrm{F}(x) = y$ equivale all'equazione

$$Ax = y,$$

che ha sempre un'unica soluzione, data da $x = A^{-1}y$. Quindi, F è invertibile, con inversa $\mathrm{F}^{-1}: \mathbb{R}^n \to \mathbb{R}^n$, $\mathrm{F}^{-1}(y) = A^{-1}y$.

Esempio.

$$\overset{\mathrm{G}}{\circlearrowright}\ \mathbb{R}^2 \xrightarrow{\mathrm{F}} \mathbb{R}^3\ \overset{\mathrm{H}}{\circlearrowleft}$$

$$\mathrm{F}(x_1, x_2) = (x_1, x_1 + x_2, x_2);$$

$$\mathrm{G}(x_1, x_2) = (x_1 + x_2, x_1 - x_2);$$

$$\mathrm{H}(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 + x_3);$$

$$[F] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix},\ [G] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},\ [H] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Consideriamo le applicazioni F e G, ci chiediamo se sono componibili e nel caso calcoliamo l'applicazione composta. Lo facciamo in due modi.

- Usando solo la definizione di applicazione composta:

$\mathrm{G} \circ \mathrm{F}$ non è definita in quanto il dominio di G è $\mathbb{R}^2$, diverso dal codominio $\mathbb{R}^3$ di F;

$\mathrm{F} \circ \mathrm{G} : \mathbb{R}^2 \to \mathbb{R}^3$,

$$\begin{aligned}(\mathrm{F} \circ \mathrm{G})(x_1, x_2) &= \mathrm{F}(\mathrm{G}(x_1, x_2)) \\ &= \mathrm{F}(x_1 + x_2, x_1 - x_2) \\ &= (x_1 + x_2, x_1 + x_2 + x_1 - x_2, x_1 - x_2) \\ &= (x_1 + x_2, 2x_1, x_1 - x_2).\end{aligned}$$

- Usando le matrici.

$[\mathrm{G}][\mathrm{F}]$ non è definita in quanto $[\mathrm{G}]$ è 2×2 e $[\mathrm{F}]$ è 3×2;

$[\mathrm{F}][\mathrm{G}]$ è definita ed è 3×2;

$$[\mathrm{F} \circ \mathrm{G}] = [\mathrm{F}][\mathrm{G}] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{bmatrix}$$

Quindi $\mathrm{G} \circ \mathrm{F}$ non è definita e $\mathrm{F} \circ \mathrm{G} : \mathbb{R}^2 \to \mathbb{R}^3$ è

$(\mathrm{F} \circ \mathrm{G})(x_1, x_2) = (x_1 + x_2, 2x_1, x_1 - x_2)$.

Consideriamo ancora le applicazioni F e G, ci chiediamo se sono biiettive e nel caso calcoliamo l'applicazione inversa.

$[\mathrm{F}]$ è 3×2, non quadrata, non possiede inversa;

$[\mathrm{G}]$ è (quadrata) non singolare; calcoliamo la sua inversa con l'algoritmo di Gauss-Jordan

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right] \rightsquigarrow \cdots \rightsquigarrow \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right]$$

$$[G]^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Quindi non esiste F^{-1} ed esiste $\mathrm{G}^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$,

$\mathrm{G}^{-1}(x_1, x_2) = (\frac{1}{2}x_1 + \frac{1}{2}x_2, \frac{1}{2}x_1 - \frac{1}{2}x_2)$.

Applicazioni fra spazi vettoriali

Dalla definizione di applicazione lineare segue che:

- Se due applicazioni $\mathrm{F} : V \to W$ e $\mathrm{G} : W \to Z$ fra spazi vettoriali sono lineari, allora anche l'applicazione composta $\mathrm{G} \circ \mathrm{F} : V \to Z$ è lineare;

- Per ogni spazio vettoriale V, l'identità id_V è lineare;

- Se un'applicazione $\mathrm{F} : V \to W$ invertibile è lineare, allora anche l'applicazione inversa $\mathrm{F}^{-1} : W \to V$ è lineare.

La rappresentazione di applicazioni lineari rispetto a basi qualsiasi si comporta bene rispetto alla composizione e inversione, precisamente:

Proposizione.

(1) Siano date due applicazioni lineari fra spazi vettoriali con basi fissate

$$V_{\mathcal{A}} \xrightarrow{\mathrm{F}} W_{\mathcal{B}} \xrightarrow{\mathrm{G}} Z_{\mathcal{C}}$$

e l'applicazione composta

$$V_{\mathcal{A}} \xrightarrow{\mathrm{G}\circ\mathrm{F}} Z_{\mathcal{C}}.$$

Allora, rispetto alle dovute basi, la matrice di $\mathrm{G} \circ \mathrm{F}$ è il prodotto delle matrici di G e F

$$[\mathrm{G} \circ \mathrm{F}]_{\mathcal{CA}} = [\mathrm{G}]_{\mathcal{CB}}[\mathrm{F}]_{\mathcal{BA}}.$$

(2) La matrice dell'applicazione identità su uno spazio vettoriale V di dimensione n, rispetto alla stessa base in dominio e codominio, è la matrice identità I_n :

$$[\mathrm{id}_V]_{\mathcal{AA}} = \mathrm{I}_n.$$

(3) Sia data un'applicazione lineare fra due spazi vettoriali con base fissata

$$V_{\mathcal{A}} \xrightarrow{\mathrm{F}} W_{\mathcal{B}}$$

che possiede inversa

$$W_{\mathcal{B}} \xrightarrow{\mathrm{F}^{-1}} V_{\mathcal{A}}.$$

Allora, rispetto alle dovute basi la matrice di F^{-1} è l'inversa dela matrice di F :

$$[\mathrm{F}^{-1}]_{\mathcal{AB}} = \left([\mathrm{F}]_{\mathcal{BA}}\right)^{-1}.$$

Dimostrazione.

(1) Per ogni $v \in V$ si ha

$$[(\mathrm{G} \circ \mathrm{F})(v)]_{\mathcal{C}} = [\mathrm{G} \circ \mathrm{F}]_{\mathcal{CA}}[v]_{\mathcal{A}}$$

$$[(\mathrm{G} \circ \mathrm{F})(v)]_{\mathcal{C}} = [\mathrm{G}(\mathrm{F}(v))]_{\mathcal{C}} = [\mathrm{G}]_{\mathcal{CB}}[\mathrm{F}(v)]_{\mathcal{B}} = [\mathrm{G}]_{\mathcal{CB}}[\mathrm{F}]_{\mathcal{BA}}[v]_{\mathcal{A}}$$

quindi $[\mathrm{G} \circ \mathrm{F}]_{\mathcal{CA}} = [\mathrm{G}]_{\mathcal{CB}}[\mathrm{F}]_{\mathcal{BA}}$.

