

Parte 1. Sistemi lineari, algoritmo di Gauss, matrici

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Brevi richiami sugli insiemi, 1
- 2 Insiemi numerici, 3
- 3 L'insieme \mathbf{R}^n , 4
- 4 Equazioni lineari, 4
- 5 Sistemi di equazioni lineari, 8
- 6 Sistemi e matrici, 9
- 7 Sistemi lineari a scalini, 11
- 8 Algoritmo di Gauss, 15
- 9 Serie di esempi, 16
- 10 Le matrici, 19
- 11 Somma e moltiplicazione per uno scalare, 20
- 12 Matrici quadrate di tipo speciale, 22
- 13 Prodotto di matrici (righe per colonne), 24
- 14 Matrice identità, matrici invertibili, 27

1 Brevi richiami sugli insiemi

Se A è un insieme la notazione

$$x \in A$$

indica che x appartiene ad A .

Esempio Se $A = \{1, 2, 3, 4\}$ allora $2 \in A$ mentre $6 \notin A$.

Gli elementi di un dato insieme sono descritti in un'espressione fra parentesi graffe. Nell'esempio precedente, gli elementi di A sono precisamente i numeri 1, 2, 3, 4 (A è quindi un insieme *finito*, cioè costituito da un numero finito di elementi, in questo caso quattro). Diamo un altro esempio. Se \mathbf{R} indica l'insieme dei numeri reali, allora

$$A' = \{x \in \mathbf{R} : 1 \leq x \leq 2\},$$

si legge: " A' è l'insieme di tutti i numeri reali tali che $1 \leq x \leq 2$ ". Notiamo che A' è un insieme *infinito* (cioè, costituito da un numero infinito di elementi).

- Se A e B sono insiemi, diremo che A è un *sottoinsieme* di B , o che A è *contenuto in* B , se ogni elemento di A è anche elemento di B ; in tal caso scriveremo:

$$A \subseteq B.$$

Ad esempio $\{1, 3\} \subseteq \{1, 2, 3, 4\}$.

- È evidente che, se $A \subseteq B$ e $B \subseteq A$, allora $A = B$, cioè A e B hanno esattamente gli stessi elementi. Questo è il cosiddetto *Principio della doppia inclusione*, spesso usato per dimostrare l'uguaglianza di due dati insiemi.
- Se A e B sono insiemi, $A \cap B$ denota l'insieme *intersezione* di A e B , formato dagli elementi comuni ad A e B , cioè da quegli elementi che appartengono sia ad A che a B . È evidente che $A \cap B$ è contenuto sia in A che in B .
- Se A e B sono insiemi, $A \cup B$ denota l'insieme *unione* di A e B , cioè l'insieme costituito dagli elementi che appartengono ad A oppure a B (o a entrambi). È evidente che $A \cup B$ *contiene* sia A che B .

Esempio Poniamo $A = \{1, 2, 3, 4\}$, $B = \{0, 1, 2, 3\}$. Allora $A \cap B = \{1, 2, 3\}$ mentre $A \cup B = \{0, 1, 2, 3, 4\}$.

Esempio Poniamo $A' = \{a, b\}$, $B' = \{c, d\}$. Allora $A' \cap B' = \emptyset$ (insieme vuoto).

- L'insieme *vuoto* è l'insieme privo di elementi: si denota con \emptyset .

Esercizio Siano $A = \{x \in \mathbf{R} : -2 \leq x \leq 4\}$, $B = \{x \in \mathbf{R} : x \geq 0\}$. Determinare $A \cup B$ e $A \cap B$.

Esercizio a) Dimostrare che, se $A \subseteq B$, allora $A \cap B = A$.

b) Dimostrare che, se $A \cap B = A$, allora $A \subseteq B$.

Osserviamo che dimostrare sia a) che b) equivale a dimostrare:

$$A \subseteq B \text{ se e solo se } A \cap B = A.$$

Il "se e solo se" indica l'equivalenza logica delle due affermazioni.

Esercizio Dimostrare che $A \subseteq B$ se e solo se $A \cup B = B$.

2 Insiemi numerici

Il termine *numero* si riferisce sempre a un numero *reale*, a meno che non indicato altrimenti. Non daremo la definizione rigorosa di numero reale; in ogni modo ricordiamo che un numero reale si rappresenta, nella scrittura decimale, da una parte intera seguita, dopo la virgola, da un numero (eventualmente illimitato) di cifre decimali.

Come già osservato, l'insieme dei numeri reali si denota con il simbolo \mathbf{R} . Oltre a \mathbf{R} , ricordiamo i principali insiemi numerici:

- L'insieme dei numeri naturali: $\mathbf{N} = \{1, 2, 3, \dots\}$.
- L'insieme dei numeri interi: $\mathbf{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$.
- L'insieme dei numeri razionali: $\mathbf{Q} = \left\{\frac{m}{n} : m, n \in \mathbf{Z}, n \neq 0\right\}$. Possiamo sempre supporre che gli interi m ed n siano privi di fattori primi comuni (che possiamo infatti eliminare con la divisione): ad esempio $\frac{4}{6} = \frac{2}{3}$ e in quest'ultima espressione m ed n non hanno fattori comuni.

Si ha che:

$$\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q} \subseteq \mathbf{R}.$$

Notiamo che vi sono numeri reali che non sono razionali, ad esempio $\sqrt{2}, \pi$. I numeri reali, non razionali, si dicono *irrazionali*. In un certo senso (che si può rendere rigoroso) i numeri irrazionali sono *molti di piú* dei numeri razionali; in particolare, i numeri irrazionali costituiscono un insieme infinito. Possiamo facilmente dimostrare l'esistenza di almeno *un* numero irrazionale.

Proposizione $\sqrt{2}$ è un numero irrazionale.

Dimostrazione. La dimostrazione si fa per assurdo. Supponiamo che

$$\sqrt{2} = \frac{m}{n}$$

con m, n interi e $n \neq 0$. Possiamo assumere che m ed n non hanno fattori primi comuni: quindi almeno uno fra m e n è un numero dispari.

Dall'ipotesi otteniamo:

$$2n^2 = m^2.$$

Il numero m^2 è pari, e quindi anche m è pari, cioè $m = 2h$ per qualche intero h . Dunque $m^2 = 4h^2$ e

$$2n^2 = 4h^2,$$

cioè $n^2 = 2h^2$ deve essere pari. Ma allora n deve essere anch'esso pari, e questo è impossibile, poichè altrimenti m ed n sono entrambi pari, contraddicendo l'ipotesi. \square

Corollario La lunghezza della diagonale di un quadrato di lato unitario è un numero irrazionale.

3 L'insieme \mathbf{R}^n

Definiamo \mathbf{R}^n come l'insieme delle n -ple ordinate di numeri reali. Le possiamo scrivere verticalmente, come colonne:

$$\mathbf{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbf{R} \right\}.$$

Ad esempio, $\begin{pmatrix} 1 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \end{pmatrix}$ sono elementi di \mathbf{R}^2 , mentre $\begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix} \in \mathbf{R}^4$. Gli elementi di \mathbf{R}^n saranno

in seguito chiamati *vettori colonna*. Spesso sarà utile scrivere tali n -ple orizzontalmente, e diremo dunque che (x_1, \dots, x_n) è un *vettore riga* di \mathbf{R}^n .

4 Equazioni lineari

4.1 Equazioni lineari in una variabile

Sono del tipo:

$$ax = b,$$

dove x è la variabile (incognita) e $a, b \in \mathbf{R}$ sono numeri reali assegnati. Una soluzione dell'equazione è un numero reale \bar{x} che, sostituito a x , soddisfa l'equazione: $a\bar{x} = b$. Risolvere l'equazione significa trovare tutte le sue soluzioni (se esistono).

Esempio L'equazione $2x = 3$ ammette l'unica soluzione $x = \frac{3}{2}$, ottenuta dividendo ambo i membri per 2 (oppure, *moltiplicando* ambo i membri per 2^{-1}).

Esempio L'equazione $0x = 1$ non ammette soluzioni.

In generale, l'equazione $ax = b$ si discute in questo modo:

- Se $a \neq 0$ si ha un'unica soluzione: $x = a^{-1}b$.
- Se $a = 0$ e $b = 0$ ogni numero è soluzione.
- Se $a = 0$ e $b \neq 0$ non ci sono soluzioni.

4.2 Equazioni lineari in due variabili

Denotiamo con x, y le due variabili. Si ha:

$$ax + by = c$$

con $a, b, c \in \mathbf{R}$. I numeri a, b sono detti i *coefficienti*, mentre c è detto il *termine noto*. Una soluzione è una coppia ordinata di numeri reali che soddisfa l'equazione data. Scriveremo una soluzione come un elemento di \mathbf{R}^2 , diciamo $\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$: dunque si deve avere

$$a\bar{x} + b\bar{y} = c.$$

Vediamo ora come risolvere questo tipo di equazioni. Se i coefficienti a, b non sono entrambi nulli avremo sempre infinite soluzioni.

Esempio Risolviamo l'equazione lineare in due variabili seguente:

$$x + 3y = 5.$$

Soluzione. Possiamo scrivere $x = 5 - 3y$ e in corrispondenza di un valore assegnato a y abbiamo una soluzione. Infatti, se poniamo $y = t$ abbiamo:

$$x = 5 - 3t$$

e quindi l'insieme delle soluzioni è dato da $\begin{cases} x = 5 - 3t \\ y = t \end{cases}$ con t parametro reale. Osserviamo che una soluzione è un elemento di \mathbf{R}^2 , e possiamo scrivere l'insieme delle soluzioni nel modo seguente:

$$\text{Sol}(S) = \left\{ \begin{pmatrix} 5 - 3t \\ t \end{pmatrix} : t \in \mathbf{R} \right\}.$$

In particolare, abbiamo infinite soluzioni (una per ogni valore di t scelto). Ad esempio $t = 0$ dà la soluzione $\begin{pmatrix} 5 \\ 0 \end{pmatrix}$ mentre $t = 2$ dà la soluzione $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$, e così via. Diremo che in questo caso abbiamo *infinito alla uno* soluzioni (in breve ∞^1 soluzioni) perchè l'insieme delle soluzioni dipende da un parametro. È chiaro che, se c'è almeno un coefficiente diverso da zero, ogni equazione lineare in due variabili ammette ∞^1 soluzioni. \square

4.3 Equazioni lineari in numero arbitrario di variabili

Siano x_1, \dots, x_n variabili reali. Discutiamo l'equazione:

$$a_1x_1 + \dots + a_nx_n = b,$$

dove a_1, \dots, a_n sono i *coefficienti* dell'equazione e b è il *termine noto*. Una soluzione è una n -pla ordinata di numeri che, sostituiti alle variabili, soddisfano l'equazione.

Osservazione Se almeno uno dei coefficienti è diverso da zero abbiamo sempre infinite soluzioni. Più precisamente l'equazione ammette ∞^{n-1} soluzioni, nel senso che l'insieme delle soluzioni dipende da $n - 1$ parametri indipendenti. Possiamo procedere in questo modo:

- Supponiamo che $a_i \neq 0$. Assegnamo a piacere i valori delle variabili diverse da x_i (in numero, sono $n - 1$) e risolviamo rispetto a x_i .

Esempio Discutiamo l'equazione nelle variabili x, y, z :

$$S : x + y - 3z = 2$$

Soluzione. Il coefficiente di x è diverso da zero; dunque, fissando a piacere i valori delle variabili y, z e risolvendo rispetto a x si ottiene una soluzione. Ponendo $y = t_1, z = t_2$ si ha

$$x = -t_1 + 3t_2 + 2.$$

Otteniamo dunque ∞^2 soluzioni, date dalle terne del tipo:

$$\begin{cases} x = -t_1 + 3t_2 + 2 \\ y = t_1 \\ z = t_2 \end{cases}$$

al variare di t_1, t_2 in \mathbf{R} . Scriveremo dunque

$$\text{Sol}(S) = \left\{ \begin{pmatrix} -t_1 + 3t_2 + 2 \\ t_1 \\ t_2 \end{pmatrix} : t_1, t_2 \in \mathbf{R} \right\}. \quad (1)$$

Il calcolo è terminato.

Osserviamo però che potevamo procedere diversamente: anche il coefficiente della y è diverso da zero, dunque possiamo assegnare a x e z valori fissati $x = s_1, z = s_2$ e ottenere $y = -x + 3z + 2 = -s_1 + 3s_2 + 2$. In conclusione, le soluzioni saranno

$$\begin{cases} x = s_1 \\ y = -s_1 + 3s_2 + 2 \\ z = s_2 \end{cases}$$

al variare di s_1, s_2 , e le possiamo scrivere in forma vettoriale:

$$\text{Sol}(S) = \left\{ \begin{pmatrix} s_1 \\ -s_1 + 3s_2 + 2 \\ s_2 \end{pmatrix} : s_1, s_2 \in \mathbf{R} \right\}. \quad (2)$$

Sembrerebbe che, guardando le espressioni (1) e (2), abbiamo ottenuto risultati diversi: in realtà entrambe le procedure sono corrette, e gli insiemi infiniti a secondo membro delle espressioni (1) e (2) sono infatti *uguali*, sebbene presentati (o *parametrizzati*) in modo diverso. In altre parole, si può verificare che ogni terna del primo insieme appartiene anche al secondo, e viceversa. Ad esempio, scegliendo $t_1 = 1, t_2 = 2$ in (1) otteniamo la soluzione $\begin{pmatrix} 7 \\ 1 \\ 2 \end{pmatrix}$; tale terna si ottiene anche dalla (2) per i valori $s_1 = 7, s_2 = 2$. \square

Esempio Risolvere l'equazione lineare nelle incognite x_1, x_2, x_3, x_4 :

$$2x_1 + x_2 - 3x_4 = 7$$

Soluzione. Osserviamo innanzitutto che il coefficiente di x_3 è nullo: in effetti, l'equazione si può riscrivere $2x_1 + x_2 + 0x_3 - 3x_4 = 7$, mostrando che x_3 è, per così dire, una variabile "fantasma", che può assumere qualunque valore, senza alterare l'equazione.

In ogni modo, il coefficiente di x_2 è non nullo e possiamo risolvere rispetto a x_2 . Poniamo $x_1 = t_1, x_3 = t_2, x_4 = t_3$ e otteniamo quindi le soluzioni:

$$\begin{cases} x_1 = t_1 \\ x_2 = -2t_1 + 3t_3 + 7 \\ x_3 = t_2 \\ x_4 = t_3 \end{cases}$$

alla variare di t_1, t_2, t_3 : troviamo, com'era naturale, ∞^3 soluzioni.

Infine scriviamo, come al solito, l'insieme delle soluzioni in forma vettoriale (siamo in \mathbf{R}^4):

$$\text{Sol}(S) = \left\{ \begin{pmatrix} t_1 \\ -2t_1 + 3t_3 + 7 \\ t_2 \\ t_3 \end{pmatrix} \in \mathbf{R}^4 : t_1, t_2, t_3 \in \mathbf{R} \right\}.$$

\square

In conclusione, conviene ricordare che

- Un'equazione lineare in n incognite, in cui almeno uno dei coefficienti è non nullo, ammette sempre ∞^{n-1} soluzioni.

5 Sistemi di equazioni lineari

Vediamo di risolvere il seguente problema:

- trovare due numeri aventi somma 2 e differenza 3.

Detti x e y i due numeri (per ora incogniti) dobbiamo imporre le seguenti condizioni:

$$\begin{cases} x + y = 2 \\ x - y = 3 \end{cases} \quad (3)$$

La parentesi graffa indica che le due equazioni devono essere soddisfatte contemporaneamente. (3) è un esempio di *sistema lineare di due equazioni in due incognite*. Risolviamo il sistema osservando che, dalla prima equazione:

$$y = 2 - x$$

e sostituendo nella seconda otteniamo $x = \frac{5}{2}$ e quindi $y = -\frac{1}{2}$. Dunque il sistema ammette un'unica soluzione:

$$\begin{cases} x = \frac{5}{2} \\ y = -\frac{1}{2} \end{cases}$$

che possiamo esprimere come elemento di \mathbf{R}^2 , cioè: $\begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \end{pmatrix}$.

Diamo ora la seguente definizione.

Definizione Un sistema lineare di m equazioni in n incognite x_1, \dots, x_n è un'espressione del tipo:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

dove a_{11}, \dots, a_{mn} sono numeri reali, detti i coefficienti del sistema, e b_1, \dots, b_m sono numeri reali, detti i termini noti del sistema. Una soluzione del sistema è una n -pla $\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{pmatrix} \in \mathbf{R}^n$ che, sostituita ordinatamente al posto delle incognite, rende soddisfatte tutte le equazioni.

- Vedremo che ci sono tre casi possibili:

- 1) Il sistema non ammette soluzioni (è cioè *incompatibile*).
- 2) Il sistema ammette una e una sola soluzione.
- 3) Il sistema ammette infinite soluzioni.

Esempio L'esempio precedente dà un sistema lineare di due equazioni in due incognite, che ammette un'unica soluzione (caso 2).

Esempio Il sistema

$$\begin{cases} x + 3y = 1 \\ x + 3y = 0 \end{cases}$$

è incompatibile.

Esempio Il sistema

$$S : \begin{cases} x + y = 2 \\ 3x + 3y = 6 \end{cases}$$

ammette infinite soluzioni: $\text{Sol}(S) = \left\{ \begin{pmatrix} 2-t \\ t \end{pmatrix} : t \in \mathbf{R} \right\}$.

Dato un sistema lineare S , esamineremo i seguenti problemi:

- Stabilire se S ammette soluzioni.
- Calcolare esplicitamente tutte le soluzioni di S .

6 Sistemi e matrici

Una *matrice* $m \times n$ è una tabella di mn numeri reali, ordinati secondo m righe ed n colonne. Ad esempio:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix}$$

sono esempi di matrici 2×2 , mentre:

$$\begin{pmatrix} 1 & 3 & 5 & 0 \\ 1 & -1 & 0 & 2 \\ 0 & \frac{1}{2} & 4 & 5 \end{pmatrix}$$

è una matrice 3×4 . Notiamo che una matrice $n \times 1$ altro non è che un vettore colonna di \mathbf{R}^n e una matrice $1 \times n$ è un vettore riga di \mathbf{R}^n .

6.1 Matrici associate a un sistema lineare

Un sistema lineare di m equazioni in n incognite si può rappresentare mediante quella che è detta la *matrice completa* del sistema, cioè la matrice

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Notiamo che l'ultima colonna corrisponde ai termini noti del sistema. Spesso sarà importante considerare la *matrice dei coefficienti* del sistema, cioè

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

Esempio Il sistema

$$S : \begin{cases} x - 2y = 3 \\ 2x + 2y - z = 4 \\ x - 2z = 0 \end{cases}$$

ha matrice dei coefficienti

$$A = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 2 & -1 \\ 1 & 0 & -2 \end{pmatrix}$$

e matrice completa

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

Esempio La matrice $\begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 2 & 0 & 1 & -1 & 2 \end{pmatrix}$ rappresenta la matrice *completa* del sistema lineare di due equazioni in quattro incognite:

$$\begin{cases} x_1 + x_2 + x_4 = 1 \\ 2x_1 + x_3 - x_4 = 2 \end{cases}$$

6.2 Matrici a scalini

Definizione Una matrice si dice a scalini se soddisfa entrambe le seguenti proprietà:

- 1) Se una riga è nulla, allora tutte le righe ad essa sottostanti sono nulle.
- 2) Sotto il primo elemento non nullo di ciascuna riga, e sotto tutti gli zeri che lo precedono, ci sono elementi nulli.

Esempio Le seguenti matrici:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 3 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

sono a scalini, mentre la matrice

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

non è a scalini: per essere a scalini, sotto l'elemento $a_{23} = -2$ ci deve essere uno zero.

Definizione Il primo elemento non nullo di una riga è detto il pivot della data riga.

- Evidenziare i pivot negli esempi precedenti.

7 Sistemi lineari a scalini

Definizione Un sistema lineare si dice a scalini se la sua matrice completa è una matrice a scalini.

I sistemi a scalini sono facilmente risolubili.

Esempio Il sistema nelle incognite x, y, z :

$$S : \begin{cases} x + y - 2z = 1 \\ y + 3z = 0 \\ 3z = 6 \end{cases}$$

è a scalini, poiché la sua matrice completa è a scalini: $\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 3 & 6 \end{pmatrix}$. Il sistema può essere risolto facilmente partendo dall'ultima equazione, e sostituendo ordinatamente il valore trovato nelle equazioni che lo precedono. Infatti dall'ultima equazione:

$$z = 2$$

Sostituendo nella seconda:

$$y = -6$$

e infine, dalla prima equazione:

$$x = 11.$$

Dunque il sistema è compatibile, e ammette l'unica soluzione: $\begin{pmatrix} 11 \\ -6 \\ 2 \end{pmatrix}$.

Esempio Risolvere (se possibile) il sistema la cui matrice completa è

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

Soluzione. Il sistema in questione si scrive

$$\begin{cases} x_1 + 2x_2 + 3x_3 = -1 \\ x_3 = 2 \\ 0 = 2 \end{cases}$$

ed è evidentemente incompatibile; infatti l'ultima equazione è $0 = 2$, chiaramente impossibile.

Il problema è che l'ultimo pivot della matrice completa appartiene all'ultima colonna (la colonna dei termini noti). Quando questo accade, si avrà sempre un sistema incompatibile. Viceversa, se l'ultimo pivot non cade nella colonna dei termini noti allora l'ultima equazione sarà sempre risolubile, e risolvendo dal basso otterremo sempre almeno una soluzione. Dunque:

Osservazione *Un sistema lineare a scalini è compatibile se e solo se l'ultimo pivot della sua matrice completa non appartiene all'ultima colonna (cioè alla colonna dei termini noti).*

Supporremo ora che il sistema a scalini sia compatibile. Nel caso precedente il numero dei pivot è uguale al numero delle incognite, e abbiamo ottenuto una unica soluzione. Questo è sempre vero:

Osservazione Se il numero dei pivot è uguale al numero delle incognite (cioè, tutti i gradini hanno larghezza uno), il sistema ammette un'unica soluzione.

Vediamo ora come procedere negli altri casi.

Esempio Sistema a scalini 3×4 :

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_3 - x_4 = 2 \\ x_4 = 1 \end{cases}$$

Matrice completa:

$$A' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

La matrice completa è a scalini, e si vede subito che il sistema è compatibile. I pivot cadono nella prima, terza e quarta colonna. Consideriamo la variabile che non corrisponde alle colonne dei pivot: la seconda, cioè x_2 . Poniamo $x_2 = t$, parametro indipendente (libero), e isoliamolo al secondo membro delle equazioni (la prima) in cui compare. Otteniamo il sistema a scalini:

$$\begin{cases} x_1 + x_3 + x_4 = 1 - t \\ x_3 - x_4 = 2 \\ x_4 = 1 \end{cases}$$

che, per ogni t fissato, ammette un'unica soluzione, che possiamo trovare risolvendo dal basso. Otteniamo:

$$\begin{cases} x_1 = -3 - t \\ x_2 = t \\ x_3 = 3 \\ x_4 = 1 \end{cases}$$

Dunque, lasciando variare t abbiamo ∞^1 soluzioni, date da $\text{Sol}(S) = \left\{ \begin{pmatrix} -3-t \\ t \\ 3 \\ 1 \end{pmatrix} : t \in \mathbf{R} \right\}$. \square

Diamo ora il procedimento generale se il numero dei pivot è inferiore al numero delle incognite:

- 1) Consideriamo le variabili che corrispondono alle colonne in cui *non* cadono i pivot. Se r è il numero dei pivot, tali variabili saranno in numero di $n - r$.
- 2) Attribuiamo a tali variabili valori arbitrari, diciamo t_1, \dots, t_{n-r} , e portiamole ai secondi membri delle equazioni in cui compaiono.

3) Il sistema così ottenuto è a scalini e il numero dei pivot uguaglia il numero delle incognite rimaste. Dunque, per ogni valore dei parametri t_1, \dots, t_{n-r} , esso ammetterà un'unica soluzione, e lasciando variare $t_1, \dots, t_{n-r} \in \mathbf{R}$ avremo ∞^{n-r} soluzioni.

In conclusione, abbiamo il seguente risultato.

Teorema *Sia S un sistema lineare a scalini di m equazioni in n incognite, con matrice completa A' . Allora S è compatibile se e solo se l'ultimo pivot di A' non appartiene alla colonna dei termini noti.*

Se S è compatibile e se r è il numero dei pivot, allora :

- a) *S ammette un'unica soluzione se $r = n$.*
- b) *S ammette ∞^{n-r} soluzioni se $r < n$.*

Osservazione *Se il dato sistema a scalini è compatibile, allora il numero dei pivot della sua matrice completa coincide con il numero di righe non nulle, cioè con il numero delle equazioni significative del sistema.*

Esempio Risolvere il sistema a scalini

$$\begin{cases} x_1 + x_2 + 2x_3 - 3x_4 + x_5 = 0 \\ x_3 + x_4 + 2x_5 = 4 \end{cases}$$

Soluzione. La matrice completa è $\begin{pmatrix} 1 & 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 & 4 \end{pmatrix}$. Il sistema è compatibile, ci sono due

pivot e cinque incognite, dunque esso ammetterà ∞^3 soluzioni. Le variabili non corrispondenti ai pivot sono x_2, x_4, x_5 : poniamo dunque $x_2 = t_1, x_4 = t_2, x_5 = t_3$ e portiamo tali variabili a secondo membro. Otteniamo il sistema a scalini:

$$\begin{cases} x_1 + 2x_3 = -t_1 + 3t_2 - t_3 \\ x_3 = 4 - t_2 - 2t_3. \end{cases}$$

Risolvendo dal basso, otteniamo le ∞^3 soluzioni:

$$\begin{cases} x_1 = -8 - t_1 + 5t_2 + 3t_3 \\ x_2 = t_1 \\ x_3 = 4 - t_2 - 2t_3 \\ x_4 = t_2 \\ x_5 = t_3 \end{cases}$$

al variare dei parametri $t_1, t_2, t_3 \in \mathbf{R}$. \square

8 Algoritmo di Gauss

L'algoritmo di Gauss è un procedimento che, applicato ad un dato sistema S , permette di costruire un sistema a scalini \tilde{S} avente lo *stesso insieme delle soluzioni* del sistema S . Risolvendo il sistema a scalini \tilde{S} (sappiamo come fare: vedi sezione precedente) abbiamo così risolto anche il sistema di partenza.

8.1 Sistemi equivalenti e operazioni elementari

Definizione Due sistemi lineari si dicono equivalenti se hanno lo stesso insieme delle soluzioni.

Consideriamo ora le seguenti operazioni su un sistema lineare, dette *operazioni elementari* :

1. Scambiare due equazioni del sistema.
2. Moltiplicare un'equazione per uno scalare non nullo.
3. Sommare a una data equazione un multiplo di un'altra equazione del sistema.

È facile vedere che, se applichiamo una qualunque di tali operazioni ad un sistema S , otteniamo un sistema S' avente le stesse soluzioni di S : cioè, le operazioni elementari *non alterano* l'insieme delle soluzioni, e producono via via sistemi equivalenti al sistema di partenza.

8.2 Operazioni elementari sulle righe di una matrice

Sappiamo che un sistema lineare è univocamente determinato dalla sua matrice completa. Le equazioni corrispondono alle righe, dunque le operazioni elementari sui sistemi corrispondono alle seguenti operazioni, dette *operazioni elementari sulle righe* di una matrice:

- a) Scambiare due righe della matrice.
- b) Moltiplicare una riga per un numero non nullo.
- c) Sommare a una data riga un multiplo di un'altra riga della matrice.

Sommare le righe $R = (a_1, \dots, a_n)$ e $R' = (b_1, \dots, b_n)$ della matrice significa sommare le entrate corrispondenti di R e R' :

$$R + R' = (a_1 + b_1, \dots, a_n + b_n),$$

mentre moltiplicare la riga $R = (a_1, \dots, a_n)$ per il numero k significa moltiplicare ogni entrata di R per k :

$$kR = (ka_1, \dots, ka_n).$$

La notazione simbolica che useremo per le operazioni elementari a), b), c) è la seguente. Indichiamo con R_1, \dots, R_m le righe della matrice. Allora:

- a) $R_i \leftrightarrow R_j$ scambia la i -esima riga con la j -esima riga.
- b) $R_i \rightarrow kR_i$ moltiplica la i -esima riga per k .
- c) $R_i \rightarrow R_i + kR_j$ sostituisce la i -esima riga con la somma della i -esima riga e la j -esima riga moltiplicata per k .

Definizione La matrice A' si dice equivalente per righe alla matrice A se A' si ottiene da A mediante una successione di operazioni elementari sulle righe. In particolare, matrici equivalenti per righe rappresentano sistemi lineari equivalenti.

L'algoritmo di Gauss permette di costruire, a partire da una data matrice A , una matrice a scalini \tilde{A} equivalente per righe ad A . Dunque, il sistema rappresentato da \tilde{A} sarà a scalini, ed equivalente a quello rappresentato da A : risolvendo quest'ultimo (sappiamo come fare) abbiamo risolto anche il sistema di partenza. Illustriamo l'algoritmo su una serie di esempi.

9 Serie di esempi

Esempio Risolviamo il sistema:

$$S : \begin{cases} x + y + z = 12 \\ x + 2y - 2z = 1 \\ 2x + y - 3z = -5. \end{cases}$$

Soluzione. La matrice completa del sistema è $A = \begin{pmatrix} 1 & 1 & 1 & 12 \\ 1 & 2 & -2 & 1 \\ 2 & 1 & -3 & -5 \end{pmatrix}$. Utilizziamo le operazioni

elementari sulle righe per ridurre A ad una matrice a gradini. Il primo passo è quello di produrre uno zero nella posizione $(2, 1)$. Questo si può fare moltiplicando la prima riga per -1 e sommando alla seconda (nella notazione simbolica, $R_2 \rightarrow R_2 - R_1$). Otteniamo

$$A_1 = \begin{pmatrix} 1 & 1 & 1 & 12 \\ 0 & 1 & -3 & -11 \\ 2 & 1 & -3 & -5 \end{pmatrix}.$$

Per definizione, A_1 è equivalente per righe ad A .