(2) Lasciata al lettore.

(3) Dalle uguaglianze fra applicazioni

$$\mathrm{F}^{-1} \circ \mathrm{F} = \mathrm{Id}_V = \mathrm{F} \circ \mathrm{F}^{-1}$$

seguono le uguaglianze fra matrici

$$[\mathrm{F}^{-1}]_{\mathcal{AB}}[\mathrm{F}]_{\mathcal{BA}} = \mathrm{I}_n = [\mathrm{F}]_{\mathcal{BA}}[\mathrm{F}^{-1}]_{\mathcal{AB}}$$

che significano proprio

$$[\mathrm{F}^{-1}]_{\mathcal{AB}} = \left([\mathrm{F}]_{\mathcal{BA}}\right)^{-1}.$$

Algebre di matrici e di applicazioni lineari

5.3 Autovettori ed autovalori

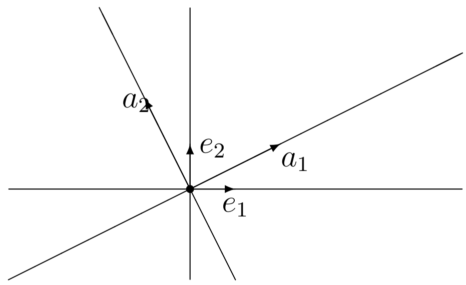
Contesto.

Uno spazio vetoriale V di un certa dimensione n. Consideriamo applicazioni lineari F di V in sè e le rappresentiamo rispetto ad una stessa base $\mathcal{A}$ sia nel dominio che nel codominio; al posto $[\mathrm{F}]_{\mathcal{A}\mathcal{A}}$, scriviamo in breve $[\mathrm{F}]_{\mathcal{A}}$.

Caso di studio

Consideriamo la base ortogonale di $\mathcal{V}^2$

$$\mathcal{A} : a_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$



Sia F la simmetria ortogonale rispetto alla retta vettoriale generata da a_1;

$$\begin{array}{l} \mathrm{F}(a_1) = a_1 \\ \mathrm{F}(a_2) = -a_2 \end{array} \quad \text{quindi } [\mathrm{F}]_{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Di seguito, prima determiniamo la matrice $[\mathrm{F}]_{\mathcal{E}}$ dell'applicazione rispetto alla base canonica, poi, a partire solo da tale matrice, vediamo come si può riconoscere che l'applicazione è una simmetria ortogonale e come si può determinare l'asse di simmetria e il suo asse ortogonale. Per determinare $[\mathrm{F}]_{\mathcal{E}}$ considereremo il problema generale della relazione fra le varie matrici che rappresentano una stessa applicazione lineare. Nel riconoscere la natura dell'applicazione, introdurremo i concetti di autovalore ed autovettore di un'applicazione lineare.

Relazione fra le matrici di una stessa applicazione lineare

Siano V uno spazio vettoriale, di dimensione n e F un'applicazione lineare di V in sé. Date due basi $\mathcal{A} : a_1, \ldots, a_n$ e $\mathcal{B} : b_1, \ldots, b_n$ di V, descriviamo la matrice $[\mathrm{F}]_{\mathcal{B}}$ in funzione della matrice $[\mathrm{F}]_{\mathcal{A}}$. Prima di tutto, consideriamo delle matrici che descrivono relazioni fra le basi.

Matrici dell'identità

L'identità $\mathrm{id} : V_{\mathcal{A}} \to V_{\mathcal{B}}$, rispetto alla base $\mathcal{A}$ di V come dominio e alla base $\mathcal{B}$ di V come codominio, è rappresentata dalla matrice che ha nelle colonne le coordinate rispetto a $\mathcal{B}$ dei valori dell'identità sui vettori di $\mathcal{A}$

$$[\mathrm{id}]_{\mathcal{B}\mathcal{A}} = [(\mathrm{id}(a_1))_{\mathcal{B}} \cdots (\mathrm{id}(a_n))_{\mathcal{B}}]$$

quindi dalla matrice che ha nelle colonne le coordinate rispetto a $\mathcal{B}$ dei vettori di $\mathcal{A}$

$[\mathrm{id}]_{\mathcal{B}\mathcal{A}} = [(a_1)_{\mathcal{B}} \cdots (a_n)_{\mathcal{B}}];$

la diciamo in breve “matrice, rispetto a $\mathcal{B}$, di $\mathcal{A}$”

Caso particolare. Se $V = \mathbb{R}^n$ ed $\mathcal{E}$ è la base canonica $\mathbb{R}^n$, essendo $(a_i)_{\mathcal{E}} = a_i$, si ha

$[\mathrm{id}]_{\mathcal{E}\mathcal{A}} = [a_1 \cdots a_n],$

in parole: la matrice rispetto ad $\mathcal{E}$ di $\mathcal{A}$ è la matrice con colonne i vettori di $\mathcal{A}$.

Caso di studio, esempio

$$[\mathrm{id}]_{\mathcal{E}\mathcal{A}} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}.$$

Fine esempio.

Dal fatto che $\mathrm{id} = \mathrm{id} \circ \mathrm{id}$ segue che

$[\mathrm{id}]_{\mathcal{A}\mathcal{A}} = [\mathrm{id}]_{\mathcal{A}\mathcal{B}}[\mathrm{id}]_{\mathcal{B}\mathcal{A}}$

$[\mathrm{id}]_{\mathcal{B}\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}\mathcal{A}}[\mathrm{id}]_{\mathcal{A}\mathcal{B}};$

essendo $[\mathrm{id}]_{\mathcal{A}\mathcal{A}} = [\mathrm{id}]_{\mathcal{B}\mathcal{B}} = \mathrm{I}$ (matrice identità), si ha che le matrici $[\mathrm{id}]_{\mathcal{B}\mathcal{A}}$ e $[\mathrm{id}]_{\mathcal{A}\mathcal{B}}$ sono una l’inversa dell’altra; in altri termini: la matrice rispetto a $\mathcal{A}$ di $\mathcal{B}$ è l’inversa della matrice rispetto a $\mathcal{B}$ di $\mathcal{A}$:

$[\mathrm{id}]_{\mathcal{A}\mathcal{B}} = ([\mathrm{id}]_{\mathcal{B}\mathcal{A}})^{-1}$

Caso di studio, esempio

$$[\mathrm{id}]_{\mathcal{A}\mathcal{E}} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix}.$$

Fine esempio.

Relazione fra le matrici di una stessa applicazione

Componendo le applicazioni

$$V_{\mathcal{B}} \xrightarrow{\mathrm{id}} V_{\mathcal{A}} \xrightarrow{\mathrm{F}} V_{\mathcal{A}} \xrightarrow{\mathrm{id}} V_{\mathcal{B}};$$

si ha l’applicazione

$$V_{\mathcal{B}} \xrightarrow{\mathrm{F}} V_{\mathcal{B}}$$

quindi la relazione fra le matrici dell’applicazione rispetto alle due basi è

$$[\mathrm{F}]_{\mathcal{B}\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}\mathcal{A}}[\mathrm{F}]_{\mathcal{A}\mathcal{A}}[\mathrm{id}]_{\mathcal{A}\mathcal{B}}$$

Caso particolare. Se $V = \mathbb{R}^n$ e $\mathcal{E}$ è la sua base canonica, allora

$$\begin{aligned} [\mathrm{F}]_{\mathcal{E}\mathcal{E}} &= [\mathrm{id}]_{\mathcal{E}\mathcal{A}}[\mathrm{F}]_{\mathcal{A}\mathcal{A}}[\mathrm{id}]_{\mathcal{A}\mathcal{E}} \\ &= [a_1 \cdots a_n][\mathrm{F}]_{\mathcal{A}\mathcal{A}}[a_1 \cdots a_n]^{-1} \end{aligned}$$

Caso di studio, esempio

$$\begin{aligned} [\mathrm{F}]_{\mathcal{E}} &= \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix} \\ &= \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}. \end{aligned}$$

Quindi

$$F(x_1, x_2) = (\tfrac{3}{5}\,x_1 + \tfrac{4}{5}\,x_1,\ \tfrac{4}{5}\,x_1 - \tfrac{3}{5}\,x_1).$$

Verifica: F è l'unica applicazione lineare di $\mathbb{R}^2$ in sè tale che $F(a_1) = a_1$ e $F(a_2) = -a_2$. Verifichiamo la 1°:

$$(\tfrac{6}{5} + \tfrac{4}{5},\ \tfrac{8}{5} - \tfrac{3}{5}) = (2, 1)? \quad \text{si}$$

Fine esempio.

Autovettori ed autovalori

Caso di studio

Abbiamo trovato che l'applicazione lineare F di $\mathbb{R}^2$ in sè, rispetto alla base canonica ha matrice

$$[F] = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}.$$

Usando solo questa informazione, possiamo ricostruire che F è una simmetria ortogonale, l'asse di simmetria (e quindi il suo asse ortogonale)? e i vettori a_1, a_2?

Osserviamo che a_1, a_2 sono accomunati dal fatto di essere mandati in loro multipli

$$F(a_1) = a_1$$
$$F(a_2) = -a_2$$

Autovettori e autovalori.

Definizione. Sia F un'applicazione lineare di uno spazio vettoriale V in sé. Si dice che un vettore $v \neq \underline{0}$ ed un numero λ sono un "autovettore" e un "autovalore" di F se e solo se il valore di F su v è il prodotto di λ per v :

$(*)\qquad F(v) = \lambda\, v;$

in altri termini:

- $v \neq \underline{0}$ è un autovettore di F se e solo se esiste un λ tale che valga $(*)$;

- λ è un autovalore di F se e solo se esiste un $v \neq \underline{0}$ tale che valga $(*)$.

Caso di studio, esempio.

λ è un autovalore di F se e solo se esiste un $x \neq \underline{0}$ tale che

$$\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} \frac{3}{5}\,x_1 + \frac{4}{5}\,x_2 = \lambda\,x_1 \\ \frac{4}{5}\,x_1 - \frac{3}{5}\,x_2 = \lambda\,x_2 \end{cases}$$

$$(\dagger)\qquad \begin{cases} (\ \frac{3}{5} - \lambda)\,x_1 + \frac{4}{5}\,x_2 = 0 \\ \frac{4}{5}\,x_1 + (-\frac{3}{5} - \lambda)\,x_2 = 0 \end{cases}$$

un sistema lineare omogeneo con matrice dei coefficienti

$$\begin{bmatrix} \frac{3}{5} - \lambda & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} - \lambda \end{bmatrix};$$

il sistema ha qualche soluzione $x \neq \underline{0}$ se e solo se la matrice è singolare, cioè

$$\det \begin{bmatrix} \frac{3}{5} - \lambda & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} - \lambda \end{bmatrix} = 0; \;\; (\tfrac{3}{5} - \lambda)(-\tfrac{3}{5} - \lambda) - \tfrac{4}{5}\tfrac{4}{5} = 0; \;\; \lambda^2 - 1 = 0, \;\; \lambda = \pm 1.$$

Quindi F ha due autovalori: 1 e -1.

L'insieme degli autovettori con autovalore 1, aggiunto $\underline{0}$, è l'insieme delle soluzioni del sistema (†) con $\lambda = 1$

$$\begin{cases} -\frac{2}{5}\,x_1 + \frac{4}{5}\,x_2 = 0 \\ \frac{4}{5}\,x_1 - \frac{8}{5}\,x_2 = 0 \end{cases}, \qquad x_1 - 2\,x_2 = 0;$$

è l'insieme i vettori

$$\begin{bmatrix} 2\,x_2 \\ x_2 \end{bmatrix}, \text{ con } x_2 \text{ libera;}$$

è una retta vettoriale, che indichiamo con V_1; abbiamo ritrovato l'asse della simmetria.

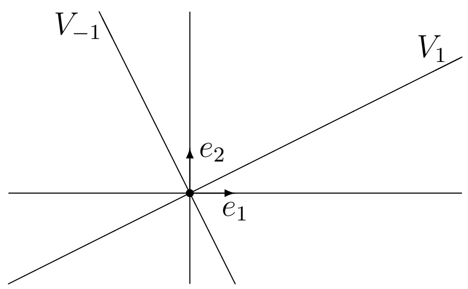
L'insieme degli autovettori con autovalore -1, aggiunto $\underline{0}$, è l'insieme delle soluzioni del sistema (†) con $\lambda = -1$

$$\begin{cases} \frac{8}{5}\,x_1 + \frac{4}{5}\,x_2 = 0 \\ \frac{4}{5}\,x_1 + \frac{2}{5}\,x_2 = 0 \end{cases}, \qquad 2\,x_1 + x_2 = 0;$$

è l'insieme dei vettori

$$\begin{bmatrix} x_1 \\ -2\,x_1 \end{bmatrix}, \text{ con } x_1 \text{ libera;}$$

è una retta vettoriale, che indichiamo con V_{-1}; è l'asse ortogonale all'asse di simmetria.