Ora dobbiamo produrre uno zero nella posizione $(3, 1)$: ciò si può fare moltiplicando per -2 la prima riga e sommando alla terza ($R_3 \rightarrow R_3 - 2R_1$). Otteniamo

$$A_2 = \begin{pmatrix} 1 & 1 & 1 & 12 \\ 0 & 1 & -3 & -11 \\ 0 & -1 & -5 & -29 \end{pmatrix}.$$

Infine (ultimo passo) dobbiamo avere uno zero nella posizione $(3, 2)$, e quindi sommiamo alla terza riga la seconda (operazione: $R_3 \rightarrow R_3 + R_2$). Il risultato è la matrice a scalini

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 12 \\ 0 & 1 & -3 & -11 \\ 0 & 0 & -8 & -40 \end{pmatrix}.$$

\tilde{A} è la matrice ridotta cercata; il sistema rappresentato da \tilde{A} è equivalente al sistema S di partenza, dunque le soluzioni di S sono le soluzioni del sistema a scalini:

$$\begin{cases} x_1 + x_2 + x_3 = 12 \\ x_2 - 3x_3 = -11 \\ -8x_3 = -40. \end{cases}$$

Risolvendo dal basso, otteniamo l'unica soluzione $\begin{cases} x_1 = 3 \\ x_2 = 4. \\ x_3 = 5 \end{cases} \square$

Esempio Risolviamo il sistema nelle incognite x, y, z :

$$S : \begin{cases} 3y - z = 1 \\ x + y + 2z = 2 \\ x + 4y + z = 0. \end{cases}$$

Soluzione. La matrice completa è $A = \begin{pmatrix} 0 & 3 & -1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 4 & 1 & 0 \end{pmatrix}$ e dobbiamo in primo luogo ridurre A

ad una matrice a scalini. Per far partire l'algoritmo, conviene scambiare le prime due righe tra loro ($R_1 \leftrightarrow R_2$), per avere

$$A_1 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & -1 & 1 \\ 1 & 4 & 1 & 0 \end{pmatrix}.$$

Per sistemare la prima colonna, dobbiamo avere uno zero nella posizione $(3, 1)$. Dunque moltiplichiamo la prima riga per -1 e sommiamo alla terza ($R_3 \rightarrow R_3 - R_1$) ottenendo

$$A_1 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & -1 & 1 \\ 0 & 3 & -1 & -2 \end{pmatrix}.$$

Infine, sottraiamo la seconda equazione dalla terza ($R_3 \rightarrow R_3 - R_2$), e otteniamo la matrice a scalini:

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

Il sistema lineare \tilde{S} la cui matrice completa è \tilde{A} è a scalini, ed è equivalente al sistema di partenza S . Poichè l'ultimo pivot di \tilde{A} cade nella colonna dei termini noti, il sistema \tilde{S} (e quindi S) *non ammette* soluzioni (infatti, l'ultima equazione è $0 = -3$). In conclusione S è incompatibile. \square

Abbiamo illustrato l'algoritmo di Gauss solamente su esempi. In realtà esso funziona sempre, e si può usare per dimostrare la seguente

Proposizione *Con un opportuno uso delle operazioni elementari sulle righe, è sempre possibile ridurre una qualunque matrice A a una matrice a scalini \tilde{A} ad essa equivalente per righe. In particolare, ogni matrice è equivalente per righe a una matrice a scalini.*

Corollario *Ogni sistema lineare è equivalente a un sistema lineare a scalini.*

Riassumiamo i punti principali dell'algoritmo di Gauss. Per risolvere un dato sistema lineare S :

- Scriviamo la matrice completa di S .
- Mediante le operazioni elementari sulle righe, riduciamo tale matrice ad una matrice a scalini.
- Risolviamo il sistema lineare a scalini corrispondente e troviamo così anche le soluzioni di S .

Se il sistema è di piccole dimensioni conviene operare direttamente sulle equazioni, senza passare per la matrice completa.

Esempio Risolviamo il sistema:

$$\begin{cases} 2x + y - 3z = 2 \\ x + 2y + z = 1. \end{cases}$$

Moltiplichiamo la prima equazione per $-\frac{1}{2}$ e sommiamo alla seconda. Dopo aver moltiplicato la seconda equazione per 2 otteniamo il sistema a scalini:

$$\begin{cases} 2x + y - 3z = 2 \\ 3y + 5z = 0 \end{cases}$$

che ammette infinite soluzioni:

$$\text{Sol}(S) = \left\{ \begin{pmatrix} \frac{7}{3}t + 1 \\ -\frac{5}{3}t \\ t \end{pmatrix} : t \in \mathbf{R} \right\}.$$

10 Le matrici

10.1 Definizione ed esempi

Una matrice a p righe e q colonne (brevemente, di tipo $p \times q$) è una tabella di numeri reali disposti su p righe e q colonne, cioè:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{pmatrix}$$

L'elemento di posto (i, j) è l'elemento a_{ij} che compare sulla riga i -esima e la colonna j -esima della matrice. Si scriverà

$$A = (a_{ij}).$$

Esempi di matrici: $\begin{pmatrix} 1 & 2 \\ 3 & 4/3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & \sqrt{5} \\ 4 & 5 & 6 \\ -8 & 8 & 9 \end{pmatrix}$, $\begin{pmatrix} 3 & 4 & -1 \\ 6 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 7 & -2 \\ 1 & 4 & 0 & -3 \end{pmatrix}$, $\begin{pmatrix} 2 & -1 \\ 2 & 3 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 1 \\ -1 \\ 2 \end{pmatrix}$, $(1, 2, -1)$.

L'insieme delle matrici $p \times q$ (a elementi reali) sarà denotato con il simbolo $M(p, q, \mathbf{R})$.

Una matrice $p \times q$ porta con sé $p \times q$ "informazioni". Ad esempio, date n città C_1, \dots, C_n , possiamo costruire la matrice $n \times n$ avente elemento a_{ij} dato dalla distanza, in chilometri, di C_i da C_j . Per esercizio, scrivere tale matrice se le città sono Roma, Napoli e Milano.

10.2 Matrici quadrate

Una matrice si dice *quadrata* se $p = q$, ovvero, se ha lo stesso numero di righe e colonne. Negli esempi precedenti, le prime due matrici sono quadrate, le altre no.

10.3 Vettori riga e vettori colonna di una matrice

Ogni matrice di tipo $p \times q$ individua p vettori riga di \mathbf{R}^q e q vettori colonna di \mathbf{R}^p .

Esempio La matrice $\begin{pmatrix} 1 & 0 & 7 & -2 \\ 1 & 4 & 0 & -3 \end{pmatrix}$ di tipo 2×4 ha vettori riga $(1, 0, 7, -2)$, $(1, 4, 0, -3) \in \mathbf{R}^4$ e vettori colonna $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -2 \\ -3 \end{pmatrix} \in \mathbf{R}^2$.

10.4 Matrice trasposta

Sia data una matrice A di tipo $p \times q$. La matrice *trasposta*, denotata con A^t , è la matrice che si ottiene scambiando le righe e le colonne di A . La trasposta è di tipo $q \times p$.

Esempio Se $A = \begin{pmatrix} 1 & 0 & 7 & -2 \\ 1 & 4 & 0 & -3 \end{pmatrix}$ allora $A^t = \begin{pmatrix} 1 & 1 \\ 0 & 4 \\ 7 & 0 \\ -2 & -3 \end{pmatrix}$.

È chiaro che, se $A = (a_{ij})$, allora $A^t = (a_{ji})$; è anche ovvio che prendendo la trasposta due volte riotteniamo la matrice di partenza: $(A^t)^t = A$.

11 Somma e moltiplicazione per uno scalare

Introdurremo le operazioni naturali di somma di due matrici e moltiplicazione di una matrice per uno scalare, e ne osserveremo le principali proprietà.

11.1 Somma di due matrici

Date due matrici $A = (a_{ij})$ e $B = (b_{ij})$, entrambe di tipo $p \times q$, definiamo la *somma* $A + B$ come la matrice di entrate c_{ij} date da $c_{ij} = a_{ij} + b_{ij}$. Anche la matrice somma è di tipo $p \times q$.

Quindi, sommare due matrici significa semplicemente sommare gli elementi corrispondenti.

Esempio Se $A = \begin{pmatrix} 1 & 2 & 0 \\ -5 & -2 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} 3 & 0 & 1 \\ -1 & 2 & 0 \end{pmatrix}$ allora

$$A + B = \begin{pmatrix} 4 & 2 & 1 \\ -6 & 0 & 1 \end{pmatrix}.$$

Esempio Se $A = \begin{pmatrix} 1 & 2 \\ -3 & 6 \end{pmatrix}$ e $B = \begin{pmatrix} -1 & -2 \\ 3 & -6 \end{pmatrix}$ allora

$$A + B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = O.$$

Notiamo che $B = -A$, la matrice opposta di A . Dunque abbiamo $A + (-A) = O$. Questo risulta vero per ogni matrice.

11.2 Proprietà della somma di matrici

In ciò che segue, A, B, C, \dots sono matrici di tipo $p \times q$.

- 1. Proprietà associativa:** $(A + B) + C = A + (B + C)$ per ogni $A, B, C \in M(p, q, \mathbf{R})$.
- 2. Proprietà commutativa:** $A + B = B + A$ per ogni $A, B \in M(p, q, \mathbf{R})$.
- 3. Esistenza dello zero:** Esiste una matrice (la matrice nulla) denotata con O , che soddisfa la proprietà: $A + O = A$ per ogni matrice $A \in M(p, q, \mathbf{R})$.
- 4. Esistenza dell'opposto:** Data $A \in M(p, q, \mathbf{R})$ esiste una matrice, denotata con $-A$ e chiamata *matrice opposta di A*, avente la proprietà $A + (-A) = O$.

È facile dimostrare le proprietà 1-2, che sono conseguenza delle analoghe proprietà della somma di numeri reali. La proprietà associativa afferma in particolare che possiamo sommare tre matrici date in qualsiasi ordine, e il risultato si potrà scrivere $A + B + C$, senza ambiguità.

La *matrice nulla* di cui alla proprietà **3** è la matrice avente tutti gli elementi (entrate) nulli. La *matrice opposta* di una data matrice $A = (a_{ij})$, è la matrice $-A = (-a_{ij})$ ottenuta cambiando di segno a tutti gli elementi di A . Esempio:

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

Esempio Si vede facilmente che $(A + B)^t = A^t + B^t$.

Esercizio Date $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$, determinare una matrice X di tipo 2×2 tale che $A + X = B$.

Soluzione. Si vede subito che, sommando ad ambo i membri la matrice $-A$ otteniamo

$$X = -A + B = \begin{pmatrix} -3 & -2 \\ -3 & 3 \end{pmatrix}.$$

Esempio Vale la seguente legge di cancellazione: $A + C = B + C$ implica $A = B$.

11.3 Moltiplicazione di una matrice per uno scalare

Data una matrice $A = (a_{ij})$ e un numero reale k , definiamo la matrice kA come la matrice con entrate (ka_{ij}) . Semplicemente, kA si ottiene moltiplicando ciascuna entrata di A per il numero k .

Esempio Se $A = \begin{pmatrix} 3 & 1 & 0 \\ -2 & -1 & 4 \end{pmatrix}$ e $k = 3$, allora $kA = \begin{pmatrix} 9 & 3 & 0 \\ -6 & -3 & 12 \end{pmatrix}$.

Valgono le seguenti proprietà (analoghe alle proprietà distributive dei numeri reali), che si dimostrano facilmente:

1. $h(kA) = (hk)A$ per ogni $A \in M(p, q, \mathbf{R})$ e $h, k \in \mathbf{R}$.
2. $(h + k)A = hA + kA$ per ogni $A \in M(p, q, \mathbf{R})$ e $h, k \in \mathbf{R}$.
3. $h(A + B) = hA + hB$ per ogni $A, B \in M(p, q, \mathbf{R})$ e $h \in \mathbf{R}$.

Si verifica immediatamente che $1A = A$, $(-1)A = -A$ (la matrice opposta), e $0A = O$.

11.4 Combinazioni lineari di matrici

Con le due operazioni appena introdotte, possiamo formare le *combinazioni lineari* di due o più matrici: se A, B sono matrici, e h, k sono numeri reali, possiamo considerare la matrice $hA + kB$ (ovviamente dello stesso tipo). Più in generale, possiamo formare combinazioni lineari di un numero qualunque di matrici.

Esempio Se $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$ allora $2A - 4B = \begin{pmatrix} -6 & 4 \\ 6 & -6 \end{pmatrix}$.

Esercizio Date $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$, trovare l'unica matrice X di tipo 2×2 tale che $2A + 3X = 4B$.

Esempio Ogni matrice 2×2 è una combinazione lineare delle matrici $E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Infatti:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aE_1 + bE_2 + cE_3 + dE_4.$$

12 Matrici quadrate di tipo speciale

Ricordiamo che una matrice si dice *quadrata di ordine n* se è di tipo $n \times n$.

12.1 Matrici simmetriche e antisimmetriche

Una matrice si dice *simmetrica* se $A^t = A$. Altrimenti detto, una matrice è simmetrica se $a_{ij} = a_{ji}$ per ogni i, j . L'insieme delle matrici simmetriche di ordine n si denota con $S(n)$.

Esempio La matrice $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 5 \\ 3 & 5 & -4 \end{pmatrix}$ è simmetrica, mentre la matrice $\begin{pmatrix} 1 & 2 \\ 8 & 0 \end{pmatrix}$ non lo è.

Una matrice quadrata si dice *antisimmetrica* se $A^t = -A$, ovvero $a_{ji} = -a_{ij}$ per ogni i, j . L'insieme delle matrici antisimmetriche di ordine n si denota con $A(n)$.

Esempio La matrice $\begin{pmatrix} 0 & 1 & -3 \\ -1 & 0 & 4 \\ 3 & -4 & 0 \end{pmatrix}$ è antisimmetrica.

Esercizio a) Scrivere la generica matrice simmetrica 2×2 .

b) Scrivere la generica matrice simmetrica 3×3 .

Esercizio a) Scrivere la generica matrice antisimmetrica di ordine 2.

a) Scrivere la generica matrice antisimmetrica di ordine 3.

Esercizio Dimostrare che, data la matrice quadrata A , la matrice $A + A^t$ è simmetrica, e la matrice $A - A^t$ è antisimmetrica.

Esercizio Dimostrare che ogni matrice quadrata si scrive (in modo unico) come somma di una matrice simmetrica e di una matrice antisimmetrica.

Esercizio Dimostrare che l'unica matrice simmetrica e antisimmetrica è la matrice nulla.

12.2 Matrici triangolari

Le matrici *triangolari superiori* sono quelle che hanno tutti zeri sotto la diagonale principale:

$a_{ij} = 0$ per ogni $i > j$. Ad esempio $\begin{pmatrix} 1 & 0 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 5 \end{pmatrix}$.

Le matrici *triangolari inferiori* sono quelle che hanno tutti zeri sopra la diagonale principale:

$a_{ij} = 0$ per ogni $i < j$. Ad esempio $\begin{pmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ 5 & 6 & 9 \end{pmatrix}$.

12.3 Matrici diagonali

Una matrice si dice *diagonale* se tutti gli elementi non appartenenti alla diagonale principale sono nulli. Dunque una matrice è diagonale se e solo se è al tempo stesso triangolare superiore e triangolare inferiore.

Esempio Matrice diagonale generica di ordine 3: $\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$.

13 Prodotto di matrici (righe per colonne)

Definiamo ora un'operazione che associa a due matrici A, B di tipo opportuno una terza matrice, chiamata *prodotto righe per colonne* di A e B . Iniziamo dal caso in cui A è un vettore riga $1 \times n$ e B è un vettore colonna $n \times 1$.

13.1 Prodotto di un vettore riga per un vettore colonna

È il numero (matrice 1×1) definito da:

$$(x_1, x_2, \dots, x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Vale a dire, si moltiplicano a due a due gli elementi corrispondenti e si somma.

Esempio $(2, 1, -1) \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} = 6 - 1 - 4 = 1.$

13.2 Definizione del prodotto di matrici

Sia A una matrice $m \times n$ e B una matrice $p \times q$. Se $n = p$, cioè se il numero delle colonne di A è uguale al numero delle righe di B , allora *definiamo* la matrice prodotto $A \cdot B$ (in quest'ordine!) con la seguente regola:

- *L'elemento di posto (i, j) della matrice $A \cdot B$ è il prodotto della i -esima riga di A per la j -esima colonna di B .*

Si noti che il risultato è una matrice $m \times q$.

In conclusione, se $A = (a_{ij})$ è $m \times n$ e $B = (b_{ij})$ è $n \times q$ allora $AB = (c_{ij})$ è per definizione la matrice $m \times q$, tale che:

$$c_{ij} = a_{i1}b_{1j} + \dots + a_{in}b_{nj}.$$

D'ora in poi scriveremo il prodotto senza puntino: AB .

Esempio Siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 2 \end{pmatrix}$. Poiché A è 2×2 e B è 2×3 , possiamo calcolare AB , che sarà una matrice 2×3 . Notiamo che BA non esiste. Si ottiene:

$$AB = \begin{pmatrix} 1 & 3 & 7 \\ 3 & 5 & 17 \end{pmatrix}.$$

Infatti, detti c_{ij} gli elementi di AB , abbiamo dalla definizione:

$$\begin{aligned} c_{11} &= (1, 2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \\ c_{12} &= (1, 2) \begin{pmatrix} -1 \\ 2 \end{pmatrix} = 3 \\ &\dots \\ c_{23} &= (3, 4) \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 17 \end{aligned}$$

Esercizio Siano A una matrice 2×3 e B una matrice 3×3 . Stabilire quali dei seguenti prodotti esiste: AB, BA, AA, BB .

13.3 Il prodotto di matrici non è commutativo

- Se A e B sono quadrate di ordine n allora AB e BA esistono entrambe, e sono anch'esse quadrate di ordine n .

Esempio Siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 2 \\ -3 & 5 \end{pmatrix}$. Allora:

$$AB = \begin{pmatrix} -7 & 12 \\ -15 & 26 \end{pmatrix}, \quad BA = \begin{pmatrix} 5 & 6 \\ 12 & 14 \end{pmatrix}.$$

Notiamo subito che $AB \neq BA$. Cioè:

- Il prodotto di matrici *non* soddisfa la proprietà commutativa.

”Non soddisfa” significa che *non sempre* si ha $AB = BA$. Ma a volte questo è possibile: se O indica la matrice nulla, allora

$$AO = OA = O$$

per ogni A .

- Diremo che le matrici A e B *commutano* se $AB = BA$.

Esercizio Sia $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Descrivere l'insieme delle matrici 2×2 che commutano con A , cioè descrivere la matrice generica B tale che $AB = BA$.

Soluzione. Partiamo dalla matrice generica 2×2 , che possiamo scrivere $B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$. Imponiamo ora che $AB = BA$; otterremo un sistema di equazioni che, risolto, mostra che B è del tipo $B = \begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$, con $x, y \in \mathbf{R}$. Dunque, esistono infinite matrici che commutano con A , ma non tutte le matrici 2×2 hanno questa proprietà.

Esercizio Sia ora $A' = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$. Descrivere l'insieme delle matrici 2×2 che commutano con A' .

Soluzione. $B = \begin{pmatrix} x & 2z \\ z & x+z \end{pmatrix}$ con $x, z \in \mathbf{R}$.

13.4 Proprietà del prodotto di matrici

Il prodotto di numeri gode delle proprietà associativa, commutativa e distributiva (rispetto alla somma). Cioè, se $a, b, c \in \mathbf{R}$ abbiamo sempre:

- 1) Proprietà *associativa*: $(ab)c = a(bc)$.
- 2) Proprietà *commutativa*: $ab = ba$.
- 3) Proprietà *distributiva*: $(a+b)c = ab + ac$.

Quali di queste proprietà si estendono al prodotto di matrici? Abbiamo già visto che la proprietà commutativa non vale in generale. Nessun problema, però, con le altre proprietà. Vale infatti la seguente:

Proposizione Per ogni terna di matrici A, B, C per le quali i prodotti indicati esistono, si ha:

1. Proprietà *associativa*: $(AB)C = A(BC)$.
2. Proprietà *distributive rispetto alla somma*: $A(B+C) = AB + AC$, $(A+B)C = AC + BC$. Inoltre il prodotto si comporta in modo naturale rispetto alla moltiplicazione per uno scalare:

3. Date due matrici A e B e un numero reale h si ha sempre:

$$h(AB) = (hA)B = A(hB).$$

La proposizione si dimostra con una verifica diretta, che omettiamo.

In conclusione, abbiamo introdotto le operazioni di somma e di prodotto di due matrici. Tutte le proprietà algebriche naturali, valide per la somma e il prodotto di numeri, si estendono

alle matrici, con la sola eccezione della proprietà commutativa, che non vale in generale per il prodotto di matrici.

Esercizio Siano $A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & -3 \end{pmatrix}$. Calcolare esplicitamente i prodotti: $AB, (AB)C, BC, A(BC)$ e verificare che effettivamente si ha $(AB)C = A(BC)$.

14 Matrice identità, matrici invertibili

14.1 Matrice identità

Il numero reale 1 ha la seguente proprietà di *neutralità* rispetto al prodotto: per ogni $a \in \mathbf{R}$ si ha

$$a \cdot 1 = a.$$

Si dice anche che 1 è *l'elemento neutro* rispetto al prodotto (di numeri).

Esiste un elemento neutro rispetto al prodotto di matrici? La risposta è affermativa.

- Definiamo *matrice identità* di ordine n la matrice quadrata $n \times n$, denotata con I_n , che ha tutti 1 sulla diagonale principale e zero altrove. Quindi $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ e in generale

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

- Per ogni matrice A di tipo $m \times n$ si ha allora:

$$AI_n = A, I_m A = A.$$

La verifica si fa con un calcolo diretto.

Spesso ometteremo di indicare l'ordine, e scriveremo semplicemente I . Si ha dunque $AI = IA = A$ per ogni matrice A .

14.2 Matrici invertibili

Se a è un numero reale *non nullo* allora esiste il suo inverso a^{-1} , che per definizione ha la proprietà:

$$aa^{-1} = 1.$$

Consideriamo ora la nozione analoga per le matrici quadrate.

- Diremo che la matrice quadrata A di ordine n è *invertibile* se esiste una seconda matrice B di ordine n tale che:

$$AB = BA = I.$$

Quando esiste, tale matrice B si denota con A^{-1} (dimostriamo che è unica). La matrice A^{-1} è detta *l'inversa* di A .

A differenza del caso dei numeri, *non è vero* che ogni matrice non nulla è invertibile. Ad esempio, la matrice $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ è non nulla, eppure non esiste alcuna matrice $X \in M(2, 2, \mathbf{R})$ tale che $AX = I$.

Infatti, poniamo $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ e tentiamo di risolvere l'equazione matriciale $AX = I$ ovvero:

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Tale equazione si traduce in un sistema lineare:

$$\begin{cases} x + 2z = 1 \\ y + 2w = 0 \\ 2x + 4z = 0 \\ 2y + 4w = 1 \end{cases}$$

che *non* ammette soluzioni: la prima e terza equazione sono chiaramente incompatibili.

Esercizio Dimostrare che la matrice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ è invertibile, e determinare la sua inversa.

Esercizio Si consideri la generica matrice diagonale di ordine 2: $D = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$. Dimostrare che D è invertibile se e solo se a e d sono *entrambi* diversi da zero. In tal caso l'inversa è $D^{-1} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/d \end{pmatrix}$. Generalizzare tale risultato alla generica matrice diagonale di ordine n .

Parte 2. Determinante e matrice inversa

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Determinante di una matrice 2×2 , 1
- 2 Teorema di Cramer (caso particolare), 3
- 3 Determinante di una matrice $n \times n$, 5
- 4 Teorema di Laplace, 7
- 5 Prime proprietà del determinante, 9
- 6 Matrice inversa, 11
- 7 Teorema di Cramer (caso generale), 15
- 8 Determinante e algoritmo di Gauss, 17

1 Determinante di una matrice 2×2

Abbiamo già osservato nella Parte 1.14 che ci sono matrici non invertibili, ad esempio $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Occorre quindi un criterio che stabilisca quando una matrice è invertibile, e quando non lo è. Tale criterio si esprime tramite la nozione di *determinante*. In ciò che segue, esamineremo il caso in cui la matrice ha ordine 2; in seguito generalizzeremo i risultati a matrici di ordine arbitrario.

Definizione Data la matrice quadrata $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ di ordine 2, si definisce determinante di A il numero reale:

$$\det A = ad - bc.$$

Il determinante di A si indica anche con il simbolo $|A|$.

Proposizione Per ogni matrice 2×2 si ha $\det A = \det(A^t)$.

Dimostrazione. Verifica immediata. \square

Osserviamo anche la seguente identità, nota come *Formula di Binet*.

Proposizione Date due matrici $A, B \in M(2, 2, \mathbf{R})$ si ha sempre $\det(AB) = \det A \cdot \det B$.

Dimostrazione. È una verifica diretta: siano $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$. Sviluppando separatamente $\det(AB)$ e $\det A \cdot \det B$ si osserva l'uguaglianza per ogni scelta di $a, b, c, d, a', b', c', d'$. \square

1.1 Un criterio per l'invertibilità

La nozione di determinante fornisce un utile criterio per vedere facilmente se una matrice è invertibile oppure no.

Teorema La matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è invertibile se e solo se $\det A \neq 0$. In tal caso l'inversa è data dalla matrice:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Dimostrazione. Il "se e solo se" significa che dobbiamo dimostrare l'equivalenza delle due affermazioni. Quindi dobbiamo dimostrare due cose:

- a) Se A è invertibile allora $\det A \neq 0$.
- b) Se $\det A \neq 0$ allora A è invertibile.

Riguardo alla implicazione a), abbiamo per ipotesi che A è invertibile e dobbiamo dimostrare che $\det A \neq 0$. Per ipotesi esiste una matrice B tale che $AB = I$. Applichiamo la formula di Binet: poiché $\det I = 1$, otteniamo:

$$1 = \det I = \det(AB) = \det A \cdot \det B.$$

Ora $\det A$ e $\det B$ sono due numeri reali che hanno prodotto uguale a 1: necessariamente ciascuno di essi è diverso da zero, in particolare $\det A \neq 0$.

Dimostriamo ora l'implicazione b): per ipotesi $\det A \neq 0$ e dobbiamo dimostrare che esiste una matrice B tale che $AB = BA = I$. Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e poniamo:

$$B = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Per le proprietà del prodotto di matrici, abbiamo:

$$AB = \frac{1}{\det A} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = I.$$

Dunque $AB = I$; in modo analogo si verifica che $BA = I$ e di conseguenza A è invertibile, con inversa $A^{-1} = B$. \square

Esempio La matrice $\begin{pmatrix} 2 & 3 \\ -4 & -6 \end{pmatrix}$ non è invertibile poichè ha determinante nullo.

Esempio La matrice $\begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix}$ ha determinante 2: dunque è invertibile. L'inversa è:

$$A^{-1} = \frac{1}{2} \begin{pmatrix} 5 & -3 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Si verifica subito che $A^{-1}A = I = AA^{-1}$.

Esercizio Stabilire se ciascuna delle seguenti matrici è invertibile oppure no. Se è invertibile, calcolare la matrice inversa.

$$A_1 = \begin{pmatrix} -2 & 1 \\ 2 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}, A_3 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

Esercizio Sia $A = \begin{pmatrix} 2 & k \\ 3 & 5 \end{pmatrix}$ dove k è un parametro reale. Per quali valori di k la matrice A è invertibile? In corrispondenza di tali valori, calcolare esplicitamente A^{-1} .

1.2 Legge di cancellazione per il prodotto di matrici

Per i numeri reali, vale la ben nota *legge di cancellazione* :

- Se $ab = ac$, e $a \neq 0$, allora $b = c$.

L'analogo per le matrici non vale, in generale; dobbiamo sostituire la condizione $A \neq 0$ con la condizione $\det A \neq 0$. Precisamente:

Proposizione Siano A, B, C matrici 2×2 . Se $AB = AC$ e $\det A \neq 0$ allora $B = C$.

Dimostrazione. Per ipotesi $\det A \neq 0$: dunque A ammette matrice inversa A^{-1} . Moltiplicando ambo i membri dell'identità $AB = AC$ (a sinistra) per A^{-1} otteniamo l'asserto. \square .

2 Teorema di Cramer (caso particolare)

In questa sezione applicheremo il determinante allo studio di sistemi lineari di due equazioni in due incognite.

2.1 Forma matriciale di un sistema

Sia S un sistema lineare di due equazioni in due incognite:

$$S : \begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}.$$

Detta $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ la matrice dei coefficienti, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ il vettore colonna delle incognite e $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ il vettore colonna dei termini noti, il sistema si scrive semplicemente:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

cioè a dire:

$$AX = B,$$

dove AX indica il prodotto righe per colonne delle matrici A e X .

- L'equazione matriciale $AX = B$ è la cosiddetta *forma matriciale* del sistema S .

La forma matriciale è un'equazione lineare in cui l'incognita è una matrice (in questo caso il vettore colonna X) e i coefficienti sono anch'essi matrici. Poiché A è invertibile, essa ammette matrice inversa A^{-1} : allora moltiplicando ambo i membri *a sinistra* per A^{-1} otteniamo:

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

Dunque la soluzione è unica, e si può calcolare facilmente.

Esempio Il sistema $\begin{cases} x_1 + 3x_2 = 2 \\ x_1 + 5x_2 = -1 \end{cases}$ ha matrice dei coefficienti $A = \begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix}$; poichè $\det A = 2$

il sistema ammette l'unica soluzione $X = A^{-1}B$. Ora $A^{-1} = \frac{1}{2} \begin{pmatrix} 5 & -3 \\ -1 & 1 \end{pmatrix}$ e l'unica soluzione è data da:

$$X = A^{-1}B = \frac{1}{2} \begin{pmatrix} 5 & -3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{13}{2} \\ -\frac{3}{2} \end{pmatrix},$$

cioè $x_1 = \frac{13}{2}, x_2 = -\frac{3}{2}$. \square

Esercizio Risolvere il sistema con l'algoritmo di Gauss e verificare il risultato.

2.2 Il teorema di Cramer per sistemi 2×2

In conclusione, abbiamo dimostrato il seguente risultato (un caso particolare di quello che sarà chiamato *Teorema di Cramer*).

Proposizione *Sia S un sistema lineare di due equazioni in due incognite avente forma matriciale $AX = B$, dove A è la matrice dei coefficienti, X è il vettore colonna delle incognite e B è il vettore colonna dei termini noti. Se $\det A \neq 0$ il sistema è risolubile e ammette l'unica soluzione $X = A^{-1}B$.*

Al momento, nulla possiamo dire a priori nel caso in cui la matrice dei coefficienti abbia determinante nullo. Di questo ci occuperemo nelle prossime lezioni.

Esempio I sistemi lineari $S_1 : \begin{cases} 2x_1 + x_2 = 4 \\ 4x_1 + 2x_2 = 8 \end{cases}$ e $S_2 : \begin{cases} 2x_1 + x_2 = 1 \\ 4x_1 + 2x_2 = 0 \end{cases}$ hanno matrice dei coefficienti a determinante nullo: non possiamo applicare il teorema precedente. Discutere le (eventuali) soluzioni usando il metodo di Gauss.

3 Determinante di una matrice $n \times n$

Abbiamo definito il determinante di una matrice 2×2 : vogliamo ora definire il determinante di una matrice quadrata di ordine arbitrario. La definizione è di tipo *induttivo*, nel senso che il determinante di una matrice di ordine n si definisce combinando opportuni determinanti di ordine $n - 1$. Dopo $n - 2$ passi, il tutto si ridurrà al calcolo di un certo numero di determinanti di ordine 2, che abbiamo già definito.

3.1 Matrice aggiunta di un elemento

Definizione *Sia A una matrice $n \times n$ e a_{ij} un elemento di A . La matrice aggiunta di a_{ij} , denotata con A_{ij} , è la matrice di ordine $n - 1$ che si ottiene sopprimendo la i -esima riga e j -esima colonna di A .*

Esempio Sia $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$. Allora le matrici aggiunte dei vari elementi saranno:

$$\begin{aligned} A_{11} &= \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, A_{12} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}, A_{13} = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}, \\ A_{21} &= \begin{pmatrix} -1 & 4 \\ 1 & 3 \end{pmatrix}, A_{22} = \begin{pmatrix} 2 & 4 \\ 2 & 3 \end{pmatrix}, A_{23} = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}, \\ A_{31} &= \begin{pmatrix} -1 & 4 \\ 3 & 1 \end{pmatrix}, A_{32} = \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix}, A_{33} = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}. \end{aligned}$$

3.2 Determinante di una matrice 3×3

Definizione Sia A una matrice 3×3 , e sia A_{ij} la matrice aggiunta dell'elemento a_{ij} . Allora definiamo

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

Osserviamo che sappiamo calcolare i tre determinanti di ordine due che compaiono nella definizione. Si tratta quindi di moltiplicare ciascun elemento della prima riga per il determinante della sua matrice aggiunta, e poi sommare, procedendo però a *segni alterni*. La procedura è anche detta *sviluppo del determinante lungo la prima riga*.