Osservazione: se b_1, b_2 sono due qualsiasi autovettori con autovalori rispettivi 1 e -1, cioè due vettori non nulli sulle rette vettoriali V_1 e V_{-1}, allora $\mathcal{B} : b_1, b_2$ è una base di $\mathbb{R}^2$ e

$$[\mathrm{F}]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Fine esempio.

Equazione caratteristica, autospazi

Sia $\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^n$, $\mathrm{F}(x) = Ax$. Allora l'equazione $(*)$ degli autovettori e autovalori diviene

$$Ax = \lambda\, x$$

per esteso

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = \lambda\, x_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = \lambda\, x_2 \\ & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = \lambda\, x_n \end{cases}$$

che equivale al sistema omogeneo

$$\begin{cases} (a_{11}-\lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + a_{2n}x_n & = 0 \\ & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n & = 0 \end{cases}$$

con matrice dei coefficienti

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix} = A - \lambda\,\mathrm{I}.$$

Il sistema ha qualche soluzione $\neq \underline{0}$ se e solo se la matrice $A - \lambda\,\mathrm{I}$ è singolare, cioè

$$\det(A - \lambda\,\mathrm{I}) = 0.$$

Quindi: gli autovalori di F sono le soluzioni di questa equazione, che si dice "equazione caratteristica" di A.

L'insieme $V_{\overline{\lambda}}$ degli autovettori con un dato autovalore $\overline{\lambda}$, con l'aggiunta di $\underline{0}$, si dice "autospazio" di F con autovalore $\overline{\lambda}$, è l'insieme delle soluzioni del sistema lineare omogeneo

$$(A - \lambda\,\mathrm{I})x = \underline{0},$$

è lo spazio nullo della matrice

$$V_{\overline{\lambda}} = \mathcal{N}(A - \lambda\mathrm{I}),$$

è un sottospazio di $\mathbb{R}^n$, con dimensione $n - \mathrm{r}(A - \lambda\,\mathrm{I})$, sempre > 0.

Diagonalizzazione

Fino ad avviso contrario, siano V uno spazio vettoriale di dimensione finita n e F un'applicazione lineare di V in sé. Di regola, negli esempi ed esercizi considereremo $V = \mathbb{R}^n$ e F data esplicitamente come $\mathrm{F}(v) = Av$ con A matrice costante (la matrice di F rispetto alla base canonica).

Applicazioni diagonalizzabili

Terminologia. Per ogni base $\mathcal{A}$ di V, l'applicazione F è rappresentata da una matrice $[\mathrm{F}]_\mathcal{A}$; diciamo che F è "rappresentabile" da una matrice M se esiste una base $\mathcal{A}$ di V tale che $[\mathrm{F}]_\mathcal{A} = M$; diciamo che F è "diagonalizzabile" se è rappresentabile da qualche matrice diagonale.

Per una data applicazione F ci poniamo le seguenti questioni:

(1) Stabilire se F è diagonalizzabile;

(2) determinare tutte le matrici diagonali che rappresentano F;

(3) per ogni matrice diagonale D che rappresenta F, determinare una base $\mathcal{A}$ di V tale che $[\mathrm{F}]_\mathcal{A} = D$.

Autovettori e autovalori.

Terminologia. Diciamo che un vettore $v \neq \underline{0}$ e un numero λ sono un autovettore e un autovalore di F fra loro "associati" se

$$\mathrm{F}(v) = \lambda\, v.$$

Fatti.

- A ciascun autovettore v è associato un unico autovalore. Infatti: $\mathrm{F}(v) = \lambda_1 v$ e $\mathrm{F}(v) = \lambda_1 v$ implicano $\lambda_1 v = \lambda_2 v$ implica $(\lambda_1 - \lambda_2)v = \underline{0}$ implica $\lambda_1 - \lambda_2 = 0$, implica $\lambda_1 = \lambda_2$ (il penultimo passaggio segue da $v \neq \underline{0}$).

- A ciascun autovalore $\overline{\lambda}$ sono associati infiniti autovettori che, con il vettore nullo, formano un sottospazio di V, che diciamo autospazio associato a $\overline{\lambda}$ e indichiamo con $V_{\overline{\lambda}}$.

Matrici diagonali, autovettori e autovalori.

Fatto. Per ciascuna base $\mathcal{A} : a_1, a_2, \ldots, a_n$ di V le due affermazioni sono equivalenti:

- $[\mathrm{F}]_\mathcal{A}$ è diagonale;

- $a_1, a_2, \ldots, a_n$ sono autovettori di F;

in tal caso, gli elementi diagonali di $[\mathrm{F}]_\mathcal{A}$ sono gli autovalori associati agli autovettori.

Infatti, $[\mathrm{F}]_\mathcal{A} = D$ diagonale se e solo se

$$[(\mathrm{F}a_1)_\mathcal{A}(\mathrm{F}a_2)_\mathcal{A} \cdots (\mathrm{F}a_n)_\mathcal{A}] = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

se e solo se

$$\begin{aligned}
&\mathrm{F}(a_1) = d_1\, a_1 + 0\, a_2 + \cdots + 0\, a_n = d_1\, a_1\\
&\mathrm{F}(a_2) = 0\, a_1 + d_2\, a_2 + \cdots + 0\, a_n = d_2\, a_2\\
&\vdots\\
&\mathrm{F}(a_n) = 0\, a_1 + 0\, a_2 + \cdots + d_n\, a_n = d_n\, a_n
\end{aligned}$$

Quindi le questioni esposte sopra si possono tradurre in termini di autovettori ed autovalori come segue

(1) Stabilire se esiste una base di V di autovettori di F;

(2) Determinare le sequenze di autovalori associate alle varie basi di autovettori;

(3) Per ogni sequenza ammissibile di autovalori, determinare una base di autovettori cui è associata.

Nel caso F applicazione lineare di $\mathbb{R}^n$ in sé, $\mathrm{F}(v) = Av$, l'equazione degli autovettori ed autovalori è

$$A\, v = \lambda\, v, \qquad (v \neq \underline{0}),$$

cioè

$$(A - \lambda\, \mathrm{I})\, v = \underline{0}, \qquad (v \neq \underline{0}),$$

gli autovalori sono le soluzioni dell'equazione caratteristica

$$\det(A - \lambda\, \mathrm{I}) = 0;$$

per ogni autovalore $\overline{\lambda}$ si ha un autospazio

$$V_{\overline{\lambda}} = \mathcal{N}(A - \overline{\lambda}\, \mathrm{I}).$$

Per ogni autospazio, si cerca una sua base, e con queste basi si cerca di costruire una base di V. Di seguito mostriamo i primi risultati della teoria di autovettori ed autovalori e usando queti risultati diamo una strategia efficace.