Esempio Calcoliamo il determinante della matrice $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$ dell'esempio precedente.

Si ha:

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} \\ &= 2 \det \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} - (-1) \det \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} + 4 \det \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \\ &= 2 \cdot 8 + (-2) + 4(-6) \\ &= -10 \end{aligned}$$

quindi $\det A = -10$.

Esempio Il determinante della matrice identità è uguale a 1.

3.3 Determinante di una matrice di ordine n

Siamo ora pronti per la definizione generale.

Definizione Sia A una matrice $n \times n$, e sia A_{ij} la matrice aggiunta dell'elemento a_{ij} . Allora definiamo

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} - \cdots + (-1)^{n+1} a_{1n} \det A_{1n}.$$

Usando il simbolo di sommatoria:

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

Notiamo i segni alterni, dovuti al termine $(-1)^{j+1}$.

Esempio Calcoliamo il determinante della matrice $A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 2 & -1 & 4 \\ 4 & 0 & 1 & 0 \\ 1 & -3 & 2 & 0 \end{pmatrix}$. Dalla definizione:

$$\det A = 1 \cdot \det \begin{pmatrix} 2 & -1 & 4 \\ 0 & 1 & 0 \\ -3 & 2 & 0 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} 1 & -1 & 4 \\ 4 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & 2 & 4 \\ 4 & 0 & 0 \\ 1 & -3 & 0 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 1 \\ 1 & -3 & 2 \end{pmatrix}$$

Procedendo al calcolo dei determinanti di ordine tre (è ovviamente sufficiente calcolare solo quelli che non vengono moltiplicati per 0), otteniamo $\det A = -84$.

4 Teorema di Laplace

Una proprietà notevole del determinante è data dal fatto che esso può essere sviluppato lungo una qualsiasi riga o colonna, nel senso del seguente teorema, noto come *primo teorema di Laplace*.

Teorema Sia A una matrice $n \times n$. Allora:

a) Per ogni $k = 1, \dots, n$ si ha:

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

detto sviluppo del determinante lungo la k -esima riga.

b) Per ogni $h = 1, \dots, n$ si ha:

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

detto sviluppo del determinante lungo la h -esima colonna.

Dimostrazione. Omessa. \square

In pratica il teorema afferma che, per calcolare il determinante, possiamo:

- 1) Scegliere una qualunque riga (o colonna).
- 2) Moltiplicare ciascun elemento della riga (o colonna) scelta per il determinante della sua matrice aggiunta, con il segno $+$ o $-$ a seconda che la somma degli indici di riga e colonna del dato elemento sia rispettivamente pari o dispari.
- 3) Sommare tutti i termini così ottenuti.

Verifichiamo il teorema di Laplace sul seguente esempio.

Esempio Torniamo alla matrice $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$, di cui avevamo già calcolato il determinante, trovando il valore -10 . Sviluppiamo il determinante lungo la seconda riga, e otteniamo dalla proposizione:

$$\begin{aligned} \det A &= -a_{21} \det A_{21} + a_{22} \det A_{22} - a_{23} \det A_{23} \\ &= -0 \cdot \det \begin{pmatrix} -1 & 4 \\ 1 & 3 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 2 & 4 \\ 2 & 3 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix} \\ &= 3 \cdot (-2) - 1 \cdot (4) \\ &= -10 \end{aligned}$$

Ora sviluppiamo il determinante lungo l'ultima colonna, e otteniamo:

$$\begin{aligned} \det A &= a_{13} \det A_{13} - a_{23} \det A_{23} + a_{33} \det A_{33} \\ &= 4 \cdot \det \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} \\ &= 4 \cdot (-6) - 1 \cdot 4 + 3 \cdot 6 \\ &= -10 \end{aligned}$$

Notiamo lo *schema dei segni* : $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$.

Per abbreviare i calcoli, possiamo dunque scegliere la riga (o colonna) con il massimo numero di zeri.

Esempio Calcoliamo il determinante della matrice $A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 2 & -1 & 4 \\ 4 & 0 & 1 & 0 \\ 1 & -3 & 2 & 0 \end{pmatrix}$ già considerata in precedenza. Conviene sviluppare il determinante lungo l'ultima colonna; dalla proposizione abbiamo che

$$\begin{aligned} \det A &= a_{24} \det A_{24} \\ &= 4 \cdot \det \begin{pmatrix} 1 & 0 & 2 \\ 4 & 0 & 1 \\ 1 & -3 & 2 \end{pmatrix} \\ &= 4 \cdot 3 \cdot (-7) \\ &= -84 \end{aligned}$$

e il determinante di ordine tre è stato sviluppato lungo la seconda colonna.

5 Prime proprietà del determinante

Osserviamo innanzitutto:

Proposizione *Si ha sempre $\det A = \det(A^t)$.*

Dimostrazione. Per matrici di ordine due l'affermazione è immediata. Supponiamo ora che la matrice abbia ordine 3, e calcoliamo il determinante con lo sviluppo lungo la prima riga. Osserviamo che tale sviluppo coincide con lo sviluppo del determinante di A^t lungo la prima colonna (spiegare perché) e dunque, per il teorema di Laplace, $\det A = \det(A^t)$. Ora è chiaro come continuare se l'ordine è 4, 5 ... \square

Se una matrice è triangolare superiore, o triangolare inferiore, il determinante si calcola immediatamente.

Proposizione *Il determinante di una matrice triangolare superiore, o triangolare inferiore, uguaglia il prodotto degli elementi diagonali.*

Dimostrazione. Supponiamo che la matrice sia triangolare superiore. Se la matrice è di ordine due l'affermazione è immediata: $\det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = ad$. Se è di ordine tre, essa si scrive

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Sviluppando lungo la prima colonna:

$$\det A = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ 0 & a_{33} \end{pmatrix} = a_{11} a_{22} a_{33}.$$

È chiaro a questo punto che, sviluppando il determinante via via lungo la prima colonna, la proposizione risulta vera per ogni ordine. Un argomento simile prova l'affermazione nel caso in cui la matrice sia triangolare inferiore (sviluppare via via lungo la prima riga). \square

Esempio Abbiamo $\det \begin{pmatrix} 1 & 5 & -9 \\ 0 & 3 & 7 \\ 0 & 0 & -2 \end{pmatrix} = 1 \cdot 3 \cdot (-2) = -6$ e $\det \begin{pmatrix} 1 & 0 & 0 \\ 7 & 0 & 0 \\ 8 & 15 & 19 \end{pmatrix} = 0$.

Proposizione Se una matrice ha una riga (o colonna) nulla, allora il suo determinante è nullo.

Dimostrazione. Basta sviluppare il determinante lungo la riga (o colonna) nulla. \square

Infine osserviamo la cosiddetta *formula di Binet*, già dimostrata per matrici di ordine due, e che afferma che il determinante del prodotto è uguale al prodotto dei determinanti.

Proposizione Date due matrici quadrate A e B di ordine n arbitrario si ha sempre: $\det(AB) = \det A \cdot \det B$.

Dimostrazione. Omessa. \square

Esercizio Dimostrare che l'uguaglianza $\det(A+B) = \det A + \det B$ non è vera in generale, cioè possiamo trovare matrici A e B tali che $\det(A+B) \neq \det A + \det B$.

Esercizio Data una matrice A di ordine n e un numero reale k , dimostrare che $\det(kA) = k^n \det A$. (*Suggerimento:* partire dai casi semplici $n = 2$ e $n = 3$ e generalizzare a n arbitrario.)

5.1 Complessità del calcolo del determinante

È evidente dagli esempi, e dalla stessa definizione, che il calcolo del determinante di una matrice $n \times n$ diventa via via più complicato al crescere dell'ordine n . Cerchiamo di quantificare questa complessità. Dalla definizione induttiva di determinante sappiamo che, dopo un certo numero di passi, il calcolo è ridotto a quello di un certo numero di determinanti di ordine due. Quanti, esattamente? Vogliamo quindi calcolare:

a_n = numero di determinanti di ordine due necessari al calcolo di un determinante di ordine n .

È evidente che $a_2 = 1$ e $a_3 = 3$. Dalla definizione di determinante con lo sviluppo lungo la prima riga ricaviamo subito la seguente relazione:

$$a_n = na_{n-1}.$$

Quindi la successione $\{a_n\}$ è l'unica avente le proprietà:

$$\begin{cases} a_2 = 1 \\ a_n = na_{n-1}. \end{cases}$$

Dunque $a_3 = 3$, $a_4 = 4 \cdot 3 = 12$, $a_5 = 5 \cdot 12 = 60$; si ha poi $a_8 = 20160$ e $a_{10} = 1814400$. Ma come cresce tale successione? È immediato verificare che, per ogni n , si ha:

$$a_n = \frac{1}{2}n!$$

dove $n!$ indica il fattoriale di n (cioè, il prodotto dei primi n numeri naturali). La crescita è estremamente rapida quando n è via via più grande.

Va poi considerato che il numero delle operazioni effettive (cioè moltiplicazioni o somme) necessarie al calcolo di un determinante di ordine n è ancora più grande, poichè oltre a calcolare i determinanti di ordine due dobbiamo anche sommarli e moltiplicarli opportunamente.

Vedremo che sarà possibile, con l'algoritmo di Gauss, ridurre i calcoli fino a un numero dell'ordine di n^3 .

6 Matrice inversa

Ricordiamo che una matrice quadrata A si dice invertibile se esiste una matrice quadrata B , dello stesso ordine, tale che $AB = BA = I$. Proviamo ora l'unicità dell'inversa.

Proposizione *Supponiamo che la matrice quadrata A sia invertibile: allora esiste un'unica matrice A^{-1} tale che $AA^{-1} = A^{-1}A = I$. La matrice A^{-1} è detta l'inversa di A .*

Dimostrazione. Per ipotesi, esiste una matrice B tale che $AB = BA = I$. Dimostriamo che B , con tale proprietà, è unica: infatti, supponiamo che esista una ulteriore matrice C tale che $AC = CA = I$. Allora si ha

$$CAB = (CA)B = IB = B$$

$$CAB = C(AB) = CI = C$$

che implica $B = C$. Possiamo dunque porre $A^{-1} = B$. \square

In questa sezione dimostreremo che una matrice quadrata di ordine n arbitrario è invertibile se e solo se ha determinante non nullo, generalizzando così l'analogo risultato già provato per le matrici di ordine 2. Inoltre daremo una formula esplicita per il calcolo della matrice inversa.

Ricordiamo che, dato l'elemento a_{ij} della matrice A , la matrice aggiunta di a_{ij} , denotata con A_{ij} , è la sottomatrice di ordine $n - 1$ ottenuta sopprimendo la i -esima riga e j -esima colonna di A .

Proposizione Supponiamo $\det A \neq 0$, e consideriamo la matrice $B = (b_{ij})$ con elementi

$$b_{ij} = (-1)^{i+j} \frac{\det A_{ji}}{\det A}.$$

Allora si ha $AB = BA = I$. In altre parole, $B = A^{-1}$.

Dimostrazione. Omessa. \square

La formula si esprime più semplicemente con l'introduzione del *complemento algebrico*.

6.1 Complemento algebrico di un elemento

Sia A una matrice quadrata e sia a_{ij} un suo elemento. Il *complemento algebrico* di a_{ij} è il numero reale:

$$\bar{a}_{ij} = (-1)^{i+j} \det A_{ij}.$$

Con questa notazione, le formule per il calcolo del determinante si semplificano. Ad esempio, lo sviluppo del determinante lungo la prima riga prende la forma:

$$\det A = a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12} + \cdots + a_{1n}\bar{a}_{1n}.$$

Più in generale si ha:

Proposizione Sia A una matrice quadrata di ordine n . Allora, per ogni $i = 1, \dots, n$ si ha:

$$\det A = a_{i1}\bar{a}_{i1} + a_{i2}\bar{a}_{i2} + \cdots + a_{in}\bar{a}_{in},$$

che rappresenta lo sviluppo del determinante lungo la i -esima riga. Inoltre, per ogni $k = 1, \dots, n$ si ha anche:

$$\det A = a_{1k}\bar{a}_{1k} + a_{2k}\bar{a}_{2k} + \cdots + a_{nk}\bar{a}_{nk},$$

che rappresenta lo sviluppo del determinante di A lungo la k -esima colonna.

Dimostrazione. Questo fatto è una conseguenza immediata delle formule per il calcolo del determinante viste nella sezione precedente. \square

In altre parole, il determinante si può calcolare in questo modo:

- Si sceglie una qualunque riga (o colonna).
- Si moltiplica ciascun elemento della riga (o colonna) scelta per il proprio complemento algebrico.

- Si somma il tutto.

Diamo ora un'espressione equivalente dell'inversa di una matrice.

Proposizione Sia $A = (a_{ij})$ una matrice quadrata di ordine n con determinante non nullo. Allora A è invertibile e l'inversa di A ha la seguente espressione:

$$A^{-1} = \frac{1}{\det A} \cdot \bar{A}^t.$$

dove $\bar{A} = (\bar{a}_{ij})$ indica la matrice dei complementi algebrici di A .

Quindi l'inversa si ottiene prendendo la *trasposta* della matrice dei complementi algebrici e moltiplicando per $\frac{1}{\det A}$.

Esempio Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix}$. Dimostriamo che A è invertibile e calcoliamo la sua matrice inversa.

Soluzione. Sviluppando lungo la terza colonna, ad esempio, si trova che $\det A = 6$. Calcoliamo la matrice dei complementi algebrici. Si ha $\bar{a}_{11} = \begin{vmatrix} 1 & 0 \\ 1 & 4 \end{vmatrix}$, etc. Tenendo conto dei segni, si ha che la matrice dei complementi algebrici è:

$$\bar{A} = \begin{pmatrix} \begin{vmatrix} 1 & 0 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} -1 & 0 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 4 & 4 & -2 \\ -5 & 1 & 1 \\ -3 & -3 & 3 \end{pmatrix}$$

Dobbiamo ora trasporre \bar{A} e moltiplicare per $\frac{1}{\det A}$. La matrice inversa è dunque:

$$A^{-1} = \frac{1}{6} \bar{A}^t = \frac{1}{6} \begin{pmatrix} 4 & -5 & -3 \\ 4 & 1 & -3 \\ -2 & 1 & 3 \end{pmatrix}.$$

Si può verificare con un calcolo diretto che $AA^{-1} = I$. \square

Esempio La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ è invertibile: infatti, essendo triangolare superiore, si vede immediatamente che ha determinante 1 (il prodotto dei suoi elementi diagonali). La sua

inversa, procedendo come nel caso precedente, risulta essere:

$$A^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

Notiamo che l'inversa è anch'essa triangolare superiore. Questo non è un caso.

- Sia A una matrice triangolare superiore (risp. inferiore) invertibile. Allora A^{-1} è triangolare superiore (risp. inferiore).

6.2 Un criterio necessario e sufficiente per l'invertibilità

Grazie alla proposizione precedente e al Teorema di Binet possiamo caratterizzare completamente le matrici invertibili.

Teorema Una matrice quadrata è invertibile se e solo se ha determinante non nullo.

Dimostrazione. Esercizio. \square

Esercizio Sia A una matrice diagonale con elementi diagonali d_1, \dots, d_n . Verificare che A è invertibile se e solo *tutti* gli elementi diagonali sono diversi da zero. In tal caso, l'inversa di A è la matrice diagonale avente elementi diagonali $\frac{1}{d_1}, \dots, \frac{1}{d_n}$.

6.3 Proprietà delle matrici invertibili

Iniziamo dalla seguente proprietà.

Proposizione Se A e B sono invertibili allora AB è invertibile e $(AB)^{-1} = B^{-1}A^{-1}$.

Dimostrazione. Esercizio. \square

Indichiamo con $\mathbf{GL}(n, \mathbf{R})$ l'insieme delle matrici di ordine n invertibili:

$$\mathbf{GL}(n, \mathbf{R}) = \{A \in M(n, n, \mathbf{R}) : \det A \neq 0\}.$$

La proposizione precedente afferma che $\mathbf{GL}(n, \mathbf{R})$ è *chiuso* rispetto al prodotto righe per colonne.

Esercizio Dimostrare che, se A è invertibile allora A^{-1} è invertibile e si ha $(A^{-1})^{-1} = A$. Inoltre $\det(A^{-1}) = \frac{1}{\det A}$.

Esercizio Verificare con un controesempio che la somma di due matrici invertibili potrebbe *non* essere invertibile.

Esercizio Dimostrare che, se A è invertibile e $k \neq 0$, allora kA è invertibile. Dare una formula per $(kA)^{-1}$.

Esercizio Dimostrare che, se A è invertibile, allora anche A^t è invertibile e si ha $(A^t)^{-1} = (A^{-1})^t$.

7 Teorema di Cramer (caso generale)

In questa sezione generalizzeremo il teorema di Cramer a sistemi *quadrati*, cioè di n equazioni in n incognite.

In generale, dato un sistema di m equazioni in n incognite, detta A la matrice dei coefficienti (di tipo $m \times n$), X il vettore colonna delle incognite (di tipo $n \times 1$) e B il vettore colonna dei termini noti (di tipo $m \times 1$) possiamo esprimere il sistema nella cosiddetta *forma matriciale* :

$$S : AX = B.$$

Se il sistema è quadrato la sua matrice dei coefficienti sarà anch'essa quadrata.

Teorema a) *Sia $S : AX = B$ un sistema di n equazioni in n incognite tale che la sua matrice dei coefficienti A sia invertibile. Allora S è risolubile e ammette una e una sola soluzione data da:*

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

b) *La soluzione è anche data dalle formule:*

$$x_i = \frac{\det A(i)}{\det A}, \quad i = 1, \dots, n$$

dove $A(i)$ è la matrice ottenuta da A sostituendo alla i -esima colonna la colonna dei termini noti.

La seconda parte del teorema è nota come *il metodo di Cramer*: esso fornisce un' importante semplificazione del calcolo della soluzione dato in a).

Dimostrazione. a) È immediata: poichè A è invertibile, esiste la matrice inversa A^{-1} . Moltiplicando ambo i membri di $AX = B$, a *sinistra*, per A^{-1} otteniamo la formula $X = A^{-1}B$.

b) Omessa. Comunque si tratta di esplicitare ciascuna delle entrate della soluzione X . \square

Definizione *Un sistema quadrato di n equazioni in n incognite, con matrice dei coefficienti avente determinante non nullo, si dice crameriano.*

Dunque, un sistema crameriano ammette sempre una e una sola soluzione.

Esempio Discutere le eventuali soluzioni del seguente sistema lineare nelle incognite x, y, z :

$$\begin{cases} x + 2y + 3z = 1 \\ -x + y = -1 \\ x + y + 4z = 2 \end{cases}$$

Soluzione. La matrice dei coefficienti è $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 1 & 1 & 4 \end{pmatrix}$. Un calcolo mostra che $\det A = 6 \neq 0$

dunque il teorema garantisce che esiste un'unica soluzione. Appliciamo dunque il metodo di Cramer descritto nella parte b) del teorema. Si ha:

$$\begin{aligned} x &= \frac{1}{6} \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 2 & 1 & 4 \end{pmatrix} = \frac{1}{2} \\ y &= \frac{1}{6} \cdot \det \begin{pmatrix} 1 & 1 & 3 \\ -1 & -1 & 0 \\ 1 & 2 & 4 \end{pmatrix} = -\frac{1}{2} \\ z &= \frac{1}{6} \cdot \det \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 2 \end{pmatrix} = \frac{1}{2} \end{aligned}$$

e l'unica soluzione è $\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Potevamo trovare la soluzione anche calcolando A^{-1} e quindi $X = A^{-1}B$, dove $B = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ è la colonna dei termini noti. Tale procedura è, normalmente, più lunga del metodo di Cramer. Ma poichè abbiamo già calcolato l'inversa (vedi sezione precedente), abbiamo, come verifica:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 4 & -5 & -3 \\ 4 & 1 & -3 \\ -2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

□

Esercizio Risolvere il sistema precedente con l'algoritmo di Gauss.

Esempio Discutere le eventuali soluzioni del seguente sistema lineare nelle incognite x, y, z :

$$\begin{cases} x + 2y + 3z = 1 \\ -x + y = -1 \\ x + 5y + 6z = 0 \end{cases}$$

Soluzione. Abbiamo in questo caso, come matrice dei coefficienti: $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 1 & 5 & 6 \end{pmatrix}$. Ma un

calcolo mostra che $\det A = 0$. Dunque *non* possiamo applicare il teorema di Cramer; il sistema potrebbe non essere compatibile. In effetti, applicando l'algoritmo di Gauss si può verificare che S non ammette soluzioni. \square

8 Determinante e algoritmo di Gauss

Vediamo ora alcune proprietà del determinante che permettono, a volte, di semplificare il calcolo.

Proposizione a) *Se si scambiano due righe (o colonne) della matrice, il determinante cambia di segno. In particolare, se due righe (o colonne) sono uguali il determinante è nullo.*

b) *Se A' si ottiene da A moltiplicando una riga (o colonna) per il numero k , allora $\det A' = k \det A$. In particolare, $\det(kA) = k^n \det A$.*

c) *Se si somma ad una riga (risp. colonna) un qualunque multiplo di un'altra riga (risp. colonna), il determinante non cambia.*

Dimostrazione. Poichè $\det A = \det(A^t)$, basta dimostrare le affermazioni riguardo alle righe della matrice. Daremo la dimostrazione nel caso in cui $n = 2, 3$: questo darà un'idea della dimostrazione nel caso generale.

a) Questo fatto è immediato dalla definizione se la matrice è 2×2 . Supponiamo ora che A abbia ordine 3, e sia A' la matrice ottenuta da A scambiando fra loro, ad esempio, le prime due righe. Sviluppiamo il determinante di A' lungo la terza riga: siccome a) è vera per le matrici di ordine due, otteniamo che $\det A' = -\det A$. Il generale, se A' si ottiene da A scambiando due righe qualunque, basta sviluppare il determinante di A' lungo la riga non interessata dallo scambio. È chiaro ora come continuare se l'ordine $n = 4, 5, \dots$.

b) Basta sviluppare il determinante lungo la riga che si è moltiplicato per k .

c) Supponiamo che $n = 2$ e che A' si ottenga da A , ad esempio, sommando alla seconda riga la prima moltiplicata per k . Esplicitamente:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A' = \begin{pmatrix} a & b \\ c + ka & d + kb \end{pmatrix}.$$

Un calcolo mostra che, in effetti, $\det A' = \det A$. È chiaro a questo punto che l'affermazione è vera per l'ordine $n = 2$. Il caso $n = 3$ si riduce al caso $n = 2$ semplicemente sviluppando il determinante di A' lungo la riga non interessata dall'operazione (ad esempio, se A' si ottiene da

A sommando alla seconda riga la prima moltiplicata per k svilupperemo il determinante di A' lungo la terza riga). Dunque c) è vera per $n = 3$ e iterando via via l'argomento appena descritto otteniamo che c) è vera per ogni ordine n . \square

Esempio Dimostriamo che, se $A = \begin{pmatrix} 100 & 101 & 102 \\ 125 & 126 & 127 \\ 150 & 151 & 152 \end{pmatrix}$, allora $\det A = 0$.

Soluzione. Moltiplichiamo la prima riga per -1 e sommiamo alla seconda ($R_2 \rightarrow R_2 - R_1$), e successivamente moltiplichiamo la prima riga per -1 e sommiamo alla terza ($R_3 \rightarrow R_3 - R_1$). Per la proposizione appena dimostrata, queste due operazioni non alterano il determinante. Dunque:

$$\det A = \det \begin{pmatrix} 100 & 101 & 102 \\ 25 & 25 & 25 \\ 50 & 50 & 50 \end{pmatrix}.$$

Ora moltiplichiamo la seconda riga per -2 e sommiamo alla terza. Si ha dunque:

$$\det A = \det \begin{pmatrix} 100 & 101 & 102 \\ 25 & 25 & 25 \\ 0 & 0 & 0 \end{pmatrix},$$

Avendo una riga nulla, il determinante è nullo. \square

Esempio Calcoliamo il determinante della matrice $A = \begin{pmatrix} 0 & 1 & 2 & -1 \\ 1 & -2 & 1 & 3 \\ 1 & 4 & 1 & 2 \\ 2 & 2 & 2 & 5 \end{pmatrix}$.

Soluzione. Scambiando le prime due righe otteniamo la matrice

$$A_1 = \begin{pmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 2 & -1 \\ 1 & 4 & 1 & 2 \\ 2 & 2 & 2 & 5 \end{pmatrix}$$

dunque $\det A = -\det A_1$. Ora applichiamo le operazioni $R_3 \rightarrow R_3 - R_1$ e $R_4 \rightarrow R_4 - 2R_1$. Otteniamo

$$A_2 = \begin{pmatrix} 1 & -2 & 1 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 6 & 0 & -1 \\ 0 & 6 & 0 & -1 \end{pmatrix}$$

e $\det A_2 = \det A_1 = -\det A$. Ma A_2 ha due righe uguali: per la proposizione si ha $\det A_2 = 0$ e quindi anche $\det A = 0$. \square

Esercizio Giustificare i seguenti passaggi, usando le parti b) e c) della Proposizione:

$$\begin{aligned} \det \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ 2 & 1 & \frac{1}{2} \\ \frac{2}{3} & 1 & -\frac{1}{2} \end{pmatrix} &= \frac{1}{72} \det \begin{pmatrix} 3 & 2 & 1 \\ 4 & 2 & 1 \\ 4 & 6 & -3 \end{pmatrix} \\ &= \frac{1}{72} \det \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 0 \\ 4 & 6 & -3 \end{pmatrix} \\ &= \frac{1}{6} \end{aligned}$$

8.1 Calcolo del determinante con l'algoritmo di Gauss

Ricordiamo ora le operazioni elementari sulle righe di una matrice, usate nell'algoritmo di Gauss:

- 1) Scambiare fra loro due righe.
- 2) Moltiplicare una riga per uno scalare $k \neq 0$.
- 3) Sommare a una data riga un multiplo di un'altra riga.

Vogliamo vedere l'effetto di ciascuna di tali operazioni sul determinante della matrice (supposta, ovviamente, quadrata). Dalla proposizione, otteniamo rispettivamente:

- 1) Il determinante cambia di segno.
- 2) Il determinante viene moltiplicato per k .
- 3) Il determinante rimane invariato.

Applicando tali operazioni in successione osserviamo che, se il determinante è non nullo all'inizio, tale rimarrà dopo aver effettuato operazioni elementari sulle righe. Dunque:

Proposizione *Supponiamo di ridurre la matrice quadrata A ad una matrice a scalini \tilde{A} mediante l'algoritmo di Gauss. Allora $\det A = 0$ se e solo se $\det \tilde{A} = 0$. In particolare, A è invertibile se e solo se \tilde{A} è invertibile.*

In realtà, nella riduzione a scalini si possono usare solamente le operazioni 1) e 3); in tal caso, occorre solamente tenere conto del numero degli scambi di riga:

Proposizione *Supponiamo di ridurre la matrice quadrata A ad una matrice a scalini \tilde{A} utilizzando solamente le operazioni 1) e 3). Allora, se s è il numero degli scambi, si ha:*

$$\det A = (-1)^s \det \tilde{A}$$

Esempio Calcolare il determinante di $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix}$ usando l'algoritmo di Gauss.

Soluzione. Applicando in successione le operazioni:

$$\begin{aligned} R_3 &\rightarrow R_3 - R_1 \\ R_3 &\rightarrow R_3 - 2R_2 \end{aligned}$$

otteniamo $\tilde{A} = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{pmatrix}$ che è a scalini, dunque triangolare. Non ci sono scambi di righe: concludiamo che $\det A = \det \tilde{A} = -6$. \square

Esempio Calcolare il determinante di $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$ usando l'algoritmo di Gauss.

Soluzione. Scambiamo le prime due righe: $R_1 \leftrightarrow R_2$. Applichiamo poi, in successione, le operazioni:

$$\begin{aligned} R_3 &\rightarrow R_3 - R_1 \\ R_4 &\rightarrow R_4 - R_1 \\ R_3 &\rightarrow R_3 + R_2 \\ R_4 &\rightarrow R_4 + R_3 \end{aligned}$$

per ottenere la matrice a scalini $\tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Abbiamo effettuato un solo scambio:

dunque

$$\det A = -\det \tilde{A} = -3.$$

Parte 3. Rango e teorema di Rouché-Capelli

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Rango di una matrice, 1
- 2 Teorema degli orlati, 3
- 3 Calcolo con l'algoritmo di Gauss, 6
- 4 Matrici dipendenti da parametri, 10
- 5 Teorema di Rouché-Capelli, 11
- 6 Sistemi lineari omogenei, 16
- 7 Dipendenza e indipendenza lineare di vettori di \mathbf{R}^n , 19
- 8 Criterio del rango, 21

1 Rango di una matrice

1.1 Minori di una matrice

- Un *minore* di una matrice A è per definizione una sottomatrice quadrata di A . Un minore si ottiene intersecando n righe ed n colonne di A .

Esempio $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ha quattro minori di ordine 1 (cioè $(1), (2), (3), (4)$) e un minore di ordine 2 (la matrice stessa).

Esempio $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Oltre ai minori di ordine 1 (gli elementi della matrice), abbiamo tre minori di ordine 2:

$$\mu_{12,12} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad \mu_{12,13} = \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}, \quad \mu_{12,23} = \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}.$$

dove $\mu_{ij,hk}$ denota il minore ottenuto scegliendo le righe i, j e le colonne h, k .

Esempio $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & l \end{pmatrix}$. Abbiamo nove minori di ordine due:

$$\mu_{12,12} = \begin{pmatrix} a & b \\ d & e \end{pmatrix}, \mu_{12,13} = \begin{pmatrix} a & c \\ d & f \end{pmatrix}, \dots, \mu_{23,23} = \begin{pmatrix} e & f \\ h & l \end{pmatrix}$$

e ovviamente solo un minore di ordine tre (la matrice).

Esercizio a) Elencare tutti i minori di ordine tre di una matrice 3×4 .

b) Quanti sono i minori di ordine p di una matrice $m \times n$? (*Risposta:* $\binom{m}{p} \cdot \binom{n}{p}$).

1.2 Definizione di rango

Definizione Una matrice A di tipo $p \times q$ ha rango n se:

1. Esiste almeno un minore di ordine n con determinante non nullo.
2. Tutti i minori di ordine $n+1$ (se esistono) hanno determinante nullo.

Espresso a parole, il rango di una matrice è l'ordine massimo di un minore di A avente determinante non nullo. Denoteremo il rango di A con il simbolo: $\text{rk}A$. Per convenzione, il rango della matrice nulla è posto uguale a zero. Tutte le altre matrici hanno rango maggiore o uguale a 1.

- Osserviamo che, se esiste un minore di ordine h con determinante non nullo, allora $\text{rk}A \geq h$.