Qualche esempio in dimensione 2

Fatto. Per ogni due autovettori a, b con autovalori associati α, β,

se a, b sono linearmente dipendenti allora $\alpha = \beta$, cioè
se $\alpha \neq \beta$ allora a, b sono linearmente indipendenti.

Infatti: a, b autovettori linearmente dipendenti implica $b = ra$ per qualhce numero r implica $\mathrm{F}(b) = \mathrm{F}(ra)$ implica $\beta b = r\alpha a$ implica $\beta b = \alpha b$ implica $\beta = \alpha$.

Esempio 1

$$\mathrm{F}(x) = \begin{bmatrix} 0 & 9 \\ 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^2)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 0 & 9 \\ 1 & 0 \end{bmatrix} x = \lambda\, x, \qquad \begin{bmatrix} -\lambda & 9 \\ 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -\lambda & 9 \\ 1 & -\lambda \end{bmatrix} = 0, \;\; \lambda^2 - 9 = 0, \;\; \lambda = \pm 3.$$

Ci sono 2 autovalori distinti; 2 autovettori ad essi associati, comunque scelti, sono linearmente indipendenti, quindi sono una base di $\mathbb{R}^2$. F è diagonalizzabile.

Autospazio V_3 : spazio delle soluzioni di

$$\begin{bmatrix} 0 & 9 \\ 1 & 0 \end{bmatrix} x = 3\,x, \qquad \begin{bmatrix} -3 & 9 \\ 1 & -3 \end{bmatrix} x = \underline{0};$$

un sistema singolare di 2 equazioni in 2 incognite, che equivale all'unica equazione

$$x_1 - 3\,x_2 = 0;\ \text{ con soluzioni } \begin{bmatrix} 3\,x_2 \\ x_2 \end{bmatrix} \ (x_2 \text{ libera}); \text{ una base è } a = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Autospazio V_{-3} : spazio delle soluzioni di

$$\begin{bmatrix} 0 & 9 \\ 1 & 0 \end{bmatrix} x = -3\,x, \qquad \begin{bmatrix} 3 & 9 \\ 1 & 3 \end{bmatrix} x = \underline{0};$$

un sistema che equivale all'unica equazione

$$x_1 + 3\,x_2 = 0;\ \text{ con soluzioni } \begin{bmatrix} -3\,x_2 \\ x_2 \end{bmatrix} \ (x_2 \text{ libera}); \text{ una base è } b = \begin{bmatrix} -3 \\ 1 \end{bmatrix}.$$

Rispetto alla base a, b, la matrice di F è

$$[\mathrm{F}]_{a,b} = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix},$$

rispetto alla base b, a è

$$[\mathrm{F}]_{b,a} = \begin{bmatrix} -3 & 0 \\ 0 & 3 \end{bmatrix}.$$

Non ci sono altre matrici diagonali che rappresentano F; infatti: 2 autovettori con autovalore associato 3, essendo nell'autospazio V_3 che ha dimensione 1, sono linearmente dipendenti quindi non sono base di $\mathbb{R}^2$; allo stesso modo 2 autovettori con autovalore associato -3 non sono base di $\mathbb{R}^2$.

Esempio 2

$$\mathrm{R}(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^2)$$

R si può iterpretare come la rotazione di $\frac{\pi}{2}$ in senso antiorario; in particolare, manda ciascun vettore non nullo in un vettore con direzione diversa, quindi non possiede alcun autovettore; non è diagonalizzabile. Svolgiamo comunque i conti.

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x = \lambda\,x, \qquad \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = 0,\ \ \lambda^2 + 1 = 0,\ \ \text{nessuna soluzione in } \mathbb{R}.$$

Non ci sono autovalori reali, quindi non ci sono autovettori, R non è diagonalizzabile sui reali (lo è sui complessi).

Esempio 3

$$G(x) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x, \quad (x \in \mathbb{R}^2)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x = \lambda\, x, \qquad \begin{bmatrix} -\lambda & 0 \\ 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -\lambda & 0 \\ 1 & -\lambda \end{bmatrix} = 0, \;\; \lambda^2 = 0, \;\; \text{due soluzioni coincidenti in } \lambda = 0.$$

Autospazio V_0 : spazio delle soluzioni di

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x = 0\, x, \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x = \underline{0};$$

un sistema che equivale all'unica equazione

$$x_1 = 0; \quad \text{con soluzioni} \begin{bmatrix} 0 \\ x_2 \end{bmatrix} (x_2 \text{ libera}).$$

V_0 è una retta vettoriale; 2 vettori di V_0 comunque scelti non sono una base di $\mathbb{R}^2$; G non è diagonalizzabile.

Tutte e tre queste applicazioni sono casi particolari di una famiglia di applicazioni, che consideriamo qui di seguito.

Esercizio 4. E' data la famiglia di applicazioni

$$F_k(x) = \begin{bmatrix} 0 & k \\ 1 & 0 \end{bmatrix} x, \quad (x \in \mathbb{R}^2)$$

dove k è un parametro $\in \mathbb{R}$.

a) Determinare per quali k l'applicazione F_k è diagonalizzabile;

b) Per ciascuno di tali k si scrivano le matrici diagonali che rappresentano F_k;

c) Per ciascuno di tali k e ciascuna di tali matrici diagonali, si scriva una base di autovettori rispetto alla quale F_k è rappresentata dalla matrice diagonale.

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 0 & k \\ 1 & 0 \end{bmatrix} x = \lambda\, x, \qquad \begin{bmatrix} -\lambda & k \\ 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -\lambda & k \\ 1 & -\lambda \end{bmatrix} = 0, \;\; \lambda^2 - k = 0.$$

Distinguiamo tre casi

$k > 0$, due soluzioni reali distinte $\lambda = \pm\sqrt{k}$;

$k = 0$, due soluzioni reali coincidenti in $\lambda = 0$;

$k < 0$; nessuna soluzione reale.

Caso $k > 0$. Ci sono 2 autovalori distinti, $\pm\sqrt{k}$; 2 autovettori ad essi associati, comunque scelti, sono linearmente indipendenti, quindi sono una base di $\mathbb{R}^2$. F_k è diagonalizzabile.