Segue immediatamente dalla definizione che:

- Se A ha p righe e q colonne si ha sempre $0 \leq \text{rk}A \leq \min\{p, q\}$ (il minimo tra p e q). Questo semplicemente perchè non ci sono minori di ordine superiore a tale numero. Se $\text{rk}A = \min\{p, q\}$ diremo che A ha rango massimo.
- Se A è quadrata, di ordine n , allora $\text{rk}A = n$ (massimo) se e solo se $\det A \neq 0$.

Esempio $A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$. Poichè $\det A \neq 0$ si ha $\text{rk}A = 2$.

Esempio $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Il rango vale 1 poichè $\det A = 0$ e A non è nulla.

Esempio $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{pmatrix}$. Il rango vale 1 (i tre minori di ordine 2 hanno determinante nullo).

Esempio $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \end{pmatrix}$. Il rango vale 2 (c'è infatti almeno un minore di ordine 2 a determinante non nullo: quale?).

Esempio $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \\ 3 & 2 & 2 \end{pmatrix}$. Il minore $\mu_{12,12}$ ha determinante non nullo, dunque $\text{rk}A$ vale 2 oppure 3. Poichè $\det A = 0$, si ha effettivamente $\text{rk}A = 2$.

Esercizio Verificare che, se tutti i minori di un certo ordine h hanno determinante nullo, allora tutti i minori di ordine più grande di h avranno determinante nullo.

Proposizione *Si ha sempre $\text{rk}A = \text{rk}(A^t)$.*

Dimostrazione. I minori di A^t si ottengono trasponendo quelli di A (e viceversa). Siccome il determinante assume lo stesso valore su un minore M e sul suo trasposto, si ha immediatamente l'asserto. \square

Esempio $A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 4 \end{pmatrix}$. Si vede subito che il rango vale almeno 2, poiché $\begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} \neq 0$

0. Esaminiamo i determinanti dei minori di ordine 3:

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 0 \\ 1 & -3 & 2 \end{vmatrix} = 0$$

$$\begin{vmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 1 & -3 & 4 \end{vmatrix} = 0$$

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \end{vmatrix} = 0$$

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ -3 & 2 & 4 \end{vmatrix} = 0$$

Sono tutti nulli, dunque $\text{rk}A = 2$.

Esempio $A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 4 \\ 2 & 3 & 1 & 2 \end{pmatrix}$. Si vede che $\text{rk}A \geq 2$. Inoltre un calcolo mostra che

$\det A = 0$. Dunque il rango può valere 2 oppure 3. A questo punto dovremmo esaminare

tutti i minori di ordine 3, ma sono 16. Lasciamo in sospeso il calcolo, poiché nella prossima sezione enunceremo un teorema che ci permetterà di abbreviare i calcoli.

2 Teorema degli orlati

2.1 Minori orlati di un minore dato

Dato un minore M di ordine n di una matrice A , diremo che il minore M' di ordine $n+1$ è un *orlato* di M se esso contiene M , se cioè si ottiene da M aggiungendo elementi di un'altra riga e un'altra colonna di A .

Esempio Sia $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$. Osserviamo che A ammette sei minori di ordine 2 (elencare tali minori). Ora fissiamo il minore di ordine 1 dato da $M = \mu_{2,1} = (5)$ ed elenchiamo i minori orlati di M . Essi sono tre; precisamente:

$$\mu_{12,12} = \begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix}, \quad \mu_{12,13} = \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}, \quad \mu_{12,14} = \begin{pmatrix} 1 & 4 \\ 5 & 8 \end{pmatrix}.$$

Esempio Sia ora $A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 4 \\ 2 & 3 & 1 & 2 \end{pmatrix}$. Fissiamo il minore di ordine 2 dato da

$$M = \mu_{12,12} = \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix},$$

ed elenchiamo i minori orlati di M . Essi sono in tutto quattro, e si ottengono aggiungendo, rispettivamente, elementi dalla:

- terza riga e terza colonna: $\mu_{123,123} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix},$
- terza riga e quarta colonna: $\mu_{123,124} = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 1 & -3 & 4 \end{pmatrix},$
- quarta riga e terza colonna: $\mu_{124,123} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 3 & 1 \end{pmatrix},$
- quarta riga e quarta colonna: $\mu_{124,124} = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 2 & 3 & 2 \end{pmatrix}.$

2.2 Teorema degli orlati ed esempi

Enunciamo ora il *teorema degli orlati*.

Teorema Sia A una matrice $p \times q$ e M un suo minore di ordine n con determinante non nullo. Se tutti gli orlati di M hanno determinante nullo allora il rango di A è esattamente n (cioè l'ordine di M).

Dimostrazione. Omessa. \square

Esempio Calcoliamo il rango della matrice $A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 4 \\ 2 & 3 & 1 & 2 \end{pmatrix}$. Il minore

$$M = \mu_{12,12} = \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix}$$

ha determinante non nullo. Esaminiamo ora i determinanti dei minori *orlati* di M , che sono solamente quattro, e sono già stati elencati precedentemente:

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 0 \\ 1 & -3 & 2 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 1 & -3 & 4 \end{vmatrix} = 0,$$
$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 3 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 2 & 3 & 2 \end{vmatrix} = 0.$$

Tutti i minori orlati hanno determinante nullo: possiamo applicare il teorema degli orlati, e concludere che il rango vale 2.

- È chiaro che, se il determinante di almeno uno dei minori orlati fosse stato diverso da zero, allora il rango della matrice risulterebbe almeno pari a 3 e avremmo dovuto continuare, esaminando i minori di ordine 4 (in questo caso, la matrice stessa).

Il teorema degli orlati ha permesso di ridurre i calcoli: invece di considerare tutti i (sedici) minori di ordine 3, è stato sufficiente considerare solo i quattro minori orlati del minore precedentemente scelto. \square

Esempio Data la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ x & y & z & w \end{pmatrix}$$

determinare x, y, z, w in modo che $\text{rk} A = 1$.

Soluzione. È chiaro che il rango vale almeno 1. Per essere proprio uguale a 1, tutti gli orlati del minore $M = (1)$ devono avere determinante nullo. Otteniamo le condizioni:

$$\begin{vmatrix} 1 & 2 \\ x & y \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ x & z \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ x & w \end{vmatrix} = 0,$$

e quindi

$$\begin{cases} y = 2x \\ z = 3x \\ w = 4x. \end{cases}$$

Dunque la matrice deve essere del tipo:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ x & 2x & 3x & 4x \end{pmatrix},$$

con x parametro reale. Notiamo che le righe (e anche le colonne) sono proporzionali. \square

- Con un argomento simile, possiamo verificare che una matrice avente solo due righe ha rango 1 se e solo se è non nulla e ha righe proporzionali. Stessa cosa per le matrici con due colonne. Ad esempio

$$\begin{pmatrix} 2 & 3 & -1 \\ 4 & 6 & -2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 \\ 2 & 6 \\ -1 & -3 \end{pmatrix}$$

hanno rango 1, mentre $\begin{pmatrix} 2 & 3 & -1 \\ 1 & 1 & -2 \end{pmatrix}$ ha rango 2.

Esercizio Dimostrare che il rango di una matrice diagonale è uguale al numero degli elementi diagonali non nulli.

Esercizio Sia A' la matrice ottenuta da A aggiungendo una riga (o una colonna). Dimostrare che $\text{rk} A' \geq \text{rk} A$ e più precisamente si ha $\text{rk} A' = \text{rk} A$ oppure $\text{rk} A' = \text{rk} A + 1$. In altre parole, aggiungendo una riga (o colonna) il rango rimane inalterato oppure aumenta di un'unità (a seconda dei casi).

3 Rango e algoritmo di Gauss

Abbiamo visto nella prima parte come l'algoritmo di Gauss permetta di risolvere i sistemi lineari. In questa sezione useremo l'algoritmo di Gauss per calcolare il rango di una matrice: tale metodo, almeno per matrici di grandi dimensioni, è molto più efficiente del metodo dei minori usato per definire il rango.

3.1 Il rango di una matrice a scalini

Ricordiamo la definizione di matrice a scalini, già data precedentemente.

Definizione Una matrice A si dice a scalini se verifica entrambe le seguenti proprietà:

1. Se una riga è nulla, tutte le righe ad essa sottostanti sono nulle.
2. Sotto il primo elemento non nullo di ciascuna riga, e sotto tutti gli zeri che lo precedono, ci sono elementi nulli.

In una matrice a scalini, il primo elemento non nullo di una riga è detto il *pivot* della data riga. Osserviamo che il numero dei pivot uguaglia il numero delle righe non nulle.

Vogliamo ora calcolare il rango di una matrice a scalini. Iniziamo con un esempio.

Esempio Calcoliamo il rango della matrice a scalini $A = \begin{pmatrix} 1 & 2 & 3 & 0 & 5 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Il numero

dei pivot (righe non nulle) è 3. Chiaramente, ogni minore di ordine 4 ha almeno una riga nulla, dunque determinante nullo. Ne segue che il rango non può essere maggiore di 3. Dimostriamo che è proprio uguale a 3. Consideriamo il minore di ordine 3 individuato dalle righe e dalle colonne corrispondenti ai pivot: nel nostro caso,

$$\mu_{123,135} = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Tale minore è triangolare superiore, con elementi diagonali dati dai pivot di A : il suo determinante, essendo il prodotto dei pivot, è sicuramente diverso da zero. Dunque: esiste un minore di ordine 3 con determinante non nullo, e tutti i minori di ordine 4 hanno determinante nullo. La conclusione è che il rango di A vale 3, esattamente il numero dei pivot di A .

Questo è sempre vero.

Teorema Il rango di una matrice a scalini uguaglia il numero dei suoi pivot.

Dimostrazione. Basta generalizzare l'argomento dell'esempio precedente. Sia r il numero dei pivot. Allora:

- ogni minore di ordine $r + 1$ ha una riga nulla, dunque determinante nullo;
- il minore di ordine r individuato dalle righe e dalle colonne cui appartengono i pivot è triangolare superiore, con elementi diagonali non nulli, dunque ha determinante non nullo.

Da queste due osservazioni segue immediatamente che il rango di A vale proprio r . \square

3.2 Rango e operazioni elementari sulle righe

Ricordiamo che le operazioni elementari sulle righe di una matrice sono:

- 1) Scambiare due righe fra loro.
- 2) Moltiplicare una riga per uno scalare non nullo.
- 3) Sommare ad una riga un multiplo di un'altra riga.

Sia A una matrice quadrata, e supponiamo che la matrice A' si ottenga da A dopo aver applicato una successione di operazioni elementari sulle righe. Non è difficile dimostrare che allora anche A si ottiene da A' mediante una successione di operazioni elementari sulle righe. Diremo che A e A' sono matrici *equivalenti per righe*.

Ricordiamo infine l'algoritmo di Gauss:

- Sia data una qualunque matrice A . Con opportune operazioni elementari sulle righe, è sempre possibile arrivare ad una matrice \tilde{A} tale che:

1. \tilde{A} è equivalente per righe ad A .
2. \tilde{A} è a scalini.

\tilde{A} è anche detta una *matrice ridotta* di A .

Il calcolo del rango mediante l'algoritmo di Gauss si basa sul seguente risultato fondamentale:

Teorema *Le operazioni elementari sulle righe non alterano il rango. Quindi, matrici equivalenti per righe hanno lo stesso rango.*

Dimostrazione. Non daremo una dimostrazione formale, ma solo un cenno. Ricordiamo l'effetto di ciascuna delle operazioni 1), 2), 3) elementari sul determinante di una matrice, dimostrate nella Parte 1:

- 1) il determinante cambia di segno,
- 2) il determinante viene moltiplicato per un numero non nullo,
- 3) il determinante rimane invariato.

Ne segue che le operazioni elementari sulle righe di una matrice possono alterare il determinante di un suo qualunque minore, ma solo (eventualmente) moltiplicandolo per un numero non nullo: dunque queste operazioni non alterano la proprietà che tale determinante sia nullo o no. Poiché solo questo conta nella definizione di rango, la conclusione è intuitivamente evidente. \square

Corollario *Sia A una matrice e \tilde{A} una ridotta di A (cioè, una matrice a scalini equivalente per righe ad A). Allora il rango di A è uguale al numero dei pivot di \tilde{A} .*

Dimostrazione. Dalla proposizione precedente, abbiamo che $\text{rk}A = \text{rk}\tilde{A}$. D'altra parte, sappiamo già che il rango di \tilde{A} è uguale al numero dei suoi pivot. \square

3.3 Calcolo del rango con l'algoritmo di Gauss

Dal corollario appena enunciato vediamo che, per calcolare il rango di una matrice A , possiamo procedere nel seguente modo:

- Con l'algoritmo di Gauss, riduciamo A ad una matrice a scalini \tilde{A} .
- Contiamo il numero dei pivot di \tilde{A} : tale numero è proprio il rango di A .

Esempio Calcoliamo il rango della matrice $A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 4 \\ 2 & 3 & 1 & 2 \end{pmatrix}$, già considerata nella

sezione precedente.

Soluzione. Riduciamo A a una matrice a scalini. Con le operazioni $R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - R_1, R_4 \rightarrow R_4 - 2R_1$ arriviamo alla matrice equivalente:

$$A_1 = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 3 & -1 & -2 \\ 0 & -3 & 1 & 2 \\ 0 & 3 & -1 & -2 \end{pmatrix}.$$

Con le operazioni $R_3 \rightarrow R_3 + R_2, R_4 \rightarrow R_4 - R_2$ arriviamo alla matrice a scalini (matrice ridotta):

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

Poiché \tilde{A} ha due pivot, il rango di A vale 2. \square

- Osserviamo che, se in una matrice la riga R_i è proporzionale (in particolare, uguale) alla riga R_j , allora possiamo sostituire R_i con una riga nulla, ottenendo sempre una matrice equivalente (perché?). In particolare, se tutte le righe sono proporzionali a una data riga (supposta non nulla), allora il rango vale 1.

Esempio Il rango della matrice $\begin{pmatrix} 1 & 2 & 0 & -1 & 2 \\ 2 & 4 & 0 & -2 & 4 \\ -2 & -4 & 0 & 2 & -4 \end{pmatrix}$ vale 1.

Esempio Il rango della matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ vale 2 (spiegare perché). In particolare, $\det A = 0$.

4 Matrici dipendenti da uno o più parametri

A volte occorre considerare matrici che dipendono da parametri.

Esempio Si consideri la matrice $A = \begin{pmatrix} 1 & k & 4-k \\ k & 4 & 4 \end{pmatrix}$ dipendente dal parametro $k \in \mathbf{R}$. Calcolare il rango di A al variare di k .

Soluzione. Risolviamo prima con il metodo degli orlati. È chiaro che $\text{rk}A$ vale 1 oppure 2. Fissiamo il minore $\mu_{1,1} = 1$ e consideriamo i determinanti dei suoi minori orlati:

$$\begin{vmatrix} 1 & k \\ k & 4 \end{vmatrix} = 4 - k^2 = (2-k)(2+k)$$

$$\begin{vmatrix} 1 & 4-k \\ k & 4 \end{vmatrix} = (k-2)^2$$

Ora, entrambi gli orlati hanno determinante nullo se e solo se $k = 2$: in tal caso il rango vale 1. Se $k \neq 2$ il secondo orlato ha determinante non nullo, dunque $\text{rk}A = 2$. Conclusione:

$$\text{rk}A = \begin{cases} 1 & \text{se } k = 2, \\ 2 & \text{se } k \neq 2. \end{cases}$$

Alla stessa conclusione potevamo arrivare usando l'algoritmo di Gauss. Infatti, riduciamo A a una matrice a scalini. Per far questo, è sufficiente l'operazione $R_2 \rightarrow R_2 - kR_1$:

$$\tilde{A} = \begin{pmatrix} 1 & k & 4-k \\ 0 & 4-k^2 & k^2-4k+4 \end{pmatrix} = \begin{pmatrix} 1 & k & 4-k \\ 0 & (2-k)(2+k) & (k-2)^2 \end{pmatrix}$$

Osserviamo che, se $k \neq 2$, abbiamo due pivot, dunque il rango è 2. Se $k = 2$ si ha $\tilde{A} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ quindi il rango è 1. In conclusione:

$$\text{rk}A = \begin{cases} 1 & \text{se } k = 2, \\ 2 & \text{se } k \neq 2. \end{cases}$$

□

Esempio Calcoliamo il rango di $A = \begin{pmatrix} 1 & k & 3 \\ 2k & 2 & 6 \\ -k & -1 & -3 \end{pmatrix}$ al variare di k .

Soluzione. Vediamo innanzitutto quando il rango è massimo, cioè 3: questo avviene se e solo se $\det A \neq 0$. Ma un calcolo mostra che

$$\det A = 0 \quad \text{per ogni } k.$$

Dunque il rango vale 1 oppure 2. Consideriamo il minore

$$\mu_{12,23} = \begin{pmatrix} k & 3 \\ 2 & 6 \end{pmatrix},$$

avente determinante $6k - 6$ che si annulla per $k = 1$. Dunque, se $k \neq 1$ il rango è 2; se $k = 1$ la matrice diventa:

$$\begin{pmatrix} 1 & 1 & 3 \\ 2 & 2 & 6 \\ -1 & -1 & -3 \end{pmatrix},$$

che ha evidentemente rango 1 (tutte le righe sono proporzionali alla prima). In conclusione:

$$\text{rk} A = \begin{cases} 2 & \text{se } k \neq 1, \\ 1 & \text{se } k = 1. \end{cases}$$

5 Il teorema di Rouché'-Capelli

La nozione di rango ci permette di dimostrare un risultato, di tipo essenzialmente teorico, sulla compatibilità o meno di un dato sistema lineare. Consideriamo un sistema lineare di m equazioni nelle n incognite x_1, \dots, x_n , che possiamo scrivere in forma matriciale

$$AX = B,$$

dove A è la matrice dei coefficienti (di tipo $m \times n$), $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ è il vettore colonna delle

incognite e B è il vettore colonna dei termini noti. Il sistema è descritto da due matrici: la matrice dei coefficienti, appunto, e la matrice completa A' , ottenuta aggiungendo ad A la colonna dei termini noti. Chiaramente A' è di tipo $m \times (n + 1)$.

Il teorema di Rouché'-Capelli permette di stabilire la compatibilità conoscendo solamente il rango di A e di A' . Precisamente:

Teorema (Rouche'-Capelli) *Sia S un sistema lineare di m equazioni in n incognite con matrice dei coefficienti A e matrice completa A' . Allora:*

a) *S è compatibile se e solo se $\text{rk}A = \text{rk}A'$.*

Supponiamo ora che S sia risolubile, e poniamo $\text{rk}A = \text{rk}A' = r$. Allora:

b) *Il sistema ammette una e una sola soluzione se e solo se $r = n$.*

c) *Il sistema ammette ∞^{n-r} soluzioni (cioè infinite soluzioni dipendenti da $n-r$ parametri indipendenti) se e solo se $r < n$.*

Dimostrazione. Dimosteremo solamente la parte a) del teorema. Lo schema è il seguente: prima verifichiamo il teorema per i sistemi a scalini, quindi usiamo il fatto che ogni sistema lineare è equivalente a un sistema a scalini. Faremo uso della seguente proprietà delle matrici a scalini:

- Sia \tilde{A}' una matrice a scalini, e sia \tilde{A} la sottomatrice ottenuta da \tilde{A}' sopprimendo l'ultima colonna. Allora anche \tilde{A} è a scalini.

Sia dunque \tilde{S} un sistema a scalini, con matrice dei coefficienti \tilde{A} e matrice completa \tilde{A}' , a scalini. Poiché \tilde{A} è la sottomatrice ottenuta da \tilde{A}' sopprimendo l'ultima colonna, vediamo che anche \tilde{A} è a scalini. Ora sappiamo che il sistema a scalini \tilde{S} è compatibile se e solo se l'ultimo pivot della sua matrice completa \tilde{A}' non cade nell'ultima colonna. Ma questo equivale a dire che:

- Il sistema a scalini \tilde{S} è compatibile se e solo se \tilde{A} e \tilde{A}' hanno lo stesso numero di pivot.

Poiché il numero di pivot di una matrice a scalini è proprio il rango, otteniamo infine:

- Il sistema a scalini \tilde{S} è compatibile se e solo se \tilde{A} e \tilde{A}' hanno lo stesso rango.

Abbiamo dunque dimostrato la parte a) del teorema per i sistemi a scalini.

Supponiamo ora che S sia un sistema arbitrario, con matrice dei coefficienti A e matrice completa A' . Applicando l'algoritmo di Gauss alla matrice completa A' vediamo che S risulterà equivalente a un sistema a scalini \tilde{S} con matrice dei coefficienti \tilde{A} e matrice completa \tilde{A}' , entrambe a scalini. È evidente che le operazioni elementari effettuate sulle righe di A' inducono operazioni elementari sulle righe di A , poichè A si ottiene da A' sopprimendo l'ultima colonna. Dunque \tilde{A} è equivalente per righe ad A e di conseguenza $\text{rk}A = \text{rk}\tilde{A}$. Ovviamente $\text{rk}A' = \text{rk}\tilde{A}'$. In conclusione, le seguenti affermazioni sono via via equivalenti:

$$\begin{aligned} S \text{ è compatibile} &\iff \tilde{S} \text{ è compatibile} \\ &\iff \text{rk}\tilde{A} = \text{rk}\tilde{A}' \\ &\iff \text{rk}A = \text{rk}A'. \end{aligned}$$

Questo dimostra la parte a). Le affermazioni b) e c) si possono dimostrare con lo stesso procedimento, ma non entreremo nei dettagli. \square

Esempio Sia

$$S : \begin{cases} x_1 + 2x_2 + x_4 = 1 \\ 2x_1 + 4x_2 + x_3 = 3 \\ x_1 + 2x_2 + x_3 - x_4 = 2 \end{cases}.$$

La matrice dei coefficienti e la matrice completa sono, rispettivamente:

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

Un calcolo (usare il teorema degli orlati) mostra che $\text{rk}A = \text{rk}A' = 2$. Dunque il sistema è compatibile; poiché $n = 4$ e $r = 2$ il teorema di Rouché-Capelli afferma che il sistema ammette ∞^2 soluzioni. Verifichiamo il risultato con l'algoritmo di Gauss, riducendo A' a una forma a scalini. Con le operazioni $R_2 \rightarrow R_2 - 2R_1$ e $R_3 \rightarrow R_3 - R_1$ arriviamo alla matrice:

$$A'_1 = \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

Con l'operazione $R_3 \rightarrow R_3 - R_2$ arriviamo alla forma a scalini:

$$\tilde{A}' = \begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il sistema \tilde{S} rappresentato da \tilde{A}' è compatibile, e poichè ci sono due pivot e quattro incognite esso ammette ∞^2 soluzioni. Notiamo anche che la matrice dei coefficienti di \tilde{S} è:

$$\tilde{A} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Notiamo infine che si ha $\text{rk}\tilde{A} = \text{rk}\tilde{A}' = 2$ e che A è equivalente ad \tilde{A} . \square

Esempio Consideriamo il sistema: $S : \begin{cases} x + y = \alpha \\ 2x + y = \beta \end{cases}$. Dimostrare che S è compatibile, e ammette un'unica soluzione, per ogni scelta di $\alpha, \beta \in \mathbf{R}$.

Soluzione. Si ha $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ e $A' = \begin{pmatrix} 1 & 1 & \alpha \\ 2 & 1 & \beta \end{pmatrix}$. Siccome $\det A = -1 \neq 0$, il rango di A è $r = 2$ (massimo), dunque il rango di A' è anch'esso 2. Dal teorema di Rouché-Capelli otteniamo l'asserto. \square

Esempio Discutere il sistema $S : \begin{cases} x + y + z = 1 \\ -x + y + 5z = 0 \\ 2y + 6z = 0 \end{cases}$.

Soluzione. $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 5 \\ 0 & 2 & 6 \end{pmatrix}$ ha rango 2: infatti $\det A = 0$ e il minore $\mu_{12,12} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

ha determinante diverso da zero. D'altra parte, la matrice completa $A' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 5 & 0 \\ 0 & 2 & 6 & 0 \end{pmatrix}$ ha rango 3, come si vede considerando il minore ottenuto sopprimendo la prima colonna. Dunque $\text{rk} A \neq \text{rk} A'$ e il sistema è incompatibile. \square

Esempio Verificare il risultato dell'esempio precedente con l'algoritmo di Gauss.

Esempio Stabilire per quali valori dei parametri a, b, c il sistema lineare

$$S : \begin{cases} x + y + z = a \\ -x + y + 5z = b \\ 2y + 6z = c \end{cases}$$

è compatibile.

Soluzione. La matrice dei coefficienti è $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 5 \\ 0 & 2 & 6 \end{pmatrix}$ e ha rango 2, indipendentemente

dai valori di a, b, c . Il sistema sarà dunque compatibile se e solo se $\text{rk} A' = 2$. Vediamo quando ciò accade. La matrice completa è

$$A' = \begin{pmatrix} 1 & 1 & 1 & a \\ -1 & 1 & 5 & b \\ 0 & 2 & 6 & c \end{pmatrix}$$

Il minore $\mu_{12,12} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ ha determinante non nullo. Dunque $\text{rk}A' = 2$ se e solo se tutti i suoi minori orlati hanno determinante nullo, dunque se e solo se:

$$\det \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 5 \\ 0 & 2 & 6 \end{pmatrix} = 0 \quad \text{e} \quad \det \begin{pmatrix} 1 & 1 & a \\ -1 & 1 & b \\ 0 & 2 & c \end{pmatrix} = 0.$$

La prima uguaglianza è verificata per ogni valore di a, b, c ; la seconda è verificata se e solo se $a + b - c = 0$. Dunque:

- il sistema è compatibile se e solo se $a + b - c = 0$.

In tal caso esso ammette ∞^1 soluzioni. \square

Esempio Discutere la compatibilità del seguente sistema al variare del parametro h :

$$S: \begin{cases} x + hy = 4 - h \\ hx + 4y = 4 \end{cases}$$

Soluzione. $A = \begin{pmatrix} 1 & h \\ h & 4 \end{pmatrix}$ e quindi $\det A = 4 - h^2$. Ora, se $\det A \neq 0$ allora $\text{rk}A = \text{rk}A' = 2$

(perché?). Dunque, se $h \neq 2$ e $h \neq -2$ il sistema è compatibile, e ammette un'unica soluzione. Discutiamo ora i casi $h = 2$ e $h = -2$. Una verifica diretta mostra che, se $h = 2$ si ha $\text{rk}A = \text{rk}A' = 1$ dunque il sistema è compatibile e ammette ∞^1 soluzioni; se $h = -2$ allora $\text{rk}A = 1$ mentre $\text{rk}A' = 2$ e il sistema è incompatibile. In conclusione:

$$\begin{cases} \text{se } h \neq 2 \text{ e } h \neq -2 \text{ si ha un'unica soluzione,} \\ \text{se } h = 2 \text{ si hanno } \infty^1 \text{ soluzioni,} \\ \text{se } h = -2 \text{ il sistema è incompatibile.} \end{cases}$$

Esercizio Verificare il risultato con l'algoritmo di Gauss.

5.1 Equazioni significative

Il teorema seguente dice che, in taluni casi, è possibile scartare un certo numero di equazioni, senza alterare l'insieme delle soluzioni del sistema.

Teorema Sia S un sistema lineare compatibile, con matrice dei coefficienti A e matrice completa A' (quindi $\text{rk}A = \text{rk}A' = p$ per ipotesi). Sia B un minore di A avente ordine p e determinante non nullo. Allora il sistema S è equivalente al sistema ridotto SR che

si ottiene considerando solo le p equazioni corrispondenti al minore B , cioè $\text{Sol}(S) = \text{Sol}(SR)$.

Dimostrazione. Omessa.

In effetti, possiamo scartare le equazioni che non corrispondono al minore B fissato, poichè queste sono conseguenza delle equazioni del sistema ridotto. Il sistema ridotto è quindi formato dalle equazioni *significative* di S . Se il rango è basso (rispetto al numero di righe), possiamo scartare molte equazioni; se il rango è massimo, nessuna: sono tutte significative.

Esempio Consideriamo il sistema:

$$S : \begin{cases} 2x + 2y + 3z = -1 \\ x + y - z = 2 \\ x + y + 4z = -3 \\ 3x + 3y + 2z = 1 \end{cases}$$

Si verifica che $\text{rk}A = \text{rk}A' = 2$, dunque S è compatibile (e ammette ∞^1 soluzioni). Ora

$$A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 1 & 4 \\ 3 & 3 & 2 \end{pmatrix},$$

e il minore di ordine due dato da $B = \mu_{12,23} = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$ ha determinante non nullo. Dunque il sistema ridotto è quello formato dalle prime due equazioni:

$$SR : \begin{cases} 2x + 2y + 3z = -1 \\ x + y - z = 2 \end{cases},$$

e si ha $\text{Sol}(S) = \text{Sol}(SR)$. Possiamo dunque scartare le ultime due equazioni: in effetti, si vede subito che la terza equazione è la differenza delle prime due, mentre la quarta ne è la somma.

Infine, osserviamo che il sistema ridotto non è unico: potevamo scartare, ad esempio, le prime due equazioni, scegliendo il minore $B' = \mu_{34,23} = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$. In ogni caso, le equazioni significative sono esattamente due.

6 Sistemi omogenei

- Un sistema lineare si dice *omogeneo* se i suoi termini noti sono tutti nulli.

I sistemi omogenei formano una classe molto importante di sistemi lineari. In forma matriciale, un sistema omogeneo si scrive:

$$S : AX = O,$$

dove A è la matrice dei coefficienti e $O = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ è il *vettore colonna nullo*. È evidente che

la matrice completa A' di un sistema omogeneo si ottiene aggiungendo ad A una colonna nulla.

Un sistema omogeneo è sempre compatibile, poiché ammette sempre la soluzione nulla (o *banale*) $X = O$, ottenuta uguagliando tutte le variabili a zero. Una soluzione non nulla è detta anche *autosoluzione*.

Esempio Il sistema $S : \begin{cases} x + y + z = 0 \\ 2x - y + 5z = 0 \\ x + 2z = 0 \end{cases}$ è omogeneo, con forma matriciale $AX = O$,

dove $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 5 \\ 1 & 0 & 2 \end{pmatrix}$ e $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Una verifica diretta mostra che il vettore $X =$

$\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ è una soluzione non nulla, dunque un'autosoluzione del sistema.

Il problema è ora quello di stabilire quando un sistema omogeneo ammette autosoluzioni. Sappiamo che, per un sistema lineare qualunque S , si hanno solamente tre possibilità:

1. S è incompatibile,
2. S ammette un' unica soluzione,
3. S ammette infinite soluzioni.

Per un sistema omogeneo la prima alternativa non si realizza mai; dunque o S ammette un'unica soluzione (necessariamente quella nulla) oppure S ammette infinite soluzioni (in particolare, almeno un'autosoluzione). In conclusione:

- Un sistema lineare omogeneo ammette autosoluzioni se e solo se esso ammette infinite soluzioni.

Vediamo ora qualche criterio per l'esistenza di autosoluzioni.

Proposizione Sia $S : AX = O$ un sistema lineare omogeneo di m equazioni in n incognite. Allora S ammette autosoluzioni se e solo se il rango di A (matrice dei coefficienti) è minore di n (numero delle incognite).

Dimostrazione. È una conseguenza immediata del teorema di Rouché-Capelli. Infatti, poiché $\text{rk}A < n$ il sistema ammetterà infinite soluzioni (precisamente $\infty^{n-\text{rk}A}$ soluzioni) dunque almeno un'autosoluzione per quanto già detto. \square

Corollario Sia $S : AX = O$ un sistema lineare omogeneo quadrato (cioè, di n equazioni in n incognite). Allora S ammette autosoluzioni se e solo se $\det A = 0$.