Autospazio $V_{\sqrt{k}}$: spazio delle soluzioni di

$$\begin{bmatrix} 0 & k \\ 1 & 0 \end{bmatrix} x = \sqrt{k}\, x, \qquad \begin{bmatrix} -\sqrt{k} & k \\ 1 & -\sqrt{k} \end{bmatrix} x = \underline{0};$$

un sistema singolare di 2 equazioni in 2 incognite, che equivale all'unica equazione

$$x_1 - \sqrt{k}\, x_2 = 0, \text{ con soluzioni } \begin{bmatrix} \sqrt{k}\, x_2 \\ x_2 \end{bmatrix} (x_2 \text{ libera}); \text{ una base è } a_k = \begin{bmatrix} \sqrt{k} \\ 1 \end{bmatrix}.$$

Autospazio $V_{-\sqrt{k}}$: spazio delle soluzioni di

$$\begin{bmatrix} 0 & k \\ 1 & 0 \end{bmatrix} x = -\sqrt{k}\, x, \qquad \begin{bmatrix} \sqrt{k} & k \\ 1 & \sqrt{k} \end{bmatrix} x = \underline{0};$$

un sistema che equivale all'unica equazione

$$x_1 + \sqrt{k}\, x_2 = 0, \text{ con soluzioni } \begin{bmatrix} -\sqrt{k}\, x_2 \\ x_2 \end{bmatrix} (x_2 \text{ libera}); \text{ una base è } b_k = \begin{bmatrix} -\sqrt{k} \\ 1 \end{bmatrix}.$$

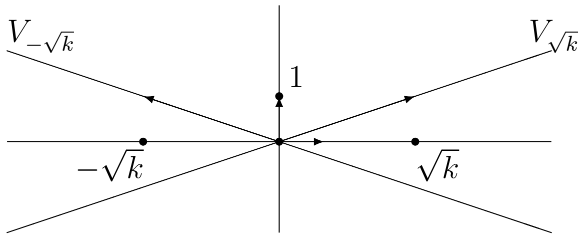
Rispetto alla base a_k, b_k, la matrice di F_k è

$$[\mathrm{F}_k]_{a_k, b_k} = \begin{bmatrix} \sqrt{k} & 0 \\ 0 & -\sqrt{k} \end{bmatrix},$$

rispetto alla base b_k, a_k è

$$[\mathrm{F}_k]_{b_k, a_k} = \begin{bmatrix} -\sqrt{k} & 0 \\ 0 & \sqrt{k} \end{bmatrix}.$$

Non ci sono altre matrici diagonali che rappresentano F_k.



Caso $k = 0$.

$$\mathrm{F}_0(x) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^2)$$

Caso già considerato nell'Esempio 3. F_0 non è diagonalizzabile.

Caso $k < 0$. F_k non ha autovalori, non ha autovettori, non è diagonalizzabile.

Generale

Fino ad avviso contrario, siano V uno spazio vettoriale di dimensione n e F una applicazione lineare di V in sé. Abbiamo visto che a due autovalori distinti corrispondono autovettori indipendenti. Questo fatto vale più in generale e la sua versione generale è all'origine della teoria di autovettori ed autovalori. Di seguito enunciamo i primi teoremi e proposizioni, senza dimostrazione, e diamo una strategia in linea di principio efficace per rispondere alle principali questioni. Illustreremo la teoria su esempi in dimensione 3.

Teorema 1. Siano $a, b, \ldots, c$ autovettori di F con autovalori associati $\alpha, \beta, \ldots, \gamma$. Se $\alpha, \beta, \ldots, \gamma$ sono distinti, allora $a, b, \ldots, c$ sono linearmente indipendenti.

Commento. Essendo V di dimensione n, in V ci sono al massimo n vettori linearmente indipendenti, quindi F ha al più n autovalori distinti.

Teorema 1.1. Se F ha n autovalori distinti, allora scegliendo un autovettore per ciascun autovalore si ha una base di V, e F è diagonalizzabile.

Esempio 5

$$\mathrm{F}(x) = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^3)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} -\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} = 0, \quad \lambda(\lambda^2 - 1) = 0, \quad \lambda = 1, 0, -1.$$

Ci sono 3 autovalori distinti, tanti quanti la dimensione di $\mathbb{R}^3$, quindi F è diagonalizzabile. Comunque scelti 3 autovettori a, b, c con autovalori associati $1, 0, -1$, si ha una base di $\mathbb{R}^3$, rispetto a questa base la matrice di F è

$$[\mathrm{F}]_{a,b,c} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

Fine esempio.

Il Teorema 1 ha una forma più generale:

Teorema 2. Siano $a_1, \ldots, a_p$ e $b_1, \ldots, b_q$ e ... e $c_1, \ldots, c_r$ sequenze linearmente indipendenti di autovettori associati ad autovalori $\alpha, \beta, \ldots, \gamma$. Se $\alpha, \beta, \ldots, \gamma$ sono distinti, allora la sequenza $a_1, \ldots, a_p, b_1, \ldots, b_q, \ldots, c_1, \ldots, c_r$ è linearmente indipendente.

Commento. Indicati con $\lambda_1, \ldots, \lambda_m$ gli autovalori di F, prendendo in ciascun autospazio $V_{\lambda_1}, \ldots, V_{\lambda_m}$ una sequenza base, e concatenando le sequenze, si ottiene una sequenza linearmente indipendente, quindi

$$\sum_{i=1}^{m} \dim(V_{\lambda_i}) \leq \dim(V).$$

Teorema 2.1. Se la somma delle dimensioni degli autospazi di F è uguale alla dimensione di V, allora prendendo una sequenza base per ciascun autospazio e concatenando si ottiene una sequenza base di V, e F è diagonalizzabile. Vale il viceversa. In altri termini, indicati con $\lambda_1, \ldots, \lambda_m$ gli autovalori di F, F è diagonalizzabile se e solo se

$$\sum_{i=1}^{m} \dim(V_{\lambda_i}) = \dim(V).$$

Esempio 6

$$\mathrm{G}(x) = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^3)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} -1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} -1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} = 0, \ \ (-1-\lambda)(\lambda^2-1) = 0, \ \ \lambda = 1, -1.$$

Autospazio V_1 : spazio delle soluzioni di

$$\begin{bmatrix} -2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} x = \underline{0};$$

$$\dim(V_1) = 3 - \mathrm{rango}(\mathit{matrice}) = 3 - 2 = 1;$$

risolvendo il sistema si trova una base di un vettore a;

Autospazio V_{-1} : spazio delle soluzioni di

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} x = \underline{0};$$

$$\dim(V_{-1}) = 3 - \mathrm{rango}(\mathit{matrice}) = 3 - 1 = 2;$$

risolvendo il sistema si trova una base di due vettori b_1, b_2.