Dimostrazione. La matrice dei coefficienti A è quadrata, e dunque il rango di A è minore di n precisamente quando il determinante di A è nullo. \square

Corollario Sia $S : AX = O$ un sistema lineare omogeneo di m equazioni in n incognite. Se $m < n$, se cioè il numero delle equazioni è minore del numero delle incognite, allora S ammette autosoluzioni.

Dimostrazione. Poiché A è $m \times n$, abbiamo che $\text{rk}A \leq \min\{m, n\}$, in particolare $\text{rk}A \leq m$. Ma per ipotesi $m < n$ dunque $\text{rk}A < n$ e il sistema ammette autosoluzioni grazie alla proposizione appena dimostrata. \square

Esempio Consideriamo il sistema lineare omogeneo $S : \begin{cases} x + y + z = 0 \\ 2x - y + 5z = 0 \\ x + 2z = 0 \end{cases}$ (quello di

prima). Un calcolo mostra che $\det A = 0$: per il primo corollario, S ammette autosoluzioni (cosa che avevamo già verificato). Calcoliamo l'insieme delle soluzioni. Riducendo la ma-

trice completa $A' = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & -1 & 5 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix}$ ad una forma a scalini otteniamo:

$$\tilde{A}' = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

dunque abbiamo ∞^1 soluzioni, del tipo

$$\text{Sol}(S) = \left\{ \begin{pmatrix} -2t \\ t \\ t \end{pmatrix} : t \in \mathbf{R} \right\}.$$

Osserviamo che la soluzione generica si scrive

$$\begin{pmatrix} -2t \\ t \\ t \end{pmatrix} = t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix},$$

ed è quindi un multiplo dell'autosoluzione $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$.

Esercizio È dato un sistema lineare omogeneo $S : AX = O$. Dimostrare che:

1. Date due soluzioni X_1 e X_2 anche la somma $X_1 + X_2$ è una soluzione.
2. Data una soluzione X e un numero $k \in \mathbf{R}$ anche kX è una soluzione.

Tutto questo si esprime dicendo che

- l'insieme delle soluzioni di un sistema lineare omogeneo è *chiuso* rispetto alla somma e alla moltiplicazione per uno scalare.

Esempio Stabilire se il sistema omogeneo $S : \begin{cases} x + y + z = 0 \\ x + 2y + 3z = 0 \\ 2x - y - 3z = 0 \end{cases}$ ammette autosoluzioni.

Soluzione. Il determinante della matrice dei coefficienti $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & -3 \end{pmatrix}$ vale 1: è diverso da zero, dunque S ammette solamente la soluzione nulla. \square

7 Dipendenza e indipendenza lineare di vettori di \mathbf{R}^n

7.1 Combinazioni lineari

Ricordiamo che i vettori di \mathbf{R}^n possono essere sommati e moltiplicati per uno scalare. Dati k vettori v_1, \dots, v_k di \mathbf{R}^n e k scalari $a_1, \dots, a_k \in \mathbf{R}$ il vettore:

$$a_1 v_1 + a_2 v_2 + \dots + a_k v_k$$

è detto la *combinazione lineare* di v_1, \dots, v_k a coefficienti a_1, \dots, a_k .

Esempio In \mathbf{R}^2 consideriamo $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$, $v_3 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$. Allora:

$$2v_1 + v_2 - 3v_3 = \begin{pmatrix} -13 \\ 6 \end{pmatrix}.$$

Esempio Scrivere, se possibile, il vettore $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, come combinazione lineare dei vettori $v_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$.

Soluzione. Risulta $v = 2v_1 - v_2$, quindi è possibile. \square

7.2 Dipendenza e indipendenza lineare

Una *relazione di dipendenza lineare* tra un certo numero di vettori v_1, \dots, v_k è un'espressione del tipo:

$$a_1v_1 + \dots + a_kv_k = O,$$

con la condizione che *almeno uno* dei coefficienti sia non nullo.

Quindi, nel nostro linguaggio, l'identità che si ottiene ponendo tutti i coefficienti uguali a zero:

$$0v_1 + \dots + 0v_k = O$$

non è una relazione di dipendenza lineare.

Esempio $2v_1 + 4v_2 - 5v_4 = O$ è una relazione di dipendenza lineare tra i vettori v_1, v_2, v_3, v_4 (si noti che non tutti i coefficienti devono essere non nulli... ne basta uno).

Notiamo che, se esiste una relazione di dipendenza lineare, allora almeno uno dei vettori sarà combinazione lineare degli altri. Nell'esempio precedente, possiamo risolvere rispetto a v_1 e osservare che:

$$v_1 = -2v_2 + \frac{5}{2}v_4.$$

Definizione a) I vettori v_1, \dots, v_k di \mathbf{R}^n si dicono linearmente dipendenti se esiste una relazione di dipendenza lineare tra di essi; se cioè esiste una combinazione lineare:

$$a_1v_1 + \dots + a_kv_k = O,$$

con almeno un coefficiente non nullo.

b) I vettori v_1, \dots, v_k si dicono linearmente indipendenti se non c'è alcuna relazione di dipendenza lineare tra di essi, se cioè l'uguaglianza

$$a_1v_1 + \dots + a_kv_k = O$$

è verificata solo quando tutti i coefficienti sono nulli: $a_1 = \dots = a_n = 0$.

- È facile verificare che i vettori v_1, \dots, v_k sono linearmente dipendenti se e solo se almeno uno di essi è combinazione lineare degli altri.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$. Notiamo che $v_2 = 3v_1$ quindi i due vettori sono legati dalla relazione di dipendenza lineare $-3v_1 + v_2 = 0$: ne segue che v_1 e v_2 sono linearmente dipendenti.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Vediamo se è possibile trovare una relazione di dipendenza lineare tra di essi. Imponiamo

$$xv_1 + yv_2 = 0,$$

con coefficienti $x, y \in \mathbf{R}$ al momento incogniti. Sviluppando l'identità, otteniamo l'equazione vettoriale:

$$\begin{pmatrix} x + y \\ 2x + 3y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

che dà luogo al sistema omogeneo:

$$\begin{cases} x + y = 0 \\ 2x + 3y = 0 \end{cases}.$$

Si vede subito che l'unica soluzione è quella nulla: $x = y = 0$. Dunque la relazione di dipendenza lineare cercata non esiste, e i due vettori sono linearmente indipendenti.

8 Criterio del rango

È possibile fornire un criterio per decidere se un'insieme di vettori di \mathbf{R}^n è linearmente indipendente oppure no. Tale criterio utilizza il rango di un'opportuna matrice.

Esempio Stabilire se i vettori $v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}$ sono linearmente indipendenti oppure no.

Soluzione. Cerchiamo una relazione di dipendenza lineare $x_1v_1 + x_2v_2 + x_3v_3 = 0$. Esplicitamente, stiamo cercando numeri reali x_1, x_2, x_3 tali che:

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tale equazione vettoriale si traduce nel sistema lineare omogeneo, di 4 equazioni in 3 incognite:

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

Notiamo che la forma matriciale del sistema è $AX = O$, con matrice dei coefficienti

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}. \text{ Notiamo anche che la matrice } A \text{ si ottiene incolonnando i vettori}$$

considerati.

Per definizione, i tre vettori saranno linearmente indipendenti se e solo il sistema S ammette solamente la soluzione nulla $x_1 = x_2 = x_3 = 0$. Dalla teoria dei sistemi omogenei ciò accade se e solo $\text{rk}A = n$, dove n è il numero delle incognite: nel nostro caso se e solo se $\text{rk}A = 3$. Un calcolo mostra che in effetti $\text{rk}A = 3$ (considerare il minore formato dalle prime tre righe). In conclusione, i vettori sono linearmente indipendenti. \square

Il procedimento è valido in generale. Dati k vettori *colonna* di \mathbf{R}^n , diciamo v_1, \dots, v_k , denotiamo con

$$\text{Mat}(v_1, \dots, v_k)$$

la matrice (di tipo $n \times k$) ottenuta incolonnando, nell'ordine, i vettori dati.

Ad esempio, se v_1, v_2, v_3 sono i vettori di \mathbf{R}^4 dell'esempio precedente, allora:

$$\text{Mat}(v_1, v_2, v_3) = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Una verifica mostra che l'equazione vettoriale

$$x_1 v_1 + \dots + x_k v_k = O$$

equivale al sistema omogeneo S di n equazioni in k incognite la cui forma matriciale è:

$$S : AX = O$$

dove $A = \text{Mat}(v_1, \dots, v_k)$ e $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Dal teorema di Rouche'-Capelli sappiamo che

S ammette soluzioni non nulle se e solo se $\text{rk}A < k$ (il numero delle incognite). Quindi

i vettori saranno linearmente dipendenti se e solo se $\text{rk}A < k$ e saranno linearmente indipendenti se e solo se $\text{rk}A = k$. In conclusione, abbiamo dimostrato la seguente

Proposizione *I vettori v_1, \dots, v_k dello spazio vettoriale \mathbf{R}^n sono linearmente indipendenti se e solo se*

$$\text{rk}A = k$$

dove $A = \text{Mat}(v_1, \dots, v_k)$ è la matrice ottenuta incolonnando i vettori dati.

Corollario *n vettori v_1, \dots, v_n di \mathbf{R}^n sono linearmente indipendenti se e solo se il determinante della matrice $\text{Mat}(v_1, \dots, v_n)$ è diverso da zero.*

Esempio Stabilire se i vettori $v_1 = \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}$ di \mathbf{R}^3 sono linearmente indipendenti oppure no.

Soluzione. Incolonniamo i vettori e otteniamo la matrice

$$A = \text{Mat}(v_1, v_2, v_3) = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 0 & 1 \\ 3 & 2 & -3 \end{pmatrix}.$$

Un calcolo mostra che $\det A = 0$ e il rango è minore di 3 (il numero dei vettori). Dunque i vettori sono linearmente dipendenti.

A questo punto sappiamo che esiste una relazione di dipendenza lineare che lega i tre vettori: cerchiamo tale relazione. Risolvendo il sistema

$$x_1 v_1 + x_2 v_2 + x_3 v_3 = O,$$

cioè il sistema omogeneo $S : AX = O$, otteniamo ∞^1 soluzioni:

$$\begin{cases} x_1 = 2t \\ x_2 = 3t \\ x_3 = 4t \end{cases}$$

con $t \in \mathbf{R}$. Prendendo ad esempio $t = 1$ vediamo che:

$$2v_1 + 3v_2 + 4v_3 = 0,$$

che è la relazione cercata.

8.1 Indipendenza lineare di righe e colonne

Da quanto detto abbiamo la seguente importante interpretazione del rango di una matrice. Una matrice A di tipo $n \times k$ individua n vettori riga di \mathbf{R}^k e k vettori colonna di \mathbf{R}^n .

Proposizione Sia A una matrice $n \times k$. Allora:

- a) I vettori colonna di A sono linearmente indipendenti se e solo se $\text{rk}A = k$ (il numero dei vettori colonna).
- b) I vettori riga di A sono linearmente indipendenti se e solo se $\text{rk}A = n$ (il numero dei vettori riga).

In particolare, se A è quadrata $n \times n$, allora gli n vettori riga (o colonna) di A sono linearmente indipendenti se e solo se $\det A \neq 0$.

Dimostrazione. a) Basta osservare che, se v_1, \dots, v_k sono i vettori colonna di A allora per definizione $A = \text{Mat}(v_1, \dots, v_k)$. Dunque a) è immediata dalla proposizione precedente. La parte b) si ottiene scambiando le righe e le colonne e ricordando che $\text{rk}A = \text{rk}A^t$. \square

Esempio La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ ha rango pari a 3. Dunque i tre vettori colonna (appartenenti allo spazio \mathbf{R}^4):

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

sono linearmente indipendenti, mentre i quattro vettori riga (che appartengono allo spazio \mathbf{R}^3)

$$w_1 = (1, 1, 0), \quad w_2 = (2, 3, 1), \quad w_3 = (1, 1, 1), \quad w_4 = (1, 0, -1)$$

sono linearmente dipendenti.

Osserviamo infine la seguente

Proposizione Siano v_1, \dots, v_k vettori di \mathbf{R}^n . Se $k > n$ tali vettori risultano sempre linearmente dipendenti.

Dimostrazione. La matrice che si ottiene incolonnando i vettori, cioè $A = \text{Mat}(v_1, \dots, v_k)$, è di tipo $n \times k$, dunque $\text{rk}A \leq \min\{n, k\} \leq n$. Per ipotesi $n < k$ dunque

$$\text{rk}A < k.$$

Per quanto detto in precedenza, i vettori saranno linearmente dipendenti. \square

Si avrà dunque che $n + 1$ (o piú) vettori di \mathbf{R}^n saranno sempre linearmente dipendenti.

Esempio I vettori $v_1 = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$, $v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} -1 \\ 11 \end{pmatrix}$ sono linearmente dipendenti, semplicemente perché i vettori sono tre, e appartengono a \mathbf{R}^2 . Trovare una relazione di dipendenza lineare, ed esprimere il terzo come combinazione lineare degli altri due.

Parte 4. Spazi vettoriali

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Spazi vettoriali, 1
- 2 Prime proprietà, 3
- 3 Dipendenza e indipendenza lineare, 4
- 4 Generatori, 6
- 5 Basi, 8
- 6 Sottospazi, 11
- 7 Teorema di esistenza di una base, 17
- 8 Dimensione, 17
- 9 Le basi di \mathbf{R}^n , 21
- 10 Spazi vettoriali di matrici, 23
- 11 Spazi vettoriali di polinomi, 24

1 Spazi vettoriali

Definiamo ora una nuova struttura algebrica: quella di spazio vettoriale. Uno spazio vettoriale è un insieme dotato di due operazioni (somma e prodotto per uno scalare) che verificano un certo numero di assiomi. Esempi di spazi vettoriali sono dati da \mathbf{R}^n e dall'insieme delle matrici $\mathbf{Mat}(m \times n)$. Ma ci sono infiniti esempi di spazi vettoriali.

1.1 Definizione

Sia V un insieme non vuoto, i cui elementi saranno chiamati *vettori*, dotato di due operazioni:

- (a) la *somma*, che associa a due vettori $u, v \in V$ un terzo vettore $u + v \in V$,
- (b) il *prodotto di un vettore per uno scalare*, che associa al vettore $v \in V$ e allo scalare $k \in \mathbf{R}$ un altro vettore denotato kv .

V si dice uno *spazio vettoriale reale* se le suddette operazioni verificano le seguenti proprietà:

- 1) $(u + v) + w = u + (v + w)$ per ogni $u, v, w \in V$.

- 2) $u + v = v + u$ per ogni $u, v \in V$.
 - 3) Esiste un vettore $O \in V$, detto *vettore nullo*, tale che $v + O = v$ per ogni $v \in V$.
 - 4) Per ogni vettore $v \in V$ esiste un'unico vettore $-v$, detto *opposto* di v , tale che $v + (-v) = -v + v = O$.
 - 5) Si ha $1v = v$ per ogni $v \in V$.
 - 6) $h(kv) = (hk)v$ per ogni $h, k \in \mathbf{R}$ e per ogni $v \in V$.
 - 7) $(h + k)v = hv + kv$ per ogni $h, k \in \mathbf{R}$ e per ogni $v \in V$.
 - 8) $h(u + v) = hu + hv$ per ogni $h \in \mathbf{R}$ e per ogni $u, v \in V$.
- Le proprietà 1)-8) sono anche dette *assiomi di spazio vettoriale*.

1.2 Esempio: lo spazio vettoriale delle matrici

Denotiamo con $\mathbf{Mat}(m \times n)$ l'insieme delle matrici $m \times n$ (a elementi reali). Abbiamo già introdotto nella Parte 1 la somma di due matrici e il prodotto di una matrice per uno scalare (numero). Per concludere che $\mathbf{Mat}(m \times n)$ è uno spazio vettoriale, occorre solamente verificare che tali operazioni verificano gli assiomi 1)-8). Tali assiomi risultano, in effetti, verificati, e dunque

- $\mathbf{Mat}(m \times n)$ è uno spazio vettoriale.

Notiamo che il *vettore nullo* di $\mathbf{Mat}(m \times n)$ è la matrice nulla O , quella avente tutte le entrate nulle.

1.3 Esempio: lo spazio vettoriale \mathbf{R}^n

Abbiamo definito \mathbf{R}^n come l'insieme dei vettori colonna ad n entrate:

$$\mathbf{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbf{R} \right\}$$

Poiché un vettore colonna è anche una matrice $n \times 1$, possiamo identificare \mathbf{R}^n con lo spazio vettoriale $\mathbf{Mat}(n \times 1)$. Somma e prodotto per uno scalare sono definiti, rispettivamente:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad k \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} kx_1 \\ \vdots \\ kx_n \end{pmatrix}.$$

Notiamo che il vettore nullo di \mathbf{R}^n è $O = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. In conclusione:

- \mathbf{R}^n è uno spazio vettoriale.

A volte sarà utile scrivere gli n numeri reali orizzontalmente, come un vettore riga: (x_1, \dots, x_n) e identificheremo dunque \mathbf{R}^n con lo spazio vettoriale $\mathbf{Mat}(1 \times n)$. Quando sarà importante distinguere i due modi di scrittura, specificheremo con la dicitura *vettore colonna di \mathbf{R}^n* , o *vettore riga di \mathbf{R}^n* .

1.4 Esempio: lo spazio vettoriale dei polinomi

Un polinomio nell'indeterminata x è un'espressione del tipo:

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n.$$

Se $a_n \neq 0$, diremo che n è il *grado* di $p(x)$. Un polinomio si dice *costante* se ha grado zero, ad esempio $p(x) = -3$.

L'insieme di tutti i polinomi (quindi, di grado arbitrario) si denota con $\mathbf{R}[x]$.

I polinomi possono essere sommati fra loro (il risultato è un polinomio) e moltiplicati per un numero, nella maniera usuale. Si verifica anche in questo caso che gli assiomi 1)-8) di spazio vettoriale sono soddisfatti. Dunque:

- $\mathbf{R}[x]$ è uno spazio vettoriale.

Notiamo che il vettore nullo di $\mathbf{R}[x]$ è il *polinomio nullo*, quello che ha tutti i coefficienti nulli $p(x) = 0$.

2 Prime proprietà

In ciò che segue, V è uno spazio vettoriale arbitrario, quindi non necessariamente uno degli esempi precedenti. Vogliamo osservare alcune proprietà comuni a tutti gli spazi vettoriali. Per dimostrare queste proprietà, dobbiamo usare *solamente* gli assiomi 1)-8) che definiscono uno spazio vettoriale.

Iniziamo dalla cosiddetta *legge di cancellazione della somma*.

Proposizione *Siano u, v, w vettori dello spazio vettoriale V . Allora si ha*

$$u + v = u + w \quad \text{se e solo se} \quad v = w.$$

In particolare $u + v = u$ se e solo se $v = O$.

Dimostrazione. Supponiamo che $u + v = u + w$. Dobbiamo dimostrare che $v = w$. Sommando ad ambo i membri l'opposto di u otteniamo:

$$-u + (u + v) = -u + (u + w).$$

Dalla proprietà associativa segue $(-u + u) + v = (-u + u) + w$ e quindi $O + v = O + w$.
Dalla proprietà che definisce il vettore nullo concludiamo che

$$v = w.$$

Il viceversa è immediato. \square

Anche le seguenti proprietà si dimostrano facilmente dagli assiomi.

Proposizione *In uno spazio vettoriale V valgono le seguenti proprietà:*

- a) *Per ogni $v \in V$ si ha $0v = O$.*
- b) *Per ogni $h \in \mathbf{R}$ si ha $hO = O$.*
- c) *Per ogni $v \in V$ si ha $(-1)v = -v$.*

Dalle proposizioni precedenti segue che, se $u + v = w$ allora $u = w - v$.
Infine osserviamo la seguente

Proposizione *Siano $h \in \mathbf{R}$ e $v \in V$. Se $hv = O$ allora $h = 0$ oppure $v = O$.*

Dimostrazione. Supponiamo $hv = O$. Se $h = 0$ abbiamo finito. Se $h \neq 0$, moltiplicando ambo i membri per h^{-1} otteniamo:

$$h^{-1}(hv) = h^{-1}O = O.$$

Ma, per le proprietà 5) e 6) abbiamo $h^{-1}(hv) = (h^{-1}h)v = 1v = v$ dunque $v = O$. \square

3 Dipendenza e indipendenza lineare

3.1 Combinazioni lineari

Avendo a disposizione la somma e il prodotto per uno scalare, possiamo definire combinazioni lineari di vettori, come abbiamo già visto in \mathbf{R}^n o in $\mathbf{Mat}(m \times n)$.

Dati k vettori v_1, \dots, v_k di uno spazio vettoriale V e k scalari $a_1, \dots, a_k \in \mathbf{R}$ il vettore:

$$a_1v_1 + a_2v_2 + \dots + a_kv_k$$

è detto la *combinazione lineare* di v_1, \dots, v_k a coefficienti a_1, \dots, a_k .

Esempio Nello spazio vettoriale dei polinomi, $\mathbf{R}[x]$, siano $p_1(x) = 1 - x + 3x^3$ e $p_2(x) = 4x + x^2$. Allora

$$4p_1(x) - 3p_2(x) = 4 - 3x^2 + 12x^3.$$

3.2 Dipendenza e indipendenza lineare

La definizione di indipendenza lineare di vettori, già vista in \mathbf{R}^n , si generalizza immediatamente al caso di uno spazio vettoriale qualunque.

Definizione a) I vettori v_1, \dots, v_k di uno spazio vettoriale V si dicono linearmente dipendenti se esiste una relazione di dipendenza lineare tra di essi; se cioè esiste una combinazione lineare:

$$a_1v_1 + \dots + a_kv_k = O,$$

con almeno un coefficiente non-nullo.

b) I vettori v_1, \dots, v_k si dicono linearmente indipendenti se non c'è alcuna relazione di dipendenza lineare tra di essi, se cioè l'uguaglianza

$$a_1v_1 + \dots + a_kv_k = O$$

è verificata solo quando tutti i coefficienti sono nulli.

Osserviamo il seguente fatto.

Proposizione I vettori v_1, \dots, v_k sono linearmente dipendenti se e solo se almeno uno di essi è combinazione lineare degli altri.

Dimostrazione. Supponiamo in primo luogo che v_1, \dots, v_k siano linearmente dipendenti: dunque esiste una relazione

$$a_1v_1 + \dots + a_kv_k = O,$$

con almeno uno dei coefficienti, diciamo a_i , diverso da zero. Possiamo risolvere rispetto a v_i , dividendo per a_i e portando tutto il resto a secondo membro:

$$v_i = -\frac{a_1}{a_i}v_1 - \dots - \frac{a_{i-1}}{a_i}v_{i-1} - \frac{a_{i+1}}{a_i}v_{i+1} - \dots - \frac{a_k}{a_i}v_k.$$

Dunque v_i risulta combinazione lineare dei rimanenti vettori. Viceversa, supponiamo che uno dei vettori, ad esempio il primo, sia combinazione lineare degli altri:

$$v_1 = b_2v_2 + \dots + b_kv_k.$$

Ma allora otteniamo la relazione di dipendenza lineare $v_1 - b_2v_2 - \dots - b_kv_k = O$ (si noti che il coefficiente di v_1 è 1, dunque non nullo) e quindi i vettori sono linearmente dipendenti. \square

• Osserviamo infine che v_1, \dots, v_k sono linearmente indipendenti se e solo se *nessuno* di essi è combinazione lineare degli altri.

3.3 Semplici conseguenze della definizione

Osserviamo che:

- Un singolo vettore v è linearmente indipendente se e solo se $v \neq O$.
- Se uno dei vettori v_1, \dots, v_k è nullo, allora v_1, \dots, v_k sono linearmente dipendenti. Infatti, se ad esempio $v_k = O$, otteniamo la relazione di dipendenza lineare

$$0v_1 + \dots + 0v_{k-1} + 1v_k = 0.$$

- Due vettori v_1, v_2 sono linearmente dipendenti se e solo se uno di essi è un multiplo dell'altro, se cioè $v_2 = kv_1$ per $k \in \mathbf{R}$ oppure $v_1 = hv_2$ con $h \in \mathbf{R}$. Diremo in tal caso che i due vettori sono *proporzionali*.

Esempio I vettori $v_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$ di \mathbf{R}^2 sono proporzionali, dunque linearmente dipendenti.

Occorre notare che la proprietà di indipendenza lineare si riferisce all'*insieme* dei vettori considerati. Sarebbe forse più preciso dire (e spesso faremo così) che *l'insieme di vettori* $\{v_1, \dots, v_k\}$ è *linearmente indipendente*.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. I tre vettori v_1, v_2, v_3 sono linearmente dipendenti, poiché $v_3 = 2v_1 + 3v_2$ (quindi l'insieme $\{v_1, v_2, v_3\}$ è linearmente dipendente) ma le tre coppie di vettori $\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}$ sono tutte linearmente indipendenti, come si verifica facilmente.

L'esempio mostra anche che, aggiungendo uno o più vettori a un'insieme linearmente indipendente, potremmo perdere questa proprietà.

Osserviamo però che la proprietà di indipendenza lineare si conserva *togliendo* un qualunque numero di vettori, nel senso della seguente proposizione.

Proposizione Sia $\mathcal{A} = \{v_1, \dots, v_k\}$ un insieme linearmente indipendente. Allora ogni sottoinsieme di \mathcal{A} è linearmente indipendente.

Dimostrazione. Esercizio. \square

4 Generatori

Definizione Diremo che i vettori v_1, \dots, v_k di uno spazio vettoriale V generano V se ogni vettore di V si può scrivere come combinazione lineare di v_1, \dots, v_k . Diremo anche che $\{v_1, \dots, v_k\}$ è un insieme di generatori di V .

Definizione Lo spazio vettoriale V si dice finitamente generato se ammette un insieme finito di generatori.

Esempio In \mathbf{R}^2 , consideriamo i vettori

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

È immediato vedere che $\{e_1, e_2\}$ è un insieme di generatori di \mathbf{R}^2 : infatti il vettore generico $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2$ si esprime:

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = xe_1 + ye_2.$$

In particolare, \mathbf{R}^2 è finitamente generato.

- e_1, e_2 sono detti i vettori della *base canonica* di \mathbf{R}^2 .

Possiamo generalizzare la precedente osservazione.

Esempio Nello spazio vettoriale \mathbf{R}^n consideriamo i vettori:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

detti i vettori della *base canonica* di \mathbf{R}^n . È evidente che essi generano \mathbf{R}^n , perché il vettore generico di \mathbf{R}^n si scrive:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n.$$

Dunque \mathbf{R}^n è finitamente generato.

Proposizione Aggiungendo vettori ad un insieme di generatori, otteniamo ancora un insieme di generatori.

Dimostrazione. Supponiamo che $\{v_1, \dots, v_k\}$ sia un insieme di generatori; dunque ogni vettore di V si scrive per ipotesi come combinazione lineare di v_1, \dots, v_k . Aggiungiamo

ora un vettore v_{k+1} alla lista. È chiaro che i vettori v_1, \dots, v_k, v_{k+1} generano V perché possiamo scrivere

$$\begin{aligned} v &= a_1 v_1 + \dots + a_k v_k \\ &= a_1 v_1 + \dots + a_k v_k + 0 v_{k+1}, \end{aligned}$$

che dimostra l'asserto. \square

Esempio I vettori $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ generano \mathbf{R}^2 (i primi due vettori hanno già questa proprietà).

Attenzione perché *togliendo* vettori ad un insieme di generatori potremmo perdere tale proprietà.

5 Basi

Definizione Un insieme finito di vettori $\{v_1, \dots, v_k\}$ si dice una base di V se:

- a) è un insieme di generatori,
- b) è un insieme linearmente indipendente.

In altre parole, una base è un insieme di generatori formato da vettori linearmente indipendenti.

Esempio I vettori della base canonica di \mathbf{R}^n formano, appunto, una base di \mathbf{R}^n . Infatti abbiamo visto che essi generano \mathbf{R}^n ; essi sono anche linearmente indipendenti, perché

$$x_1 e_1 + \dots + x_n e_n = 0$$

solo quando $x_1 = \dots = x_n = 0$.

Esempio I vettori $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ generano \mathbf{R}^2 . Essi però non sono linearmente indipendenti, poichè $v = 2e_1 + 3e_2$. Dunque tali vettori non formano una base di \mathbf{R}^2 .

Esempio I vettori $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ sono linearmente indipendenti, ma non generano

\mathbf{R}^3 (infatti, ad esempio, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ non è combinazione lineare di e_1, e_2). Dunque, essi non formano una base di \mathbf{R}^3 .

La proprietà importante di una base è espressa nel seguente

Teorema *I vettori v_1, \dots, v_k formano una base di V se e solo se ogni vettore di V si scrive, in modo unico, come combinazione lineare di v_1, \dots, v_k .*

Dimostrazione. Prima parte. Supponiamo che v_1, \dots, v_k formino una base di V . Poiché i vettori generano V per ipotesi, potremo scrivere ogni vettore $v \in V$ nella forma:

$$v = a_1 v_1 + \dots + a_k v_k,$$

con $a_1, \dots, a_k \in \mathbf{R}$. Dimostriamo che i coefficienti a_1, \dots, a_k di tale combinazione lineare sono unici. Supponiamo infatti che v si possa esprimere in altro modo, diciamo:

$$v = b_1 v_1 + \dots + b_k v_k$$

Uguagliando le due espressioni, otterremo $a_1 v_1 + \dots + a_k v_k = b_1 v_1 + \dots + b_k v_k$ dunque:

$$(a_1 - b_1)v_1 + \dots + (a_k - b_k)v_k = O.$$

Poiché i vettori v_1, \dots, v_k sono linearmente indipendenti, dobbiamo avere necessariamente $a_i - b_i = 0$ per ogni i , dunque $a_i = b_i$ per ogni i . Questo dimostra che i coefficienti della combinazione lineare sono univocamente determinati dalla base scelta.

Seconda parte. Supponiamo ora che ogni vettore di V si scriva, in modo unico, come combinazione lineare di v_1, \dots, v_k . Dimostriamo che tali vettori formano una base. Ora, è immediato che essi generano V ; occorre solo dimostrare che essi sono linearmente indipendenti. Supponiamo allora che

$$a_1 v_1 + \dots + a_k v_k = O.$$

Si ha anche, banalmente:

$$0v_1 + \dots + 0v_k = O$$

e quindi, dalla proprietà di unicità, otteniamo $a_1 = \dots = a_k = 0$. Dunque i vettori v_1, \dots, v_k sono linearmente indipendenti. \square

Esempio Abbiamo visto che i vettori $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ generano \mathbf{R}^2 , ma non formano una base. Questo si poteva dimostrare anche osservando che, ad esempio, il vettore $w = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ è combinazione lineare di v_1, v_2, v_3 in almeno due modi diversi:

$$\begin{aligned} \begin{pmatrix} 2 \\ 4 \end{pmatrix} &= v_1 + v_2 + v_3, \\ \begin{pmatrix} 2 \\ 4 \end{pmatrix} &= 2v_1 + 4v_2 + 0v_3. \end{aligned}$$

Vedremo che uno spazio vettoriale finitamente generato ha diverse (in realtà, infinite!) basi.

5.1 Coordinate di un vettore rispetto a una base

Fissiamo ora una base $\{v_1, \dots, v_n\}$ di V . Includendo i vettori tra parentesi tonde:

$$(v_1, \dots, v_n)$$

intenderemo che i vettori sono ordinati nel modo indicato: v_1 , poi v_2 etc. Diremo anche che (v_1, \dots, v_n) è una *base ordinata* di V .

Dato un vettore $v \in V$, possiamo scrivere

$$v = a_1 v_1 + \dots + a_n v_n,$$

con coefficienti a_1, \dots, a_n univocamente determinati. Tali coefficienti sono detti *le coordinate di v rispetto alla base (v_1, \dots, v_n)* . Esprimeremo le coordinate del vettore v mediante il vettore colonna di \mathbf{R}^n :

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Esempio Le coordinate di un vettore $v = \begin{pmatrix} x \\ y \end{pmatrix}$ rispetto alla base canonica di \mathbf{R}^2 sono $\begin{pmatrix} x \\ y \end{pmatrix}$, il vettore stesso.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$.

- a) Dimostrare che v_1, v_2 formano una base di \mathbf{R}^2 .
- b) Calcolare le coordinate del vettore $w = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$ rispetto alla base (v_1, v_2) .

Soluzione. a) Dimostriamo che ogni vettore $v = \begin{pmatrix} a \\ b \end{pmatrix}$ di \mathbf{R}^2 si scrive in modo unico:

$$v = x v_1 + y v_2.$$

L'equazione si traduce nel sistema $S : \begin{cases} x + y = a \\ 3x + 4y = b \end{cases}$ con matrice dei coefficienti $A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}$ avente determinante diverso da zero. Dunque, per il teorema di Cramer, S ammette un'unica soluzione *per ogni* scelta di a, b e (v_1, v_2) è quindi una base di \mathbf{R}^2 .

b) Dobbiamo esprimere w come combinazione lineare

$$w = xv_1 + yv_2$$

(questo è possibile perché v_1, v_2 formano una base) e le coordinate di w saranno per definizione $X = \begin{pmatrix} x \\ y \end{pmatrix}$. Tali coordinate si trovano risolvendo il sistema della parte a) con $a = 1$ e $b = 6$. Otteniamo $x = -2, y = 3$ dunque $w = -2v_1 + 3v_2$ e le coordinate sono $X = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$. \square

- Osserviamo che le coordinate di un vettore *dipendono dalla base scelta*: nell'esempio precedente, le coordinate di $w = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$ sono:

$$\begin{pmatrix} -2 \\ 3 \end{pmatrix}, \quad \text{rispetto alla base } (v_1, v_2),$$

$$\begin{pmatrix} 1 \\ 6 \end{pmatrix}, \quad \text{rispetto alla base canonica } (e_1, e_2).$$

- Generalizzando l'argomento usato nel precedente esempio vediamo anche che \mathbf{R}^2 ha infinite basi diverse: una qualunque coppia di vettori $v_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, v_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$, tali che la matrice $\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ abbia determinante diverso da zero, formerà una base di \mathbf{R}^2 .

6 Sottospazi

Definizione Un sottoinsieme E di uno spazio vettoriale V si dice un sottospazio di V se verifica le seguenti proprietà.

- Il vettore nullo appartiene a E .
- Se $u, v \in E$ allora $u + v \in E$.
- Se $u \in E$ e $k \in \mathbf{R}$ allora $ku \in E$.

Le proprietà b) e c) si esprimono dicendo anche che un sottospazio è *chiuso rispetto alla somma e al prodotto per uno scalare*.

Esempi immediati di sottospazi sono dati dal *sottospazio nullo* $\{O\}$, costituito dal solo vettore nullo, e da tutto lo spazio V : questi sono i cosiddetti *sottospazi banali* di V .

Proposizione *Ogni sottospazio di uno spazio vettoriale V è esso stesso uno spazio vettoriale (rispetto alle operazioni di somma e prodotto per uno scalare definite in V).*

Dimostrazione. Gli assiomi di spazio vettoriale sono verificati per V , e quindi anche per E , perché le operazioni di somma e prodotto per uno scalare sono le stesse. \square

Abbiamo così molti altri esempi di spazi vettoriali.

6.1 Esempi di sottospazi di \mathbf{R}^n

Proposizione *Sia E l'insieme delle soluzioni di un sistema lineare omogeneo di m equazioni in n incognite. Allora E è un sottospazio di \mathbf{R}^n .*

Dimostrazione. Un sistema lineare omogeneo si scrive, in forma matriciale, come $AX = O$, dove A è la matrice dei coefficienti, di tipo $m \times n$, e X è il vettore colonna delle incognite. Quindi

$$E = \text{Sol}(S) = \{X \in \mathbf{R}^n : AX = O\}.$$

Verifichiamo le proprietà di chiusura. È evidente che il vettore nullo $O \in E$. Se $X_1, X_2 \in E$ allora

$$A(X_1 + X_2) = AX_1 + AX_2 = O + O = O,$$

dunque anche $X_1 + X_2 \in E$, e abbiamo la chiusura rispetto alla somma. La chiusura rispetto al prodotto per uno scalare si verifica in modo simile. \square

Esempio Consideriamo il sistema lineare omogeneo $S : \begin{cases} x + y - z = 0 \\ 2x - y + 4z = 0 \end{cases}$. Trovare una base di E , il sottospazio di \mathbf{R}^3 formato dalle soluzioni di S .

Soluzione. Dobbiamo innanzitutto risolvere il sistema. Le soluzioni si esprimono

$$E = \text{Sol}(S) = \left\{ \begin{pmatrix} -t \\ 2t \\ t \end{pmatrix} : t \in \mathbf{R} \right\}$$

e sono ∞^1 . Il vettore generico di E si scrive:

$$\begin{pmatrix} -t \\ 2t \\ t \end{pmatrix} = t \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix},$$

ed è un multiplo del vettore $\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$, che è non nullo, dunque linearmente indipendente.

In conclusione una base di E è formata dal vettore

$$\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Esercizio Si consideri il sistema omogeneo di due equazioni in quattro incognite

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

Trovare una base di $E = \text{Sol}(S)$.

Soluzione. Notiamo che E è un sottospazio di \mathbf{R}^4 . Risolvendo il sistema, otteniamo ∞^2 soluzioni:

$$E = \left\{ \begin{pmatrix} t \\ t \\ -2s \\ s \end{pmatrix} : t, s \in \mathbf{R} \right\}.$$

Il vettore generico di E si scrive:

$$\begin{pmatrix} t \\ t \\ -2s \\ s \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

Dunque i due vettori $w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $w_2 = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}$ generano E . Essi sono anche linearmente

indipendenti (come si vede subito) dunque w_1, w_2 formano una base di E . \square

- È evidente che, se S non è omogeneo, il suo insieme delle soluzioni *non* è un sottospazio, poiché il vettore nullo non appartiene a E .

6.2 Sottospazio generato da un insieme di vettori

Sia V uno spazio vettoriale e v_1, \dots, v_k vettori assegnati di V . Consideriamo l'insieme di tutte le combinazioni lineari di v_1, \dots, v_k :

$$L[v_1, \dots, v_k] \doteq \{a_1 v_1 + \dots + a_k v_k : a_1, \dots, a_k \in \mathbf{R}\}.$$

Si verifica facilmente che $L[v_1, \dots, v_k]$ è un sottospazio di V , detto il *sottospazio generato* da v_1, \dots, v_k .

Le seguenti affermazioni sono quindi equivalenti:

- 1) I vettori v_1, \dots, v_k generano V .
- 2) $V = L[v_1, \dots, v_k]$.

Esempio Siano $w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Descrivere il sottospazio E di \mathbf{R}^3 generato dai vettori w_1, w_2 , cioè

$$E = L[w_1, w_2].$$

Soluzione. Il vettore generico di E è combinazione lineare di w_1, w_2 . Dunque esso si scrive $v = tw_1 + sw_2$ con $t, s \in \mathbf{R}$. Esplicitamente:

$$v = \begin{pmatrix} t \\ 0 \\ s \end{pmatrix},$$

con $t, s \in \mathbf{R}$ arbitrari. A parole, E è formato da tutti i vettori di \mathbf{R}^3 aventi seconda entrata nulla. Notiamo che E è anche l'insieme delle soluzioni dell'equazione $x_2 = 0$ nelle variabili x_1, x_2, x_3 :

$$E = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbf{R}^3 : x_2 = 0 \right\}.$$

Quindi, ad esempio: $\begin{pmatrix} 4 \\ 0 \\ -1 \end{pmatrix} \in E$ ma $\begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} \notin E$. Notiamo infine che i generatori di E , cioè i vettori w_1, w_2 , sono linearmente indipendenti: dunque w_1, w_2 formano una base di E . \square

• È evidente che, se $E = L[v_1, \dots, v_k]$, e se i generatori v_1, \dots, v_k sono anche linearmente indipendenti, allora essi formano una base di E .

Esempio Siano $w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, w_3 = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$. Descrivere il sottospazio di \mathbf{R}^3 generato da w_1, w_2, w_3 , cioè $L[w_1, w_2, w_3]$.

Soluzione. $F = L[w_1, w_2, w_3]$ è formato dai vettori del tipo:

$$v = aw_1 + bw_2 + cw_3,$$

con $a, b, c \in \mathbf{R}$. Ponendo $c = 0$ vediamo che F contiene tutte le combinazioni lineari di w_1, w_2 , cioè

$$L[w_1, w_2] \subseteq L[w_1, w_2, w_3].$$

Osserviamo ora che $w_3 = 2w_1 - 3w_2$. Sostituendo nell'espressione di v , otteniamo che il vettore generico di F si scrive anche

$$v = (a + 2c)w_1 + (b - 3c)w_2,$$

e questo dimostra che

$$L[w_1, w_2, w_3] \subseteq L[w_1, w_2].$$

Per il principio della doppia inclusione, abbiamo allora

$$L[w_1, w_2, w_3] = L[w_1, w_2].$$

Dunque $L[w_1, w_2, w_3]$ coincide con il sottospazio dell'esempio precedente. \square

L'argomento usato nell'esempio motiva la seguente proposizione di carattere generale.

Proposizione Siano v_1, \dots, v_{k+1} vettori di un qualunque spazio vettoriale V . Allora si ha:

- a) $L[v_1, \dots, v_k] \subseteq L[v_1, \dots, v_k, v_{k+1}]$.
- b) $L[v_1, \dots, v_k] = L[v_1, \dots, v_k, v_{k+1}]$ se e solo se v_{k+1} è combinazione lineare di v_1, \dots, v_k .

Dimostrazione. a) È sufficiente osservare che $a_1v_1 + \dots + a_kv_k = a_1v_1 + \dots + a_kv_k + 0v_{k+1}$: quindi ogni combinazione lineare di v_1, \dots, v_k è anche combinazione lineare di v_1, \dots, v_{k+1} e la prima parte è dimostrata.

b) Supponiamo che v_{k+1} sia combinazione lineare di v_1, \dots, v_k . Questo implica che ogni combinazione lineare di v_1, \dots, v_k, v_{k+1} è anche combinazione lineare dei primi k vettori v_1, \dots, v_k : dunque

$$L[v_1, \dots, v_{k+1}] \subseteq L[v_1, \dots, v_k].$$

Dalla parte a) sappiamo però che $L[v_1, \dots, v_k] \subseteq L[v_1, \dots, v_{k+1}]$. Dunque, per il principio della doppia inclusione:

$$L[v_1, \dots, v_{k+1}] = L[v_1, \dots, v_k].$$

Ora dimostriamo il viceversa. Supponiamo che $L[v_1, \dots, v_{k+1}] = L[v_1, \dots, v_k]$: allora per ipotesi $v_{k+1} \in L[v_1, \dots, v_k]$ quindi v_{k+1} è una combinazione lineare di v_1, \dots, v_k . \square

In conclusione, supponiamo di aggiungere un vettore v_{k+1} a un insieme di generatori v_1, \dots, v_k . Allora il sottospazio generato

- rimane lo stesso, se e solo se v_{k+1} è combinazione lineare dei generatori precedenti;
- diventa più grande, se e solo se v_{k+1} non è combinazione lineare dei generatori precedenti.

In particolare:

- se un generatore è combinazione lineare degli altri, allora può essere rimosso dalla lista, senza alterare il sottospazio generato.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}$, $v_4 = \begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}$. Dimostrare che

- $L[v_1, v_2, v_3, v_4] = L[v_1, v_2]$.
- v_1, v_2 formano una base di $L[v_1, v_2, v_3, v_4]$.

Soluzione. Osserviamo che $v_4 = 3v_1$; dunque possiamo eliminare v_4 . A questo punto abbiamo:

$$L[v_1, v_2, v_3, v_4] = L[v_1, v_2, v_3].$$

Possiamo eliminare qualche altro generatore? Si ha: $v_3 = 2v_1 - v_2$ dunque possiamo eliminare anche v_3 . A questo punto

$$L[v_1, v_2, v_3, v_4] = L[v_1, v_2, v_3] = L[v_1, v_2].$$

Ora osserviamo che v_1, v_2 sono linearmente indipendenti: nessuno di essi è combinazione lineare (multiplo) dell'altro. Dunque non possiamo più scartare niente.

- Poiché i vettori v_1, v_2 sono linearmente indipendenti, e generano il sottospazio, essi formano una base. \square

- L'esempio precedente mostra come una base sia il modo più "economico" per generare un dato sottospazio: i generatori v_3 e v_4 sono inutili, e possono essere eliminati. I generatori rimasti, v_1 e v_2 sono linearmente indipendenti: nessuno di essi può essere scartato, perché altrimenti il sottospazio generato diventerebbe più piccolo.

7 Teorema di esistenza di una base

Il procedimento di "eliminazione dei generatori inutili" può essere applicato in ogni spazio vettoriale finitamente generato, e ci permetterà di ricavare una base da un qualunque insieme di generatori. Enunciamo quindi il cosiddetto *teorema di esistenza di una base*.

Teorema *Sia $V \neq \{O\}$ uno spazio vettoriale finitamente generato. Allora V ammette almeno una base.*

Dimostrazione. Per ipotesi V è generato da un certo numero (finito) di vettori: $V = L[v_1, \dots, v_k]$. Se i generatori sono linearmente indipendenti, allora essi formano una base e abbiamo finito. Altrimenti, almeno uno di essi è una combinazione lineare degli altri: supponiamo che tale vettore sia v_k (questo non lede la generalità). Possiamo dunque eliminarlo dalla lista, e $V = L[v_1, \dots, v_{k-1}]$. Se v_1, \dots, v_{k-1} sono linearmente indipendenti, essi formano una base. Altrimenti, possiamo scartare qualche altro generatore, e così via. Ora non possiamo scartare tutti i generatori, perché per ipotesi V contiene almeno un vettore non nullo. Dunque prima o poi ci dobbiamo fermare, e ci fermiamo esattamente quando i generatori rimasti sono linearmente indipendenti. Tali generatori formano la base cercata. \square

Il metodo di eliminazione mostra anche che

Corollario *Ogni insieme di generatori contiene una base.*

8 Dimensione

Uno spazio vettoriale ammette infinite basi diverse. In questa sezione dimostreremo che tutte le basi hanno lo stesso numero di vettori, ciò che ci permetterà di definire la dimensione di V come il numero di vettori di una sua base qualunque. In ciò che segue, V è uno spazio vettoriale finitamente generato.

Lemma *Supponiamo che i vettori v_1, \dots, v_m siano linearmente indipendenti, e i vettori w_1, \dots, w_n generino V . Allora si ha necessariamente $m \leq n$.*

Dimostrazione. La dimostrazione si fa per assurdo: faremo vedere che l'ipotesi $m > n$ porta a una contraddizione. In ciò che segue supporremo dunque:

$$m > n.$$

Poiché w_1, \dots, w_n generano V , possiamo esprimere i primi n vettori della lista v_1, \dots, v_m come combinazione lineare di w_1, \dots, w_n : esistono dunque degli scalari a_{ij} tali che

$$\begin{cases} v_1 = a_{11}w_1 + \dots + a_{1n}w_n \\ \dots \\ v_n = a_{n1}w_1 + \dots + a_{nn}w_n \end{cases}$$

Conviene esprimere queste relazioni nella forma matriciale seguente:

$$(v_1, \dots, v_n) = (w_1, \dots, w_n)A, \quad (1)$$

dove $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ e l'espressione $(w_1, \dots, w_n)A$ indica il prodotto del vettore

riga (w_1, \dots, w_n) (le cui entrate sono vettori) per la matrice A , che è quadrata $n \times n$.

Supponiamo che $\det A = 0$. Dal teorema di Rouché-Capelli, sappiamo che il sistema

omogeneo $AX = O$ ammette almeno una soluzione *non nulla* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Moltiplichiamo

ambo i membri di (??) per X (a destra) e otteniamo:

$$(v_1, \dots, v_n)X = (w_1, \dots, w_n)AX = O \quad (\text{il vettore nullo})$$

da cui

$$x_1v_1 + \dots + x_nv_n = O,$$

con x_1, \dots, x_n non tutti nulli: questo, però, è impossibile, perché per ipotesi i vettori v_1, \dots, v_n sono linearmente indipendenti.

Dunque $\det A \neq 0$ e la matrice A risulta invertibile.

Moltiplichiamo ambo i membri della (??) per A^{-1} , a destra. Otteniamo

$$(v_1, \dots, v_n)A^{-1} = (w_1, \dots, w_n).$$

Se $A^{-1} = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$ tale relazione si scrive esplicitamente

$$\begin{cases} w_1 = b_{11}v_1 + \dots + b_{n1}v_n \\ \dots \\ w_n = b_{1n}v_1 + \dots + b_{nn}v_n \end{cases}$$

e mostra che ogni vettore w_j è combinazione lineare di v_1, \dots, v_n . Per ipotesi i vettori w_1, \dots, w_n generano V . In particolare, il vettore v_{n+1} è combinazione lineare di w_1, \dots, w_n , quindi anche di v_1, \dots, v_n : ma questo è impossibile, perché i vettori v_1, \dots, v_{n+1} sono, per ipotesi, linearmente indipendenti.

In conclusione, abbiamo un assurdo in entrambi i casi: $\det A = 0, \det A \neq 0$. Tale assurdo è conseguenza dell'ipotesi $m > n$. Dunque $m \leq n$. \square

Proposizione *Tutte le basi di uno spazio vettoriale V hanno lo stesso numero di vettori.*

Dimostrazione. Siano $\mathcal{B} = \{v_1, \dots, v_m\}$ e $\mathcal{B}' = \{w_1, \dots, w_n\}$ due basi di V . Vogliamo dimostrare che

$$m = n.$$

Per ipotesi, i vettori v_1, \dots, v_m sono linearmente indipendenti e i vettori w_1, \dots, w_n generano V : dal lemma otteniamo $m \leq n$. D'altra parte è anche vero che i vettori v_1, \dots, v_m generano V , e i vettori w_1, \dots, w_n sono linearmente indipendenti, dunque sempre grazie al lemma otteniamo $m \geq n$. La conclusione è che $m = n$. \square

Definizione *Sia V uno spazio vettoriale finitamente generato. Definiamo dimensione di V il numero di vettori di una qualunque base di V (tale numero è sempre lo stesso).*

Per calcolare la dimensione di uno spazio vettoriale, basta quindi:

- trovare una base di V ,
- contare i vettori che la compongono.

Teorema *Sia V uno spazio vettoriale di dimensione n . Allora:*

- a) *Non esistono più di n vettori linearmente indipendenti.*
- b) *Dati comunque k vettori, con $k < n$, essi non possono generare V .*
- c) *n vettori linearmente indipendenti sono anche generatori (quindi formano una base).*
- d) *n vettori generatori sono anche linearmente indipendenti (quindi formano una base).*

Dimostrazione. Per ipotesi, esiste una base (v_1, \dots, v_n) costituita da n vettori.

- a) Siano w_1, \dots, w_k vettori linearmente indipendenti; poichè per ipotesi v_1, \dots, v_n generano V si ha $k \leq n$ per il lemma precedente.
- b) Supponiamo che w_1, \dots, w_k generino V ; poichè per ipotesi v_1, \dots, v_n sono linearmente indipendenti allora $k \geq n$ per il lemma precedente.
- c) Supponiamo che i vettori w_1, \dots, w_n siano linearmente indipendenti: dobbiamo dimostrare che, dato un qualunque vettore $u \in V$, esso si esprime come combinazione

lineare di w_1, \dots, w_n . Ora i vettori u, w_1, \dots, w_n devono essere linearmente dipendenti per la parte a) del teorema: dunque esiste una combinazione lineare:

$$bu + b_1w_1 + \dots + b_nw_n = O,$$

con coefficienti b, b_1, \dots, b_n non tutti nulli. Ora, se $b = 0$, avremmo $b_1w_1 + \dots + b_nw_n = O$ quindi $b_1 = \dots = b_n = 0$: ma questo è impossibile per quanto appena detto. Dunque $b \neq 0$: dividendo per b e isolando u a primo membro possiamo esprimere u come combinazione lineare di w_1, \dots, w_n .

d) Supponiamo per assurdo che i vettori w_1, \dots, w_n generino V e non siano linearmente indipendenti. Dunque almeno uno di essi è combinazione lineare degli altri, e lo possiamo eliminare. Otteniamo in questo modo un insieme di $n - 1$ generatori, e questo contraddice la parte b) del teorema. \square

La parte a) dice che la dimensione è il *numero massimo* di vettori linearmente indipendenti. La parte b) dice che la dimensione è il *numero minimo* di generatori.

8.1 Teorema del completamento di una base

Il teorema seguente dice che possiamo sempre estendere un insieme di vettori linearmente indipendenti fino ad ottenere una base.

Teorema Siano v_1, \dots, v_k vettori linearmente indipendenti di uno spazio vettoriale V di dimensione n . Se $k < n$ allora possiamo trovare $n - k$ vettori w_1, \dots, w_{n-k} tali che gli n vettori

$$v_1, \dots, v_k, w_1, \dots, w_{n-k}$$

formano una base di V .

Il teorema è conseguenza del seguente risultato.

Lemma Supponiamo che i vettori v_1, \dots, v_k siano linearmente indipendenti, e sia E il sottospazio da essi generato: $E = L[v_1, \dots, v_k]$. Allora, se $w \notin E$, i $k + 1$ vettori

$$v_1, \dots, v_k, w$$

risultano linearmente indipendenti.

Dimostrazione. Supponiamo che

$$a_1v_1 + \dots + a_kv_k + bw = O.$$

Se $b \neq 0$, possiamo dividere ambo i membri per b ed esprimere w come combinazione lineare di v_1, \dots, v_k : questo però è impossibile, poiché altrimenti $w \in E$. Dunque $b = 0$: ma allora

$$a_1 v_1 + \dots + a_k v_k = 0$$

e poiché v_1, \dots, v_k sono linearmente indipendenti per ipotesi, si avrà $a_1 = \dots = a_k = 0$. \square

Dimostriamo ora il teorema.

Sia $E = L[v_1, \dots, v_k]$. Poiché $k < n$, i vettori v_1, \dots, v_k non possono generare tutto lo spazio vettoriale V (altrimenti essi formerebbero una base, e quindi $k = n$). Dunque possiamo trovare almeno un vettore $w_1 \notin E$. Per il lemma, i $k + 1$ vettori

$$v_1, \dots, v_k, w_1$$

risultano linearmente indipendenti. Ora, se $k + 1 = n$ essi formano una base per il teorema precedente (parte c), e abbiamo finito. Se $k + 1 < n$ possiamo ripetere l'argomentazione precedente e trovare almeno un vettore w_2 tale che i $k + 2$ vettori

$$v_1, \dots, v_k, w_1, w_2$$

risultano linearmente indipendenti. Continuando in questo modo arriviamo a un insieme di n vettori linearmente indipendenti $v_1, \dots, v_k, w_1, \dots, w_{n-k}$: per la parte c) del teorema della sezione precedente, tali vettori formano una base. \square

9 Le basi di \mathbf{R}^n

Sia ora $V = \mathbf{R}^n$. Abbiamo già osservato che gli n vettori:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

formano una base, detta *base canonica* di \mathbf{R}^n . Dunque:

- \mathbf{R}^n ha dimensione n .

Ci proponiamo ora di caratterizzare le (infinite) basi di \mathbf{R}^n .

Proposizione 1) *Tutte le basi di \mathbf{R}^n sono composte da n vettori.*

2) *I vettori v_1, \dots, v_n di \mathbf{R}^n formano una base se e solo la matrice di colonne v_1, \dots, v_n ha determinante non nullo.*

Dimostrazione. 1) è una conseguenza immediata del fatto che la dimensione di \mathbf{R}^n è n .

2) Se la matrice di colonne v_1, \dots, v_n ha determinante non nullo allora i vettori v_1, \dots, v_n sono linearmente indipendenti e quindi, poiché sono esattamente n , formano una base. Il viceversa si dimostra in modo analogo.

Esempio I vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ formano una base di \mathbf{R}^3 , poiché

$$\text{Mat}(v_1, v_2, v_3) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

ha determinante -1 , quindi non nullo.

Illustriamo il teorema di estensione di una base sul seguente esempio.

Esempio a) Verificare che i vettori $v_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$ sono linearmente indipendenti.

b) Estendere v_1, v_2 ad una base di \mathbf{R}^3 .

Soluzione. a) Basta osservare che il rango della matrice $A = \begin{pmatrix} 2 & 4 \\ 1 & 2 \\ 2 & 1 \end{pmatrix}$ vale 2.

b) Basta aggiungere un solo vettore $w_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ tale che la matrice

$$A' = \begin{pmatrix} 2 & 4 & a \\ 1 & 2 & b \\ 2 & 1 & c \end{pmatrix}$$

abbia rango 3 (cioè, determinante non nullo). Questo si può fare in infiniti modi diversi. Forse il modo più semplice è il seguente. Fissiamo un minore di ordine 2 di A con determinante non nullo, ad esempio $M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ e scegliamo $a = 1, b = 0, c = 0$: in questo modo si vede subito che

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

ha determinante non nullo. Dunque il vettore che possiamo aggiungere per ottenere una base di \mathbf{R}^3 è $w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. \square

10 Spazi vettoriali di matrici

Consideriamo ora lo spazio vettoriale delle matrici $m \times n$, che abbiamo denotato con il simbolo $\mathbf{Mat}(m \times n)$. Ricordiamo che il vettore nullo di $\mathbf{Mat}(m \times n)$ è la matrice nulla O .

Vogliamo dimostrare che $\mathbf{Mat}(m \times n)$ è finitamente generato, ed ha infatti dimensione pari a mn .

Fissiamo gli indici i, j e consideriamo la *matrice elementare* E_{ij} , cioè la matrice che ha 1 nell'entrata (i, j) e zero altrove. È evidente che le matrici elementari di tipo $m \times n$ sono, in numero, mn , e generano $\mathbf{Mat}(m \times n)$: infatti, se $A = (a_{ij})$ allora, per definizione

$$A = \sum_{i,j} a_{ij} E_{ij}$$

dove $i = 1, \dots, m$ e $j = 1, \dots, n$. Quindi le matrici elementari formano un insieme di generatori. Tali matrici sono anche linearmente indipendenti perché è evidente che

$$\sum_{i,j} a_{ij} E_{ij} = O \quad (\text{matrice nulla})$$

se e solo se $a_{ij} = 0$ per ogni i, j . Concludiamo che:

Proposizione *Le matrici elementari formano una base di $\mathbf{Mat}(m \times n)$, che dunque ha dimensione mn .*

Esempio Nello spazio vettoriale $\mathbf{Mat}(2 \times 2)$ abbiamo quattro matrici elementari:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Ogni matrice 2×2 si scrive:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aE_{11} + bE_{12} + cE_{21} + dE_{22}$$

in modo unico: la quaterna $(E_{11}, E_{12}, E_{21}, E_{22})$ è una base di $\mathbf{Mat}(2 \times 2)$, che ha di conseguenza dimensione pari a 4.

- $(E_{11}, E_{12}, E_{21}, E_{22})$ (con l'ordinamento indicato) sarà chiamata la *base canonica* di $\mathbf{Mat}(2 \times 2)$.

Notiamo anche che le coordinate della matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ rispetto alla base canonica sono date da

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

11 Spazi vettoriali di polinomi

11.1 Combinazioni lineari di polinomi

Ricordiamo che $\mathbf{R}[x]$ denota lo spazio vettoriale dei polinomi nella indeterminata x a coefficienti reali, con le operazioni naturali di somma di due polinomi e moltiplicazione di un polinomio per un numero. Il vettore nullo di $\mathbf{R}[x]$ è per definizione il polinomio nullo $p(x) = 0$.

Possiamo quindi formare, come in ogni spazio vettoriale, combinazioni lineari.

Esempio Se $p_1(x) = 1 - x + 6x^4$ e $p_2(x) = x - x^3 - 4x^4 + x^6$ allora

$$2p_1(x) + 3p_2(x) = 2 + x - 3x^3 + 3x^6.$$

Osserviamo che l'operazione di formare combinazioni lineari di polinomi non può *aumentare il grado* nel senso che:

- Una qualunque combinazione lineare dei polinomi $p_1(x), \dots, p_k(x)$, di grado, rispettivamente, d_1, \dots, d_k , ha grado minore o uguale del massimo fra d_1, \dots, d_k .

11.2 $\mathbf{R}[x]$ non è finitamente generato

Denotiamo con $E_i(x)$ il monomio di grado i , cioè il polinomio:

$$E_i(x) = x^i,$$

dove $i = 0, 1, 2, \dots$. In particolare, $E_0(x) = 1$, il polinomio costante 1. È evidente che ogni polinomio $p(x)$ è combinazione lineare dei monomi $E_i(x)$:

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2 + \dots + a_nx^n \\ &= a_0E_0(x) + a_1E_1(x) + \dots + a_nE_n(x) \end{aligned}$$

Dunque:

- I monomi $E_0(x), E_1(x), \dots$ formano un insieme di generatori di $\mathbf{R}[x]$.

Tali generatori però, sono in numero infinito. A questo punto ci chiediamo se sia possibile trovare un insieme *finito* di generatori di $\mathbf{R}[x]$. La risposta è negativa.

Proposizione $\mathbf{R}[x]$ non è finitamente generato.

Dimostrazione. La dimostrazione si fa per assurdo. Supponiamo che ci sia un insieme finito di generatori, diciamo $p_1(x), \dots, p_k(x)$. Siano d_1, \dots, d_k i gradi rispettivi dei polinomi generatori e sia D il massimo fra d_1, \dots, d_k . Da quanto detto in precedenza, una qualunque combinazione lineare di $p_1(x), \dots, p_k(x)$ avrà grado non superiore a D . Dunque il polinomio x^{D+1} non sarà in alcun modo combinazione lineare dei polinomi dati, e abbiamo così trovato una contraddizione. \square

11.3 Lo spazio vettoriale $\mathbf{R}^n[x]$

Abbiamo visto che $\mathbf{R}[x]$ non è finitamente generato: il motivo è che esso contiene polinomi di grado arbitrariamente grande. Fissiamo ora un intero positivo n , e consideriamo il sottoinsieme $\mathbf{R}^n[x]$ di $\mathbf{R}[x]$ formato dai polinomi di grado minore di n :

$$\mathbf{R}^n[x] = \{a_0 + a_1x + \dots + a_{n-1}x^{n-1} : a_0, \dots, a_{n-1} \in \mathbf{R}\}.$$

È facile verificare che $\mathbf{R}^n[x]$ è chiuso rispetto alla somma e al prodotto per uno scalare, e ovviamente contiene il polinomio nullo. Dunque $\mathbf{R}^n[x]$ è un sottospazio di $\mathbf{R}[x]$, e come tale è esso stesso uno spazio vettoriale. Il polinomio generico di $\mathbf{R}^n[x]$ è combinazione lineare dei monomi $1, x, \dots, x^{n-1}$, che sono linearmente indipendenti. Dunque tali monomi formano una base (finita) di $\mathbf{R}^n[x]$. In conclusione:

- $\mathbf{R}^n[x]$ è uno spazio vettoriale di dimensione n , con base $\{1, x, \dots, x^{n-1}\}$.