La sequenza a, b_1, b_2 è una base di $\mathbb{R}^3$, rispetto a questa base la matrice di G è

$$[\mathrm{G}]_{a,b_1,b_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

Esempio 7

$$\mathrm{H}(x) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^3)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} 1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} = 0, \ \ (1-\lambda)(\lambda^2-1) = 0, \ \ \lambda = 1, -1.$$

Autospazio V_1 : spazio delle soluzioni di

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} x = \underline{0};$$

$$\dim(V_1) = 3 - \text{rango}(matrice) = 3 - 2 = 1;$$

risolvendo il sistema si trova una base di un vettore a.

Autospazio V_{-1} : spazio delle soluzioni di

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} x = \underline{0};$$

$$\dim(V_{-1}) = 3 - \text{rango}(matrice) = 3 - 2 = 1;$$

risolvendo il sistema si trova una base di un vettore b.

Riassumendo, si ha:

$$\dim(V_1) + \dim(V_{-1}) = 1 + 1 < 3 = \dim(\mathbb{R}^3),$$

quindi H non è diagonalizzabile.

Esempio 8. E' data la famiglia di applicazioni

$$\mathrm{F}_k(x) = \begin{bmatrix} k & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x, \quad (x \in \mathbb{R}^3)$$

dove k è un parametro $\in \mathbb{R}$.

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} k-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} k-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} = 0, \;\; (k-\lambda)(\lambda^2 - 1) = 0, \;\; \lambda = k, 1, -1.$$

Se $k \neq \pm 1$, allora ci sono 3 autovalori distinti, tanti quanti la dimensione di $\mathbb{R}^3$, quindi F è diagonalizzabile. Comunque scelti 3 autovettori a, b, c con autovalori associati $k, 1, -1$, si ha una base di $\mathbb{R}^3$, rispetto a questa base la matrice di F_k è

$$[\mathrm{F}_k]_{a,b,c} = \begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

Se $k = -1$, allora F_k è diagonalizzabile (cfr. Esempio 6).

Se $k = 1$, allora F_k non è diagonalizzabile (cfr. Esempio 7).

Diagonalizzazione ortogonale, teorema spettrale

Matrici simmetriche.

Le matrici simmetriche sono le matrici quadrate del tipo

$$[a], \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}, \ldots$$

in altri termini, A quadrata $n \times n$ è diagonale se e solo se gli elementi simmetrici rispetto alla diagonale discendente sono uguali:

$$A_{ij} = A_{ji} \qquad \forall i, j,$$

in breve, se e solo se la matrice A coincide con la sua trasposta:

$$A^T = A.$$

Osservazione. Ogni matrice diagonale è simmetrica.

Diagonalizzazione ortogonale. Teorema spettrale.

Fino ad avviso contrario, sia F un'applicazione lineare dello spazio vettoriale Euclideo $\mathbb{R}^n$ in sé, data esplicitamente come $\mathrm{F}(x) = Ax$ con A matrice costante.

Diciamo che F è "ortogonalmente diagonalizzabile" se esiste una base ortogonale di $\mathbb{R}^n$ tale che la matrice di F rispetto ad essa sia diagonale, equivalentemente, se esiste una base ortogonale di $\mathbb{R}^n$ di autovettori di F. In tal caso, normalizzando i vettori della base si ha una base ortonormale $\mathcal{A} : a_1, \ldots, a_n$; dunque la matrice di questa base ha per inversa la sua trasposta (cfr. p.70-71), specificamente

$$[a_1 \cdots a_n]^{-1} = [a_1 \cdots a_n]^T.$$

La matrice di F rispetto alla base canonica è legata alla matrice diagonale rispetto alla base $\mathcal{A}$ dalla relazione

$$\begin{aligned}
[\mathrm{F}]_{\mathcal{E}\mathcal{E}} &= [\mathrm{id}]_{\mathcal{E}\mathcal{A}}[\mathrm{F}]_{\mathcal{A}\mathcal{A}}[\mathrm{id}]_{\mathcal{A}\mathcal{E}} \\
&= [a_1 \cdots a_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} [a_1 \cdots a_n]^{-1} \\
&= [a_1 \cdots a_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} [a_1 \cdots a_n]^T.
\end{aligned}$$

Possiamo sintetizzare questa relazione con

$$A = PDP^T.$$

Osserviamo che

$$A^T = (PDP^T)^T = P^{T^T} D^T P^T = PDP^T = A.$$

In sintesi:

Se $\mathrm{F} \circlearrowleft \mathbb{R}^n$, $\mathrm{F}(x) = Ax$ *è ortogonalmente diagonalizzabile, allora* A *è simmetrica.*

Vale il viceversa:

Teorema spettrale. Se F $\circlearrowleft \mathbb{R}^n$, $\mathrm{F}(x) = Ax$ con A è simmetrica, allora F è ortogonalmente diagonalizzabile.

Di seguito illustriamo questo teorema su un paio di esempi, diamo un frammento della dimostrazione, e diamo un'applicazione del teorema. In ciascuno degli esempi consideriamo un'applicazione lineare data da una matrice simmetrica e descriviamo le basi ortogonali e le basi ortonormali di autovettori.

Esempio 1.

$$\mathrm{F}(x) = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix} x, \qquad (x \in \mathbb{R}^2)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} 3-\lambda & 2 \\ 2 & -\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det \begin{bmatrix} 3-\lambda & 2 \\ 2 & -\lambda \end{bmatrix} = 0, \;\; \lambda^2 - 3\,\lambda - 4 = 0, \;\; \lambda = 4, -1.$$

Ci sono 2 autovalori distinti; 2 autovettori ad essi associati, comunque scelti, sono linearmente indipendenti, quindi sono una base di $\mathbb{R}^2$. F è diagonalizzabile. Rispetto a ciascuna base di autovettori, il 1° in V_4 e il 2° in V_{-1}, la matrice di F è

$$\begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}.$$

Il teorema spettrale afferma che esiste qualche base ortogonale, e qualche anche ortonormale, di autovettori.

Autospazio V_4 : spazio delle soluzioni di

$$\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} x = \underline{0},$$

che equivale all'unica equazione

$$-x_1 + 2\,x_2 = 0; \;\; \text{con soluzioni} \begin{bmatrix} 2\,x_2 \\ x_2 \end{bmatrix} (x_2 \text{ libera}); \text{ una retta vettoriale.}$$

Autospazio V_{-1} : spazio delle soluzioni di

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} x = \underline{0},$$

che equivale all'unica equazione

$$2\,x_1 + x_2 = 0; \;\; \text{con soluzioni} \begin{bmatrix} x_1 \\ -2\,x_1 \end{bmatrix} (x_1 \text{ libera}); \text{ una retta vettoriale.}$$

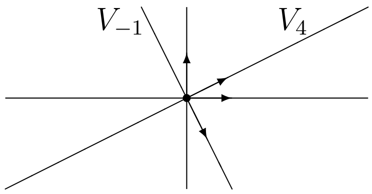
Osservazione. Ogni due vettori, uno in V_4 e uno in V_{-1}, sono fra loro ortogonali

$$\begin{bmatrix} 2x_2 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ -2x_1 \end{bmatrix} = 2x_2x_1 - x_2 2x_1 = 0.$$

Quindi: ogni base di $\mathbb{R}^2$ di autovettori è ortogonale. Per ottenere una base ortonormale, scegliamo una base ortogonale e la normalizziamo. Una delle possibili base ortonormali è

$$\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \ \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Questione per il lettore: quante sono le basi ortonormali di autovettori?



Esempio 2.

$$\mathrm{P}(x) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} x, \quad (x \in \mathbb{R}^3)$$

Equazione degli autovettori ed autovalori:

$$\begin{bmatrix} \frac{1}{3}-\lambda & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3}-\lambda & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}-\lambda \end{bmatrix} x = \underline{0};$$

equazione caratteristica degli autovalori:

$$\det(matrice) = 0, \ \ -\lambda^2(\lambda - 1) = 0, \ \ \lambda = 0, 1.$$

Autospazio V_1 : spazio delle soluzioni di

$$\begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} x = \underline{0};$$

$$\dim(V_1) = 3 - \text{rango}(matrice) = 3 - 2 = 1;$$

Autospazio V_0 : spazio delle soluzioni di

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} x = \underline{0};$$

$$\dim(V_0) = 3 - \text{rango}(matrice) = 3 - 1 = 2;$$

Si ha

$$\dim(V_1) + \dim(V_0) = 1 + 2 = 3 = \dim(\mathbb{R}^3),$$

quindi P è diagonalizzabile. Un qualsiasi autovettore in V_1 e due qualsiasi autovettori linermente indipendenti in V_0 formano una base di $\mathbb{R}^3$, rispetto a questa base la matrice di P è

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Il teorema spettrale afferma che esiste qualche base ortogonale, e qualche anche ortonormale, di autovettori.

Risolvendo le equazioni degli autospazi, si trova:

V_1 è l'insieme dei $\begin{bmatrix} x_3 \\ x_3 \\ x_3 \end{bmatrix}$ (x_3 libera); una base è $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$;

V_0 è l'insieme dei $\begin{bmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix}$ (x_2, x_3 libere); una base è $b_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Si verifica che ogni due vettori, uno in V_1 e uno in V_0 sono fra loro ortogonali. Normalizzando a, applicando la procedura di Gram-Schmidt a b_1, b_2 e normalizzando si ottiene una base ortonormale di $\mathbb{R}^3$ di autovettori di P.

Frammenti di dimostrazione.

Fatto. Ogni applicazione lineare $\mathrm{F} :\circlearrowleft \mathbb{R}^2$ con matrice canonica simmetrica

$$\mathrm{F}(x) = \begin{bmatrix} a & b \\ b & c \end{bmatrix} x \qquad (x \in \mathbb{R}^2)$$

ha due autovalori $\in \mathbb{R}$, distinti o coincidenti; e' sempre diagonalizzabile.

Proviamo l'affermazione sull'esistenza degli autovalori reali. L'equazione caratteristica degli autovalori è

$$\det \begin{bmatrix} a - \lambda & b \\ b & c - \lambda \end{bmatrix} = 0, \ \ \lambda^2 - (a + c)\lambda + ac - b^2 = 0.$$

Il discriminante del trinomio è

$$\Delta = (a + c)^2 - 4(ac - b^2) = (a - c)^2 + 4b^2 \geq 0 \text{ sempre.}$$

Teorema. Sia $\mathrm{F} :\circlearrowleft \mathbb{R}^n$, $\mathrm{F}(x) = Ax$, con A simmetrica. Ogni due autovettori a, b di F con autovalori associati α, β distinti sono fra loro ortogonali.

Dimostrazione. Moltiplicando la riga $(1 \times n)$ associata al vettore a per la matrice A per la colonna $(n \times 1)$ associata al vettore b si ottiene una matrice 1×1, un numero, che coincide col suo trasposto; tenendo conto della simmetria di A si ottiene

$$a^T A\ b = (a^T A\ b)^T = b^T A^T a^{T^T} = b^T A\ a.$$

Da una parte si ha

$$a^T (A\ b) = a^T (\beta b) = \beta (a^T b)$$

Dall'altra parte si ha

$$b^T (A\ a) = b^T (\alpha a) = \alpha (b^T a)$$

Quindi

$\beta(a^T b) = \alpha(b^T a)$

Si ha $a^T b = a \cdot b = b \cdot a = b^T a$ ed essendo $\beta \neq \alpha$ si ricava $a \cdot b = 0$, cioè $a \perp b$.

Esempio 2, ripresa.

$$\mathrm{P}(x) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} x, \quad (x \in \mathbb{R}^3)$$

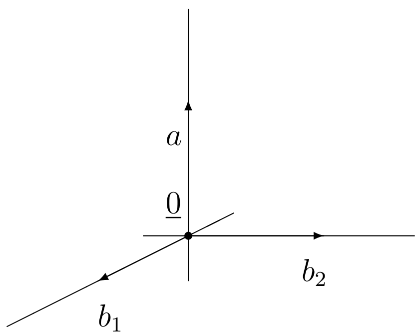
Osservazione:

P manda la base a, b_1, b_2 (con $a \perp b_1, b_2$) nella sequenza $a, \underline{0}, \underline{0}$;

Pr_a, proiezione ortogonale su a, manda la base a, b_1, b_2 nella sequenza $a, \underline{0}, \underline{0}$;

assumendo gli stessi valori su una base, le due applicazioni sono uguali:

$\mathrm{P} = \mathrm{Pr}_a$.



Usando autovettori ed autovalori abbiamo "capito" l'applicazione.

Si lascia al lettore di verificare direttamente l'uguaglianza delle due applicazioni, usando la formula per la proiezione ortogonale di un vettore su una retta vettoriale in uno spazio Euclideo (cfr. p.63 e seguenti).