Parte 5. Sottospazi

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Sottospazi di \mathbf{R}^n , 1
- 2 Equazioni di un sottospazio di \mathbf{R}^n , 3
- 3 Sottospazio intersezione, 6
- 4 Sottospazio somma, 8
- 5 Formula di Grassmann, 10
- 6 Somma diretta, 11
- 7 Serie di esempi, 14
- 8 Coordinate e criterio del rango, 17
- 9 Spazi vettoriali di matrici e polinomi, 19

1 Sottospazi di \mathbf{R}^n

Premettiamo la seguente

Proposizione *Sia V uno spazio vettoriale di dimensione finita n ed E un suo sottospazio. Allora:*

$$\dim E \leq n.$$

Inoltre $\dim E = n$ se e solo se $E = V$.

Dimostrazione. Sia v_1, \dots, v_k una base di E , cosicché $\dim E = k$. I k vettori della base di E sono linearmente indipendenti dunque $k \leq n$ (perché n è il massimo numero di vettori linearmente indipendenti di V). Supponiamo che $\dim E = n = \dim V$: allora esiste una base di E con n vettori. Essendo tali vettori linearmente indipendenti, e in numero pari alla dimensione di V , essi formano una base anche di V . In conclusione, V ed E hanno in comune una base, dunque coincidono. \square

Quindi, l'unico sottospazio di V che ha dimensione massima (cioè pari a quella di V) è tutto lo spazio. Ricordiamo anche il *sottospazio nullo* $E = \{O\}$ composto dal solo vettore nullo, che per convenzione ha dimensione zero.

1.1 Dimensione e rango

Veniamo ora al seguente problema: dato un sottospazio E di \mathbf{R}^n , descritto con un insieme di generatori, calcolare la sua dimensione e trovare una base.

Lemma a) *I vettori colonna v_1, \dots, v_k di una matrice A sono linearmente indipendenti se e solo se $\text{rk}A = k$.*

b) *Sia A una matrice, e sia A' è la matrice ottenuta da A aggiungendo la colonna v_{k+1} . Allora: v_{k+1} è combinazione lineare delle colonne precedenti se e solo se $\text{rk}A' = \text{rk}A$.*

Dimostrazione. La parte a) è già stata dimostrata. Dimostriamo la b). Supponiamo che le colonne di A siano v_1, \dots, v_k . Allora v_{k+1} è combinazione lineare delle colonne precedenti se e solo se possiamo trovare numeri reali x_1, \dots, x_k tali che:

$$v_{k+1} = x_1 v_1 + \dots + x_k v_k.$$

Questa equazione vettoriale si traduce nel sistema lineare $AX = v_{k+1}$, che è compatibile se e solo se $\text{rk}A' = \text{rk}A$ (teorema di Rouche'-Capelli). \square

In altre parole, se aggiungiamo una colonna v_{k+1} si ha che:

- $\text{rk}A' = \text{rk}A$, se v_{k+1} è combinazione lineare delle colonne precedenti,
- $\text{rk}A' = \text{rk}A + 1$, altrimenti.

Quindi, aggiungendo via via colonne che sono combinazioni lineari delle precedenti, il rango non cambia. Il teorema che segue dà una ulteriore caratterizzazione del rango di una matrice.

Teorema *Sia A una matrice $m \times n$, con vettori colonna $v_1, \dots, v_n \in \mathbf{R}^m$. Allora:*

$$\text{rk}A = \dim L[v_1, \dots, v_n].$$

Cioè, il rango di A uguaglia la dimensione del sottospazio di \mathbf{R}^m generato dalle colonne di A . Inoltre, una base di $L[v_1, \dots, v_n]$ è data dalle colonne corrispondenti ad un minore di ordine massimo di A con determinante non nullo.

Dimostrazione. Il sottospazio $E = L[v_1, \dots, v_n]$ ha generatori v_1, \dots, v_n per ipotesi. Sappiamo che ogni insieme di generatori contiene una base, ottenuta scartando (eventualmente) i generatori inutili. Riordinando, possiamo supporre che una base di E sia formata dalle prime k colonne v_1, \dots, v_k . Quindi tutte le colonne successive saranno combinazione lineare di v_1, \dots, v_k e aggiungendole via via il rango rimarrà costante, per la parte b) del

lemma. Dunque abbiamo:

$$\begin{aligned}
 \dim E &= k \\
 &= \text{rango della matrice } \text{Mat}(v_1, \dots, v_k) \\
 &= \text{rango della matrice } \text{Mat}(v_1, \dots, v_k, v_{k+1}) \\
 &\dots \\
 &= \text{rango della matrice } \text{Mat}(v_1, \dots, v_n) \\
 &= \text{rk} A
 \end{aligned}$$

Questo dimostra la prima parte del teorema. Sia ora M un minore di ordine massimo di A con determinante non nullo. Le colonne corrispondenti a quelle di M sono evidentemente linearmente indipendenti, e formano quindi una base del sottospazio E . \square

Poiché il rango di una matrice uguaglia quello della sua trasposta, avremo anche che

- Il rango di una matrice A uguaglia la dimensione del sottospazio generato dalle righe di A .

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 2 \\ -3 \\ 4 \\ -3 \end{pmatrix}$. Calcolare la dimensione del sottospazio E di \mathbf{R}^4 generato da v_1, v_2, v_3 .

Soluzione. Il rango della matrice $\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & -3 \\ 2 & 0 & 4 \\ 1 & 1 & -3 \end{pmatrix}$ vale 2, dunque $\dim E = 2$. Una base può

essere formata dai primi due vettori v_1, v_2 , corrispondenti al minore $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Notiamo che in effetti $v_3 = 2v_1 - 5v_2$. \square

Esempio Calcolare la dimensione del sottospazio E di \mathbf{R}^3 generato dai vettori

$$u_1 = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}, u_4 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Soluzione. Incolonniamo i vettori, e osserviamo che il rango vale 2. Dunque $\dim E = 2$.

Una base possibile è (u_2, u_3) . Notiamo, a posteriori, qualche relazione fra le colonne:

$$\begin{aligned}
 u_2 &= -2u_1 \\
 u_4 &= u_1 - u_3
 \end{aligned}$$

Dunque potevamo eliminare i generatori u_2 e u_4 , e un'altra base è (u_1, u_3) . \square

2 Equazioni di un sottospazio di \mathbf{R}^n

Sappiamo che l'insieme delle soluzioni di un sistema lineare omogeneo è sempre un sottospazio di \mathbf{R}^n . Vogliamo ora far vedere che è anche vero il viceversa, cioè:

- Ogni sottospazio di \mathbf{R}^n è l'insieme delle soluzioni di un opportuno sistema lineare omogeneo.

Iniziamo con un esempio.

Esempio Consideriamo il sottospazio E di \mathbf{R}^3 generato dai vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$. Esprimiamo E con equazioni.

Per definizione, un vettore $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ appartiene a E se e solo esso è combinazione lineare di v_1 e v_2 . I generatori sono linearmente indipendenti, quindi formano una base di E . Incolonnando i tre vettori, otteniamo la matrice

$$A = \begin{pmatrix} 1 & 2 & x \\ 1 & 1 & y \\ 0 & -1 & z \end{pmatrix}.$$

Ora la terza colonna è combinazione lineare delle precedenti se e solo se $\text{rk}A = 2$. Quindi dobbiamo semplicemente imporre la condizione

$$\begin{vmatrix} 1 & 2 & x \\ 1 & 1 & y \\ 0 & -1 & z \end{vmatrix} = 0.$$

Sviluppando, otteniamo l'equazione:

$$S : x - y + z = 0.$$

- L'equazione $S : x - y + z = 0$ si dice *equazione cartesiana* del sottospazio E . Essa rappresenta E nel senso che $E = \text{Sol}(S)$.

Esempio Vediamo di rappresentare il seguente sottospazio di \mathbf{R}^4 mediante equazioni:

$$E = L \left[\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix} \right].$$

Innanzitutto osserviamo che i due generatori sono linearmente indipendenti, dunque formano una base di E . Consideriamo la matrice ottenuta incolonnando la base, con l'aggiunta del vettore colonna delle incognite: $(x_1, x_2, x_3, x_4)^t$:

$$A = \begin{pmatrix} 1 & 0 & x_1 \\ 1 & 2 & x_2 \\ 0 & -1 & x_3 \\ 0 & 1 & x_4 \end{pmatrix}.$$

Affinché il vettore colonna delle incognite sia combinazione lineare delle prime due colonne, occorre e basta che $\text{rk} A = 2$ (il rango della matrice ottenuta sopprimendo l'ultima colonna).

Orlando il minore $\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ (che ha determinante non nullo) otteniamo due condizioni:

$$\begin{vmatrix} 1 & 0 & x_1 \\ 1 & 2 & x_2 \\ 0 & -1 & x_3 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & x_1 \\ 1 & 2 & x_2 \\ 0 & 1 & x_4 \end{vmatrix} = 0,$$

che danno luogo alle equazioni richieste:

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

Quindi un vettore appartiene a E se e solo se le sue coordinate soddisfano il sistema S , e si ha

$$E = \text{Sol}(S).$$

I due esempi precedenti giustificano il risultato generale che segue.

Proposizione Sia E un sottospazio di \mathbf{R}^n di dimensione k . Allora esiste un sistema lineare omogeneo S di $n - k$ equazioni in n incognite tale che

$$E = \text{Sol}(S).$$

- Le equazioni che compongono il sistema S si dicono *equazioni del sottospazio E* . Tali equazioni non sono uniche; esse però sono sempre, in numero, $n - k = n - \dim E$ (tale numero è detto la *codimensione* del sottospazio E).

Dimostrazione. La dimostrazione fornisce anche un metodo per trovare le equazioni del sottospazio. Fissiamo una base di E , diciamo (v_1, \dots, v_k) , e consideriamo il vettore generico di \mathbf{R}^n , diciamo $v = (x_1, \dots, x_n)^t$. Ora $v \in E$ se e solo se v è combinazione lineare di v_1, \dots, v_k . Questo avviene se e solo se la matrice

$$A = \text{Mat}(v_1, \dots, v_k, v)$$

ha rango uguale a k . La sottomatrice $\text{Mat}(v_1, \dots, v_k)$ ha rango k per ipotesi; dunque possiamo trovare un minore M di tale sottomatrice che ha ordine k e determinante non nullo. Consideriamo i minori orlati di M : essi sono, in numero, $n - k$. Si noti che l'ultima colonna di tali minori è formata da un certo numero di incognite. Annullando i determinanti di tutti i minori orlati otterremo $n - k$ equazioni, che sono le equazioni del sottospazio.

Si noti anche che le equazioni trovate dipendono dalla base scelta, dunque non sono uniche; però, in numero, sono sempre $n - k$. \square

3 Sottospazio intersezione

Consideriamo due sottospazi E, F di uno spazio vettoriale V . Si verifica facilmente che

- l'intersezione $E \cap F$ è un sottospazio di V .

Inoltre $E \cap F$ è un sottospazio sia di E che di F , dunque:

$$\dim(E \cap F) \leq \min\{\dim E, \dim F\}.$$

Esempio Dati i vettori di \mathbf{R}^3

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, v_5 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

consideriamo i sottospazi

$$E = L[v_1, v_2, v_3], F = L[v_4, v_5].$$

Calcolare:

- La dimensione e una base di E e di F ;
- Una base di $E \cap F$.

Soluzione. a) E ha dimensione 2 con base (v_1, v_2) . Infatti, $v_3 = v_2 - v_1$. I generatori di

F sono linearmente indipendenti, dunque una base di F è (v_4, v_5) . Entrambi i sottospazi hanno dimensione 2.

b) Il vettore generico di E si scrive

$$a_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} a_1 + 2a_2 \\ a_1 + a_2 \\ -a_2 \end{pmatrix},$$

con $a_1, a_2 \in \mathbf{R}$. Analogamente, il vettore generico di F è del tipo:

$$\begin{pmatrix} a_3 + a_4 \\ 2a_4 \\ a_4 \end{pmatrix}$$

con $a_3, a_4 \in \mathbf{R}$. Uguagliando, otterremo condizioni affinché un vettore appartenga all'intersezione:

$$\begin{cases} a_1 + 2a_2 = a_3 + a_4 \\ a_1 + a_2 = 2a_4 \\ -a_2 = a_4 \end{cases}.$$

Questo è un sistema omogeneo di tre equazioni in quattro incognite le cui soluzioni sono:

$$\begin{cases} a_1 = 3t \\ a_2 = -t \\ a_3 = 0 \\ a_4 = t \end{cases}$$

con $t \in \mathbf{R}$. Sostituendo nel vettore generico di E (o anche F) otteniamo che il vettore generico dell'intersezione è del tipo:

$$\begin{pmatrix} t \\ 2t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Dunque $E \cap F$ ha dimensione 1, con base $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

Potevamo procedere anche nel modo seguente. Esprimiamo i due sottospazi con equazioni. Poiché E ha dimensione 2, ed è un sottospazio di \mathbf{R}^3 , possiamo descrivere E con una sola equazione, ottenuta annullando il determinante della matrice di colonne v_1, v_2 (la base di E) e il vettore generico:

$$\begin{vmatrix} 1 & 2 & x \\ 1 & 1 & y \\ 0 & -1 & z \end{vmatrix} = 0.$$

Dunque E è descritto dall'equazione:

$$x - y + z = 0.$$

Analogamente F è descritto dall'equazione:

$$y - 2z = 0.$$

Risulterà quindi che le equazioni di $E \cap F$ sono:

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

e risolvendo il sistema otteniamo la base $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. \square

Esercizio Supponiamo che vettori v_1, \dots, v_n di uno spazio vettoriale V siano linearmente indipendenti, e poniamo:

$$E = L[v_1, \dots, v_k], F = L[v_{k+1}, \dots, v_n].$$

Dimostrare che $E \cap F = \{O\}$.

4 Sottospazio somma

L'unione di due sottospazi E ed F non è in generale un sottospazio.

Esempio Siano $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ i vettori della base canonica di \mathbf{R}^2 , e poniamo:

$$E = L[e_1], F = L[e_2].$$

Allora $e_1 + e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ non appartiene né a E né a F , dunque $E \cup F$ non è chiuso rispetto alla somma.

Possiamo però definire il *sottospazio somma* $E + F$, semplicemente sommando, in tutti i modi possibili, un vettore di E e un vettore di F :

$$E + F = \{u + v \in V : u \in E, v \in F\}$$

Proposizione $E + F$ è un sottospazio di V .

Dimostrazione. Esercizio. \square

È evidente che $E + F$ contiene E . Infatti se $v \in E$ possiamo scrivere $v = v + O$ e sappiamo che $O \in F$, poiché F è un sottospazio. Analogamente, $E + F$ contiene anche F . Dunque:

$$\dim(E + F) \geq \max\{\dim E, \dim F\}.$$

Se conosciamo i generatori di E e di F , allora è facile trovare dei generatori della somma.

Proposizione Se E è generato dai vettori u_1, \dots, u_k ed F è generato dai vettori w_1, \dots, w_h allora $E + F$ è generato da $u_1, \dots, u_k, w_1, \dots, w_h$.

Dimostrazione. Un vettore della somma $E + F$ è sempre somma di un vettore u di E e di un vettore w di F . Il primo è combinazione lineare di u_1, \dots, u_k , e il secondo è combinazione lineare di w_1, \dots, w_h . Dunque la loro somma sarà combinazione lineare di $u_1, \dots, u_k, w_1, \dots, w_h$. \square

Quindi:

- Generatori del sottospazio somma si ottengono prendendo l'unione dei generatori degli addendi.

Esempio In \mathbf{R}^3 consideriamo i sottospazi:

$$E = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right], \quad F = L \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right].$$

Determiniamo il sottospazio somma. Dalla proposizione, vediamo che

$$E + F = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right],$$

e quindi la dimensione di $E + F$ è pari al rango della matrice

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

Tale rango è tre e quindi $\dim(E + F) = 3$. L'unico sottospazio di \mathbf{R}^3 di dimensione 3 è tutto \mathbf{R}^3 , dunque:

$$E + F = \mathbf{R}^3.$$

Esercizio Sia V uno spazio vettoriale con base v_1, \dots, v_n . Sia $1 \leq k \leq n$ e poniamo

$$E = L[v_1, \dots, v_k], F = [v_{k+1}, \dots, v_n].$$

Dimostrare che $E + F = V$ e $E \cap F = \{O\}$.

Esercizio Dati i sottospazi E, F di uno spazio vettoriale V , dimostrare che $E \subseteq F$ se e solo se $E + F = F$.

5 Formula di Grassmann

C'è una relazione fra le dimensioni dei sottospazi $E, F, E + F, E \cap F$, detta *formula di Grassmann*.

Teorema Siano E, F sottospazi di uno spazio vettoriale V . Allora si ha:

$$\dim(E + F) + \dim(E \cap F) = \dim E + \dim F.$$

Dimostrazione. Omessa.

La formula di Grassmann permette di abbreviare il calcolo delle dimensioni.

Esempio In \mathbf{R}^3 , consideriamo i sottospazi

$$E = L \left[\begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \right], F = L \left[\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right].$$

Determiniamo $E \cap F$ e $E + F$.

Notiamo che $\dim E = 1$ e $\dim F = 2$. Ora $E + F$ ha dimensione data da:

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

Dalla formula di Grassmann otteniamo $\dim(E \cap F) = 0$, dunque $E \cap F = \{O\}$.

Esempio Supponiamo che E ed F siano due sottospazi di \mathbf{R}^6 , tali che:

$$\dim E = 3, \dim F = 5.$$

Quali valori può assumere $\dim(E \cap F)$?

Soluzione. Sappiamo che $E \cap F$ è un sottospazio sia di E che di F . Dunque si ha sicuramente:

$$0 \leq \dim(E \cap F) \leq 3.$$

La formula di Grassmann ci permette di essere più precisi. Infatti, F è un sottospazio di $E + F$, che è a sua volta un sottospazio di \mathbf{R}^6 , dunque:

$$5 \leq \dim(E + F) \leq 6,$$

inoltre $\dim(E + F) = 5$ se e solo se $E \subseteq F$. Dalla formula di Grassmann otteniamo

$$\begin{aligned} \dim(E \cap F) &= \dim E + \dim F - \dim(E + F) \\ &= 8 - \dim(E + F). \end{aligned}$$

Quindi $\dim(E \cap F)$ può valere 2 oppure 3, e vale 3 esattamente quando $E \subseteq F$. In conclusione:

$$\dim(E \cap F) = \begin{cases} 2 & \text{se } E \text{ non è contenuto in } F, \\ 3 & \text{se } E \text{ è contenuto in } F. \end{cases}$$

6 Somma diretta

Diremo che la somma di due sottospazi $U + W$ è *somma diretta* se $U \cap W = \{O\}$. Scriveremo in tal caso:

$$U + W = U \oplus W.$$

Proposizione Sia V uno spazio vettoriale e U, W due suoi sottospazi. Allora $V = U \oplus W$

se e solo se ogni vettore $v \in V$ si decompone, in modo unico

$$v = u + w,$$

dove $u \in U$ e $w \in W$.

Dimostrazione. È un caso particolare del teorema enunciato più avanti. \square

Esempio La somma di due sottospazi U, W di \mathbf{R}^3 , entrambi di dimensione 2, non è mai diretta: infatti si verifica che $\dim(U + W) \geq 1$.

Esempio In \mathbf{R}^3 consideriamo $U = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right], W = L \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right]$. Si verifica che $U + W$ ha dimensione 3, dunque $U + W = \mathbf{R}^3$. Dalla formula di Grassmann otteniamo

$U \cap W = \{O\}$. Dunque $\mathbf{R}^3 = U \oplus W$. Decomponiamo il vettore $\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$ nella somma $u + w$ con $u \in U$ e $w \in W$. Il vettore generico di U si scrive

$$\begin{pmatrix} a+b \\ 2a+b \\ a \end{pmatrix}$$

mentre il vettore generico di W è $\begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix}$. Imponiamo:

$$\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} a+b \\ 2a+b \\ a \end{pmatrix} + \begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix},$$

e otteniamo $a = 5, b = -7, c = 4$. Dunque

$$\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix},$$

in cui il primo vettore appartiene a U mentre il secondo appartiene a W . \square

Esempio Nello spazio vettoriale $\mathbf{Mat}(n \times n)$ si consideri il sottospazio $S(n)$ formato dalle matrici simmetriche, e il sottospazio $W(n)$ formato dalle matrici antisimmetriche. Allora

$$\mathbf{Mat}(n \times n) = S(n) \oplus W(n).$$

Dimostrare inoltre che

$$\dim S = \frac{n(n+1)}{2}, \quad \text{e} \quad \dim W = \frac{n(n-1)}{2}.$$

Soluzione. Occorre dimostrare che ogni matrice è la somma di una matrice simmetrica e di una matrice antisimmetrica, e che $S(n) \cap W(n) = \{0\}$. Ora possiamo scrivere, per ogni $A \in \mathbf{Mat}(n \times n)$:

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t),$$

il che dimostra la prima parte; se $A \in S(n) \cap W(n)$ allora $A = A^T$ e, al tempo stesso, $A = -A^T$. Quindi A è necessariamente la matrice nulla, il che dimostra la seconda parte.

La seconda parte è lasciata per esercizio (partire dalle dimensioni basse, e generalizzare alla dimensione arbitraria).

Esempio Decomporre la matrice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ nella somma di una matrice simmetrica e di una matrice antisimmetrica.

Soluzione.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 5/2 \\ 5/2 & 4 \end{pmatrix} + \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix}.$$

Diamo ora dei criteri per la somma diretta.

Teorema *Le seguenti condizioni sono equivalenti:*

- a) $V^n = U \oplus W$.
- b) $\dim U + \dim W = n$ e inoltre $U \cap W = \{O\}$.
- c) Se $B = (u_1, \dots, u_k)$ è una base di U e (w_1, \dots, w_h) è una base di W allora i vettori $u_1, \dots, u_k, w_1, \dots, w_h$ formano una base di V^n .
- d) Ogni vettore di V si decompone, in modo unico, nella somma $u + w$, con $u \in U$ e $w \in W$.

Dimostrazione. È sufficiente dimostrare le implicazioni:

$$a \implies b, b \implies c, c \implies d, d \implies a.$$

$a) \implies b)$. Basta verificare che $\dim U + \dim W = n$, e questo discende immediatamente dall'ipotesi a) e dalla formula di Grassmann.

$b) \implies c)$. Per ipotesi, $k+h = n$ e quindi basta dimostrare che i vettori $u_1, \dots, u_k, w_1, \dots, w_h$ sono linearmente indipendenti. Supponiamo

$$a_1 u_1 + \dots + a_k u_k + b_1 w_1 + \dots + b_h w_h = 0,$$

dunque:

$$a_1 u_1 + \dots + a_k u_k = -(b_1 w_1 + \dots + b_h w_h).$$

L'uguaglianza mostra che il vettore a sinistra è in $U \cap W$, dunque:

$$a_1 u_1 + \dots + a_k u_k = 0$$

per l'ipotesi $U \cap W = \{O\}$. Siccome u_1, \dots, u_k sono linearmente indipendenti, necessariamente $a_1 = \dots = a_k = 0$. Analogamente si dimostra che $b_1 = \dots = b_h = 0$.

$c) \implies d)$. Per l'ipotesi c), si ha che $(u_1, \dots, u_k, w_1, \dots, w_h)$ è una base di V^n . Dato un vettore v in V^n , possiamo scrivere:

$$\begin{aligned} v &= a_1 u_1 + \dots + a_k u_k + b_1 w_1 + \dots + b_h w_h \\ &= (a_1 u_1 + \dots + a_k u_k) + (b_1 w_1 + \dots + b_h w_h). \end{aligned}$$

Se poniamo $u = a_1 u_1 + \dots + a_k u_k$ e $w = b_1 w_1 + \dots + b_h w_h$, allora $u \in U, w \in W$ e

$$v = u + w,$$

e questa decomposizione è evidentemente unica, poiché abbiamo usato basi di U e di W .

$d) \implies a)$. È evidente che $V^n = U + W$. Occorre solamente dimostrare che $U \cap W = \{O\}$. Supponiamo che $v \in U \cap W$. Allora possiamo scrivere:

$$\begin{aligned} v &= v + O \\ v &= O + v \end{aligned}$$

dove nella prima uguaglianza O è pensato come vettore di W , mentre nella seconda come vettore di U . Dall'unicità della decomposizione, si $v = O$. Dunque $U \cap W = \{O\}$. \square

7 Serie di esempi

Dati i sottospazi U, W dello spazio vettoriale V indicato, determinare in ciascuno dei casi base e dimensione di U e W , della somma $U + W$ e dell'intersezione $U \cap W$. Stabilire inoltre se la somma è diretta.

7.1 Esempio

In $V = \mathbf{R}^3$:

$$U = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right], W = L \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right].$$

Soluzione. Risposte: $U \cap W$ è generato da $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$, $U + W = \mathbf{R}^3$, ma la somma non è

diretta.

Per trovare l'intersezione determiniamo prima le equazioni dei due sottospazi. I generatori sono linearmente indipendenti e quindi formano una base in ciascun caso. Equazione di

U :

$$\begin{vmatrix} 1 & 1 & x \\ 2 & 1 & y \\ 1 & 0 & z \end{vmatrix} = 0, \quad \text{quindi} \quad x - y + z = 0.$$

Equazione di W :

$$\begin{vmatrix} 1 & 2 & x \\ 0 & 1 & y \\ 0 & -1 & z \end{vmatrix} = 0, \quad \text{quindi} \quad y + z = 0.$$

Equazioni di $U \cap W$:

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

da cui, risolvendo, otteniamo la base $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ di $U \cap W$.

Dalla formula di Grassmann otteniamo immediatamente $\dim(U+W) = 3$ dunque $U+W = \mathbf{R}^3$. La somma non è diretta perché l'intersezione non è nulla. L'esercizio è finito.

Osserviamo che per determinare la somma potevamo anche procedere come segue. Sappiamo che

$$U + W = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right],$$

e quindi la dimensione è pari al rango della matrice

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

Tale rango vale tre, quindi $\dim(U+W) = 3$ e necessariamente $U+W = \mathbf{R}^3$.

7.2 Esempio

In $V = \mathbf{R}^3$:

$$U = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right], \quad W = L \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right].$$

Soluzione. Risposte: $U \cap W = \{O\}$, $U+W = \mathbf{R}^3$, e la somma è diretta: $\mathbf{R}^3 = U \oplus W$.

Calcoliamo la dimensione di U . I tre generatori sono linearmente dipendenti perché il terzo è la somma dei primi due: dunque possiamo scartare il terzo generatore e

$$U = L \left[\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right],$$

che coincide con il sottospazio U dell'esempio precedente. Dunque $\dim U = 2$. È immediato che $\dim W = 1$. Sappiamo già che l'equazione di U è $x - y + z = 0$. Le equazioni di W sono immediatamente date da $y = 0, z = 0$. Dunque $U \cap W$ è descritto dalle equazioni:

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

che ammettono solamente la soluzione comune nulla, dunque $U \cap W = \{0\}$. Dalla formula di Grassmann otteniamo $\dim(U + W) = 3$ dunque $U + W = \mathbf{R}^3$. La somma è diretta perché l'intersezione è nulla.

7.3 Esempio

In $V = \mathbf{R}^4$:

$$U = L \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 2 \end{pmatrix} \right], W = L \left[\begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right]$$

Soluzione. Risposte: $U \cap W = \{0\}$, $U + W = \mathbf{R}^4$ e quindi $\mathbf{R}^4 = U \oplus W$.

Siano u_1, u_2 i generatori di U e w_1, w_2 i generatori di W . In entrambi i casi essi sono linearmente indipendenti, quindi $\dim U = \dim W = 2$. Si ha che

$$U + W = L[u_1, u_2, w_1, w_2]$$

e poiché i quattro generatori sono linearmente indipendenti (ciò si verifica osservando che il rango della matrice corrispondente è quattro) otteniamo che $\dim(U + W) = 4$ e quindi $U + W = \mathbf{R}^4$. Dalla formula di Grassmann otteniamo che $U \cap W = \{0\}$, e la somma è diretta.

Si può però procedere in modo alternativo. Sia $v \in U \cap W$. Poiché $v \in U$, possiamo trovare $a, b \in \mathbf{R}$ tali che:

$$v = au_1 + bu_2.$$

Analogamente, poiché $v \in W$ esisteranno numeri reali c e d tali che

$$v = cw_1 + dw_2.$$

Uguagliando, a, b, c, d verificano l'equazione vettoriale:

$$au_1 + bu_2 = cw_1 + dw_2,$$

e quindi

$$au_1 + bu_2 - cw_1 - dw_2 = 0.$$

Ora sappiamo che i vettori u_1, u_2, w_1, w_2 sono linearmente indipendenti, e quindi necessariamente $a = b = c = d = 0$. Dunque l'unico vettore nell'intersezione è il vettore nullo.

7.4 Esempio

In $V = \mathbf{R}^4$:

$$U = L \left[\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right], W = L \left[\begin{pmatrix} 5 \\ 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -1 \\ -1 \end{pmatrix} \right]$$

Soluzione. Risposte: $\dim U = \dim W = \dim(U \cap W) = \dim(U + W) = 2$. Quindi $U = W$.

La somma non è diretta.

Chiamiamo u_1, u_2, u_3 i generatori di U , nell'ordine. Poiché il rango della matrice da essi formata è 2, la dimensione di U è 2 e una base è data da u_1, u_2 . I generatori di W sono linearmente indipendenti e formano una base. Dunque entrambi i sottospazi hanno dimensione 2. Procedendo come negli esempi precedenti, si trova (dopo qualche calcolo) che il rango della matrice di righe u_1, u_2, w_1, w_2 è 2, e quindi

$$\dim(U + W) = 2.$$

Dalla formula di Grassmann $\dim(U \cap W) = 2 = \dim U$; poiché $U \cap W$ è un sottospazio di U di dimensione pari a quella di U si dovrà avere

$$U = U \cap W,$$

quindi $W \subseteq U$. Analogamente $U \subseteq W$ e quindi $U = W$. I due sottospazi coincidono! Possiamo anche dire che (u_1, u_2) e (w_1, w_2) sono due basi dello *stesso* sottospazio.

Per dimostrare che $U = W$ si poteva anche osservare che

$$\begin{aligned} w_1 &= u_1 + 2u_2 \\ w_2 &= 3u_1 - u_2 \end{aligned},$$

che dimostra $W \subseteq U$. Poiché U e W hanno la stessa dimensione, necessariamente $U = W$.

8 Coordinate e criterio del rango

Nei precedenti esempi abbiamo studiato questioni di dipendenza e indipendenza lineare, e questioni riguardanti i sottospazi, principalmente nello spazio vettoriale \mathbf{R}^n . In tal caso possiamo utilizzare il rango di opportune matrici per risolvere i problemi. In questa sezione faremo vedere come sia possibile utilizzare il criterio del rango in ogni spazio vettoriale (di dimensione finita).

In ciò che segue, utilizzeremo la notazione V^n per indicare uno spazio vettoriale di dimensione n .

8.1 Coordinate di un vettore rispetto a una base

Ricordiamo che, se $\mathcal{B} = (e_1, \dots, e_n)$ è una base (ordinata) di uno spazio vettoriale V^n allora possiamo esprimere ogni vettore $v \in V$ come combinazione lineare:

$$v = x_1 e_1 + \dots + x_n e_n,$$

in modo unico. Questo significa che i coefficienti x_1, \dots, x_n sono univocamente determinati dal vettore v (e dalla base \mathcal{B}). Il vettore di \mathbf{R}^n dato dai coefficienti:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

è detto il *vettore delle coordinate di v rispetto alla base \mathcal{B}* .

Se scriviamo

$$\mathcal{F}(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \tag{1}$$

abbiamo così una corrispondenza tra i vettori di V e quelli di \mathbf{R}^n , che esprimeremo con la notazione:

$$\mathcal{F} : V \rightarrow \mathbf{R}^n.$$

- \mathcal{F} è un esempio di *applicazione* dall'insieme V nell'insieme \mathbf{R}^n .

Notiamo che $\mathcal{F}(O) = O$, e $\mathcal{F}(v_j) = e_j$ (j -esimo vettore della base canonica di \mathbf{R}^n). È chiaro che l'applicazione \mathcal{F} dipende in modo essenziale dalla base scelta: lo stesso vettore ha coordinate diverse in basi diverse.

- In generale, se A e B sono insiemi, un'applicazione f dall'insieme A nell'insieme B è una legge che associa ad ogni elemento $a \in A$ uno ed un solo elemento di B , denotato $f(a)$. Un'applicazione si scrive:

$$f : A \rightarrow B,$$

dove A (insieme di destra) è l'insieme di partenza e B (insieme di sinistra) è l'insieme di arrivo.

Nelle prossime lezioni studieremo in dettaglio una classe naturale di applicazioni fra spazi vettoriali: le cosiddette *applicazioni lineari*.

8.2 Proprietà delle coordinate

L'applicazione \mathcal{F} ha le seguenti proprietà.

Proposizione a) Sia V^n uno spazio vettoriale di dimensione n e sia $\mathcal{B} = (v_1, \dots, v_n)$ una base ordinata di V^n . Se $v \in V$, indichiamo con

$$\mathcal{F}(v) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

il vettore delle coordinate di v rispetto a \mathcal{B} . Allora:

- a) I vettori w_1, \dots, w_k sono linearmente indipendenti (in V) se e solo se le rispettive coordinate $\mathcal{F}(w_1), \dots, \mathcal{F}(w_k)$ sono vettori linearmente indipendenti (in \mathbf{R}^n).
- b) I vettori w_1, \dots, w_k generano V se e solo se le rispettive coordinate $\mathcal{F}(w_1), \dots, \mathcal{F}(w_k)$ generano \mathbf{R}^n .
- c) In particolare, \mathcal{F} trasforma basi di V in basi di \mathbf{R}^n .

Dimostrazione. La dimostrazione sarà data in seguito. \square

Dunque, l'applicazione delle coordinate permette di trasferire problemi di dipendenza e indipendenza lineare dallo spazio vettoriale V allo spazio vettoriale \mathbf{R}^n , dove possiamo usare il criterio del rango.

Proposizione Sia V un qualunque spazio vettoriale e $\mathcal{B} = (v_1, \dots, v_n)$ una sua base. Dati i vettori w_1, \dots, w_k di V , consideriamo la matrice A ottenuta incolonnando le coordinate di w_1, \dots, w_k rispetto a \mathcal{B} . Allora:

- a) I vettori w_1, \dots, w_k sono linearmente indipendenti se e solo se $\text{rk} A = k$.
- b) Più in generale si ha: $\dim L[w_1, \dots, w_k] = \text{rk} A$.

9 Spazi vettoriali di matrici e polinomi

9.1 Esempi su spazi di matrici

Per semplicità ci ridurremo a considerare lo spazio vettoriale $V = \mathbf{Mat}(2 \times 2)$. La base piu' semplice di $\mathbf{Mat}(2 \times 2)$ è la cosiddetta *base canonica* di $\mathbf{Mat}(2 \times 2)$, cioè la base $(E_{11}, E_{12}, E_{21}, E_{22})$ dove:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Quindi $\dim \mathbf{Mat}(2 \times 2) = 4$. La matrice generica si scrive

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aE_{11} + bE_{12} + cE_{21} + dE_{22},$$

e scelta la base canonica, abbiamo

$$\mathcal{F} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Esempio Sono date le matrici $A_1 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}.$

- a) Stabilire se A_1, A_2, A_3, A_4 sono linearmente indipendenti oppure no.
- b) Calcolare la dimensione del sottospazio E di $\mathbf{Mat}(2 \times 2)$ generato dalle quattro matrici.

Soluzione. a) Incolonniamo le coordinate delle quattro matrici (rispetto alla base canonica $(E_{11}, E_{12}, E_{21}, E_{22})$) e otteniamo la matrice

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

Se il rango di A vale 4 le matrici sono linearmente indipendenti, altrimenti no. Un calcolo mostra che $\det A = 0$, dunque $\text{rk} A < 4$ e le matrici risultano linearmente dipendenti. In effetti, possiamo notare che $A_4 = A_2 + 2A_3$.

b) Basta calcolare il rango di A . Sappiamo che $\text{rk} A \leq 3$; ora il minore di ordine 3 in alto a sinistra ha determinante non nullo, dunque il rango vale 3, e tale è la dimensione del sottospazio cercato. Possiamo verificare che le matrici A_1, A_2, A_3 sono linearmente indipendenti: poiché tali matrici sono in numero pari alla dimensione di E , esse formano una base di E .

Esempio Sono date le matrici $M_1 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, M_2 = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, M_3 = \begin{pmatrix} 1 & 4 \\ 0 & 0 \end{pmatrix}, M_4 = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$. Verificare che le quattro matrici formano una base di $\mathbf{Mat}(2 \times 2)$.

Soluzione. In effetti, la matrice delle coordinate è:

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

Si verifica facilmente che il suo determinante è non nullo, dunque le quattro matrici sono linearmente indipendenti. Poiché $\dim \mathbf{Mat}(2 \times 2) = 4$, questo è sufficiente per affermare che esse formano una base. Altrimenti, potevamo osservare che le coordinate delle quattro matrici formano una base di \mathbf{R}^4 . \square

Esempio Data la matrice $N = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}$, consideriamo il sottoinsieme E di $\mathbf{Mat}(2 \times 2)$ costituito dalle matrici $X \in \mathbf{Mat}(2 \times 2)$ tali che $NX = O$, dove O è la matrice nulla. In altre parole:

$$E = \{X \in \mathbf{Mat}(2 \times 2) : NX = O\}.$$

- a) Dimostrare che E è un sottospazio di $\mathbf{Mat}(2 \times 2)$.
- b) Trovare una base di E e calcolare la sua dimensione.

Soluzione. a) È chiaro che la matrice nulla appartiene a E , perché $NO = O$. Se $X_1, X_2 \in E$ allora per ipotesi $NX_1 = NX_2 = O$. Dunque

$$N(X_1 + X_2) = NX_1 + NX_2 = O + O = O,$$

e anche $X_1 + X_2 \in E$. Di conseguenza E è chiuso rispetto alla somma. La chiusura rispetto al prodotto per uno scalare si dimostra in modo analogo. E è un sottospazio.

b) Cerchiamo un'espressione per la matrice generica del sottospazio E . Partiamo dalla matrice generica di $\mathbf{Mat}(2 \times 2)$

$$X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

e imponiamo la condizione $NX = O$. Si ottiene

$$\begin{pmatrix} x - 2z & y - 2w \\ -x + 2z & -y + 2w \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

che equivale al sistema lineare omogeneo

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

La terza (risp. quarta) equazione è equivalente alla prima (risp. seconda). Dunque il sistema si riduce a

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

Ponendo $z = t$ e $w = s$ otteniamo le ∞^2 soluzioni

$$\begin{cases} x = 2t \\ y = s \\ z = t \\ w = s \end{cases}$$

con $t, s \in \mathbf{R}$ e la matrice generica di E si scrive $\begin{pmatrix} 2t & 2s \\ t & s \end{pmatrix}$, con $t, s \in \mathbf{R}$. Per trovare una base di E , basta scrivere la matrice generica nel seguente modo:

$$\begin{pmatrix} 2t & 2s \\ t & s \end{pmatrix} = t \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} + s \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}.$$

Dunque le matrici $\begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$ generano E , e si vede subito che sono anche linearmente indipendenti. In conclusione, E ha dimensione 2 e una sua base è $\left(\begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix} \right)$.

□

Esercizio Sia $M \in \mathbf{Mat}(2 \times 2)$ una matrice invertibile. Dimostrare che il sottospazio

$$E = \{X \in \mathbf{Mat}(2 \times 2) : MX = O\}$$

è costituito solo dalla matrice nulla, cioè $E = \{O\}$.

Esercizio Nello spazio vettoriale $\mathbf{Mat}(n \times n)$, si considerino il sottoinsieme $T^+(n)$ (rispettivamente $T^-(n)$) formato dalle matrici triangolari superiori (rispettivamente, triangolari inferiori).

- a) Dimostrare che $T^+(n)$ e $T^-(n)$ sono entrambi sottospazi di $\mathbf{Mat}(n \times n)$.
- b) Trovare una base e la dimensione di $T^+(2)$ e $T^-(2)$.
- c) Descrivere il sottospazio intersezione $T^+(2) \cap T^-(2)$.
- d) Dimostrare che $\mathbf{Mat}(2 \times 2) = T^+(2) + T^-(2)$.
- e) Verificare che $\mathbf{Mat}(2 \times 2) \neq T^+(2) \oplus T^-(2)$.

Esercizio Si consideri il sottoinsieme

$$E = \{A \in \mathbf{Mat}(n \times n) : \det A = 0\}.$$

Verificare che E non è un sottospazio di $\mathbf{Mat}(n \times n)$.

(In effetti, E contiene la matrice nulla ed è chiuso rispetto al prodotto per uno scalare. Mostrare però che E non è chiuso rispetto alla somma, fornendo un controesempio).

Esercizio Data una matrice quadrata A , si definisce *traccia* di A la somma di tutti gli elementi diagonali. Esempio:

$$\text{tr} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9 = 15.$$

La traccia di una matrice è quindi un numero. Dimostrare che il sottoinsieme

$$E = \{A \in \mathbf{Mat}(2 \times 2) : \text{tr} A = 0\}$$

è un sottospazio di $\mathbf{Mat}(2 \times 2)$, e trovare una base di E .

9.2 Esempi su spazi di polinomi

Consideriamo lo spazio vettoriale $\mathbf{R}^n[x]$ dei polinomi di grado minore di n . Una base di $\mathbf{R}^n[x]$ è

$$(1, x, x^2, \dots, x^{n-1}),$$

detta anche la *base canonica* di $\mathbf{R}^n[x]$. Quindi $\dim \mathbf{R}^n[x] = n$. Rispetto a tale base, il

polinomio $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$ ha coordinate $\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}$. L'applicazione

delle coordinate $\mathcal{F} : \mathbf{R}^n[x] \rightarrow \mathbf{R}^n$ è data da

$$\mathcal{F}(a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}) = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}.$$

Esempio Stabilire se i polinomi $p_1(x) = 1 + x$, $p_2(x) = 1 - x + 2x^2$, $p_3(x) = 1 + x^2$ sono linearmente indipendenti oppure no, e calcolare la dimensione del sottospazio E di $\mathbf{R}^3[x]$ da essi generato.

Soluzione. Incolonniamo le coordinate rispetto alla base $(1, x, x^2)$ (la base canonica) e otteniamo la matrice

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

Il determinante si annulla, e il rango vale 2. Dunque i polinomi sono linearmente dipendenti e il sottospazio da essi generato ha dimensione 2. In effetti si ha la relazione di dipendenza lineare

$$p_1(x) + p_2(x) - 2p_3(x) = 0.$$

Una base di E sarà data da una qualunque coppia di polinomi linearmente indipendenti, ad esempio $(p_1(x), p_2(x))$. Notiamo infine che E è un sottospazio propriamente contenuto in $\mathbf{R}^3[x]$, poiché ha dimensione minore di 3 (la dimensione di $\mathbf{R}^3[x]$). \square

Esempio Verificare che i polinomi $p_1(x) = x$, $p_2(x) = 1 - x + 2x^2$, $p_3(x) = 1 + 2x^2$ sono linearmente indipendenti e formano una base di $\mathbf{R}^3[x]$.

Soluzione. La matrice delle coordinate:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

ha determinante non nullo e i tre polinomi sono linearmente indipendenti; essi formano automaticamente una base dato che $\dim \mathbf{R}^3[x] = 3$. \square

Parte 6. Applicazioni lineari

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Applicazioni fra insiemi, 1
- 2 Applicazioni lineari tra spazi vettoriali, 2
- 3 Applicazioni lineari da \mathbf{R}^n a \mathbf{R}^m , 4
- 4 Omomorfismo assegnato su una base, 8
- 5 Matrice associata, 11
- 6 Nucleo, 14
- 7 Immagine, 15
- 8 Esempi, 16
- 9 Teorema della dimensione, 18
- 10 Isomorfismi, 23

1 Applicazioni fra insiemi

Se A e B sono due insiemi, un' *applicazione* da A in B è una legge che associa a ciascun elemento $a \in A$ un elemento $f(a) \in B$, e scriveremo

$$f : A \rightarrow B.$$

A è detto *insieme di partenza* e B *insieme di arrivo*. L'*immagine* di f è il sottoinsieme di B costituito da tutti gli elementi di B che sono immagine di almeno un elemento di A :

$$\text{Im}f = \{b \in B : b = f(a) \text{ per qualche } a \in A\}.$$

L'applicazione f si dice *suriettiva* se $\text{Im}f = B$, se cioè ogni elemento di B è l'immagine di qualche elemento di A .

L'applicazione f si dice *iniettiva* se trasforma elementi distinti di A in elementi distinti di B , quindi se:

$$a \neq a' \implies f(a) \neq f(a').$$

Un modo equivalente di esprimere l'iniettività di f è il seguente:

$$f(a) = f(a') \implies a = a'.$$

Un'applicazione f si dice *biiettiva* (o anche *biunivoca*) se è sia iniettiva che suriettiva.

Esempio Stabilire l'iniettività e la suriettività di ciascuna delle seguenti applicazioni $f_i : \mathbf{R} \rightarrow \mathbf{R}$.

$$f_1(x) = x^2$$

$$f_2(x) = e^x$$

$$f_3(x) = x^3 - x$$

$$f_4(x) = x^3$$

Soluzione. (a) $f_1(x)$ non è iniettiva, poichè $f_1(1) = f_1(-1) = 1$, e non è suriettiva, poichè

$\text{Im} f_1 = [0, \infty)$, quindi in particolare $-1 \notin \text{Im} f_1$.

(b) f_2 è iniettiva perché la funzione esponenziale è crescente; non è però suriettiva perché $\text{Im} f_2 = (0, \infty)$, quindi $-1 \notin \text{Im} f_2$.

(c) f_3 non è iniettiva, poichè $f_3(-1) = f_3(0) = f_3(1) = 0$; è però suriettiva (disegnare il grafico di $f_3(x)$).

(d) f_4 è iniettiva poichè crescente, e anche suriettiva: l'equazione $x^3 = y$ nell'incognita x ammette una soluzione (per di più unica) per ogni $y \in \mathbf{R}$:

$$y = x^{1/3}$$

Dunque f_4 è biiettiva.

Esercizio Dimostrare che l'applicazione

$$f : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{R}$$

definita da $f(A) = \det A$ è suriettiva ma non iniettiva.

Esercizio Sia ora $A = \{1, 2, 3\}$. Elencare tutte le applicazioni biettive da A in A .

2 Applicazioni lineari tra spazi vettoriali

Fissiamo ora due spazi vettoriali V e V' , che d'ora in poi supporremo di dimensione finita. Ci proponiamo di studiare in dettaglio un tipo particolare di applicazioni da V in V' : le *applicazioni lineari*, che sono quelle applicazioni che "rispettano" le operazioni definite in uno spazio vettoriale (somma e prodotto per uno scalare).

Definizione Sia $f : V \rightarrow V'$ un'applicazione da uno spazio vettoriale V in uno spazio vettoriale V' . Tale f si dice *lineare* se verifica le seguenti proprietà.

$$\begin{cases} f(u+v) = f(u) + f(v) & \text{per ogni } u, v \in V, \\ f(au) = af(u) & \text{per ogni } a \in \mathbf{R}, u \in V. \end{cases}$$

Osserviamo che le due condizioni sono equivalenti all'unica condizione:

$$f(au + bv) = af(u) + bf(v) \quad \text{per ogni } a, b \in \mathbf{R}, u, v \in V.$$

Per induzione, si dimostra che un'applicazione lineare trasforma combinazioni lineari in combinazioni lineari:

$$f(a_1v_1 + \cdots + a_kv_k) = a_1f(v_1) + \cdots + a_kf(v_k),$$

per ogni $a_1, \dots, a_k \in \mathbf{R}, v_1, \dots, v_k \in V$.

- Un'applicazione lineare da V in V' si dice anche un *omomorfismo* da V in V' .

Vedremo molti esempi più in avanti. Osserviamo poi che ogni applicazione lineare trasforma il vettore nullo di V nel vettore nullo di V' , e porta vettori opposti in vettori opposti:

Proposizione Sia $f : V \rightarrow V'$ un'applicazione lineare. Allora:

- $f(O) = O$.
- $f(-v) = -f(v)$ per ogni $v \in V$.

Dimostrazione. a) Basta osservare che

$$f(O) = f(O + O) = f(O) + f(O)$$

dunque $f(O) = O$.

b) Abbiamo

$$O = f(O) = f(v - v) = f(v) + f(-v)$$

dunque $f(-v) = -f(v)$. \square

Esempio Un esempio banale di applicazione lineare è dato dall'applicazione *nulla*, denotata con $O : V \rightarrow V'$, che trasforma tutti i vettori di V nel vettore nullo di V' :

$$O(v) = O,$$

per ogni $v \in V$.

Esempio L'applicazione $f : \mathbf{Mat}(n \times n) \rightarrow \mathbf{Mat}(n \times n)$ definita da $f(A) = A^T$ è lineare, poiché, per ogni $A, B \in \mathbf{Mat}(n \times n)$ e ogni $k \in \mathbf{R}$:

$$\begin{cases} f(A + B) = (A + B)^T = A^T + B^T = f(A) + f(B) \\ f(kA) = (kA)^T = kA^T = kf(A) \end{cases}$$

Esempio L'applicazione $f : \mathbf{Mat}(n \times n) \rightarrow \mathbf{R}$ definita da $f(A) = \det A$ non è lineare. Basta infatti osservare che, se

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

allora $\det(A + B) = \det I = 1$, mentre $\det A + \det B = 0$, quindi

$$f(A + B) \neq f(A) + f(B).$$

Esempio L'operatore di derivazione $D : \mathbf{R}^n[x] \rightarrow \mathbf{R}^n[x]$, definito come:

$$D(p(x)) = p'(x)$$

è un'applicazione lineare.

3 Applicazioni lineari da \mathbf{R}^n a \mathbf{R}^m

Iniziamo descrivendo una classe di applicazioni lineari da \mathbf{R}^n a \mathbf{R}^m . Sia A una matrice $m \times n$, e $v \in \mathbf{R}^n$, scritto in forma colonna. Allora il prodotto Av è un vettore di \mathbf{R}^m . Dunque possiamo definire un'applicazione $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ nel seguente modo:

$$f(v) = Av.$$

Dalle proprietà del prodotto di matrici, vediamo subito che f è un'applicazione lineare.

Esempio La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix}$ definisce la seguente applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^2$:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ -y + 2z \end{pmatrix}.$$

Notiamo che ogni entrata di $f \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ è una funzione lineare omogenea di x, y, z . Notiamo anche che se (e_1, e_2, e_3) è la base canonica di \mathbf{R}^3 , allora:

$$f(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad f(e_3) = \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

sono i vettori colonna della matrice A . \square

Esempio Ora vogliamo dimostrare che ogni applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ si scrive

$$f(v) = Av$$

con A un'opportuna matrice 2×3 . Consideriamo le immagini dei vettori della base canonica:

$$f(e_1) = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}, \quad f(e_3) = \begin{pmatrix} a_3 \\ b_3 \end{pmatrix}.$$

Allora, se $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ è il vettore generico di \mathbf{R}^3 :

$$\begin{aligned} f(v) &= f \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= f(xe_1 + ye_2 + ze_3) \\ &= xf(e_1) + yf(e_2) + zf(e_3) \\ &= x \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + y \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} + z \begin{pmatrix} a_3 \\ b_3 \end{pmatrix} \\ &= \begin{pmatrix} a_1x + a_2y + a_3z \\ b_1x + b_2y + b_3z \end{pmatrix} \\ &= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

quindi $f(v) = Av$ con $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$. La matrice A è detta *matrice canonica* di f , e ha colonne $f(e_1), \dots, f(e_n)$.

Esempio a) Scrivere l'unica applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ tale che

$$f(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad f(e_3) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

b) Calcolare i seguenti vettori: $f \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, f \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$.

Soluzione. La matrice canonica di f è $A = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -2 & 1 \end{pmatrix}$. Dunque

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 3z \\ x - 2y + z \end{pmatrix}.$$

Si ha: $f \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, f \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$

In conclusione:

• Sia $f : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ un'applicazione lineare, e sia A la matrice 2×3 le cui colonne sono $f(e_1), f(e_2), f(e_3)$. A è detta la matrice canonica di f . Allora f si scrive:

$$f(v) = Av.$$

Viceversa, ogni applicazione del tipo precedente è lineare.

Quanto appena detto per le applicazioni lineari $\mathbf{R}^3 \rightarrow \mathbf{R}^2$ si generalizza in modo ovvio, per ottenere la seguente

Proposizione a) *Sia f un'applicazione lineare da \mathbf{R}^n a \mathbf{R}^m . Se A è la matrice $m \times n$ di colonne $f(e_1), \dots, f(e_n)$, detta la matrice canonica di f , allora f si scrive:*

$$f(v) = Av.$$

per ogni vettore colonna $v \in \mathbf{R}^n$.

b) *Viceversa, per ogni matrice A di tipo $m \times n$ l'applicazione $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ definita da $f(v) = Av$ è lineare.*

Quindi, le applicazioni lineari da \mathbf{R}^n a \mathbf{R}^m sono in corrispondenza biunivoca con le matrici $m \times n$.

Notiamo che un'applicazione $f : \mathbf{R}^n \rightarrow \mathbf{R}$ è lineare se e solo se è del tipo

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \dots + a_n x_n,$$

con coefficienti $a_1, \dots, a_n \in \mathbf{R}$. Una tale applicazione si dice *funzione lineare omogenea* di x_1, \dots, x_n . Più in generale, un'applicazione $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ è lineare se e solo se ogni

entrata del vettore immagine $f(v)$, dove $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, è una funzione lineare omogenea di x_1, \dots, x_n .

Esempio L'applicazione $f_1 : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ definita da:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 3y \\ x + y \end{pmatrix}$$

è lineare. Essa si scrive

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

e la sua matrice canonica è $\begin{pmatrix} 2 & -1 \\ 0 & 3 \\ 1 & 1 \end{pmatrix}$.

Esempio L'applicazione $g : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ definita da

$$g \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + z \\ 0 \\ x + 4z \end{pmatrix}$$

è lineare con matrice canonica

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 4 \end{pmatrix}.$$

Esempio L'applicazione $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definita da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y + 3 \\ 2x - y \end{pmatrix}$$

non è lineare, in quanto la prima entrata di $f \begin{pmatrix} x \\ y \end{pmatrix}$ non è una funzione lineare omogenea di x, y .

4 Omomorfismo assegnato su una base

Abbiamo visto che ogni applicazione lineare da \mathbf{R}^n a \mathbf{R}^m si può rappresentare con una matrice. Vedremo che ciò è vero per ogni applicazione lineare da uno spazio vettoriale V di dimensione n in un'altro spazio vettoriale W di dimensione n .

Iniziamo con l'osservare che un'applicazione lineare $f : V \rightarrow W$ è determinata dai valori che assume su una base di V .

Teorema *Sia V uno spazio vettoriale e $\mathcal{B} = (v_1, \dots, v_n)$ una sua base. Siano inoltre w_1, \dots, w_n vettori arbitrari di W . Allora esiste un'unica applicazione lineare $f : V \rightarrow W$ tale che:*

$$\begin{cases} f(v_1) = w_1 \\ f(v_2) = w_2 \\ \dots \\ f(v_n) = w_n \end{cases}.$$

In particolare, se f si annulla su una base allora $f = 0$.

Dimostrazione. Dimostriamo che tale applicazione esiste. Sia v un vettore di V , quindi: $v = x_1 v_1 + \dots + x_n v_n$. Definiamo $f(v)$ nel seguente modo:

$$f(v) = x_1 w_1 + \dots + x_n w_n.$$

Si dimostra che f è lineare, e soddisfa le condizioni.

Facciamo ora vedere che tale applicazione è unica. Se infatti g fosse una seconda applicazione lineare che verifica le condizioni date, cioè $f(v_i) = g(v_i) = w_i$ per ogni i , allora:

$$f(v) = x_1 w_1 + \dots + x_n w_n = g(v)$$

per ogni $v \in V$, quindi $f = g$. \square

Esempio Trovare l'unica applicazione lineare $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ tale che:

$$\begin{cases} f(e_1) = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ f(e_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{cases}$$

dove (e_1, e_2) è la base canonica di \mathbf{R}^2 .

Soluzione. Si ha $\begin{pmatrix} x \\ y \end{pmatrix} = xe_1 + ye_2$ quindi:

$$\begin{aligned} f\begin{pmatrix} x \\ y \end{pmatrix} &= f(xe_1 + ye_2) \\ &= xf(e_1) + yf(e_2) \\ &= x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 2x \\ 3x - y \end{pmatrix} \end{aligned}$$

In conclusione:

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3x - y \end{pmatrix}.$$

Potevamo anche procedere direttamente: sappiamo che i trasformati dei vettori della base canonica sono proprio le colonne della matrice canonica di f . Dunque la matrice canonica di f è:

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

e quindi $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3x - y \end{pmatrix}$. \square

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Trovare l'unica applicazione lineare $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ tale che $f(v_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $f(v_2) = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$.

Soluzione. Dobbiamo prima vedere quali valori assume la f sui vettori della base canonica. Esprimiamo i vettori della base canonica di \mathbf{R}^2 rispetto alla base (v_1, v_2) :

$$\begin{cases} e_1 = \frac{1}{2}v_1 + \frac{1}{2}v_2 \\ e_2 = \frac{1}{2}v_1 - \frac{1}{2}v_2 \end{cases}$$

Quindi:

$$\begin{cases} f(e_1) = \frac{1}{2}f(v_1) + \frac{1}{2}f(v_2) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \\ f(e_2) = \frac{1}{2}f(v_1) - \frac{1}{2}f(v_2) = \begin{pmatrix} -1 \\ -2 \end{pmatrix} \end{cases}$$

La matrice canonica di f è $\begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix}$ dunque

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 4x - 2y \end{pmatrix}.$$

Metodo alternativo: partiamo dalla generica applicazione lineare $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

Imponiamo le condizioni e otteniamo:

$$\begin{pmatrix} a + b \\ c + d \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} a - b \\ c - d \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix},$$

e risolvendo otteniamo $a = 2, b = -1, c = 4, d = -2$, quindi

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 4x - 2y \end{pmatrix}.$$

□

Se i vettori sui quali si assegna f non formano una base, potremmo avere dei problemi.

Esempio Siano $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. Dimostrare che non esiste alcuna applicazione lineare $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ tale che:

$$\begin{cases} f(v_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ f(v_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{cases}$$

Soluzione. Notiamo che $v_2 = 2v_1$ dunque i vettori di partenza non formano una base. Poiché f è lineare, si dovrebbe avere

$$f(v_2) = 2f(v_1) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

Ma questo contraddice la condizione $f(v_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Quindi una tale f non esiste. □

Esempio Nella notazione dell'esempio precedente, dimostrare che esistono infinite applicazioni lineari $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ tali che

$$\begin{cases} f(v_1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ f(v_2) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \end{cases}$$

5 Matrice associata

Sia $f : V \rightarrow W$ lineare, con V e W di dimensione finita, e fissiamo una base $\mathcal{B} = (v_1, \dots, v_n)$ di V e una base $\mathcal{B}' = (w_1, \dots, w_m)$ di W .

Con questi dati, definiamo la *matrice associata a f rispetto alle basi $\mathcal{B}, \mathcal{B}'$* con la regola seguente:

- La i -esima colonna di A è data dalle coordinate del vettore $f(v_i)$ rispetto alla base \mathcal{B}' , per $i = 1, \dots, n$.

Osseviamo che, se $\dim V = n$ e $\dim W = m$, allora A è di tipo $m \times n$. Ovviamente A dipende, oltre che da f , anche dalle basi scelte.

Esempio Sia $f : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{R}^2$ definita da

$$f \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + 2b \\ c - 3d \end{pmatrix}.$$

Trovare la matrice associata a f rispetto alla base canonica $(E_{11}, E_{12}, E_{21}, E_{22})$ di $\mathbf{Mat}(2 \times 2)$ e alla base canonica di \mathbf{R}^2 , rispettivamente.

Soluzione. Abbiamo:

$$f(E_{11}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, f(E_{12}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, f(E_{21}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, f(E_{22}) = \begin{pmatrix} 0 \\ -3 \end{pmatrix}.$$

Dunque la matrice associata è $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & -3 \end{pmatrix}$.

Esempio Calcoliamo la matrice associata all'omomorfismo "derivazione" $D : \mathbf{R}^3[x] \rightarrow \mathbf{R}^3[x]$ definito da $D(p(x)) = p'(x)$. Scegliamo la base canonica sia nello spazio di partenza che nello spazio di arrivo. La base canonica di $\mathbf{R}^3[x]$ è formata dai tre monomi

$$E_0(x) = 1, \quad E_1(x) = x, \quad E_2(x) = x^2.$$

Si ha quindi

$$\begin{cases} D(E_0(x)) = 0 \\ D(E_1(x)) = 1 = 1 \cdot E_0(x) + 0 \cdot E_1(x) + 0 \cdot E_2(x) \\ D(E_2(x)) = 2x = 0 \cdot E_0(x) + 2 \cdot E_1(x) + 0 \cdot E_2(x) \end{cases}$$

Incolonnando le coordinate, otteniamo la matrice associata

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

Esempio Sia $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ lineare. Allora la matrice associata a f rispetto alle basi canoniche è quella che abbiamo chiamato la *matrice canonica* di f .

Ecco la proprietà importante della matrice associata.

Teorema Siano $f : V \rightarrow W$ un'applicazione lineare, $\mathcal{B} = (v_1, \dots, v_n)$ una base di V e $\mathcal{B}' = (w_1, \dots, w_m)$ una base di W . Consideriamo la matrice associata a f rispetto alle basi scelte. Se $X = (x_1, \dots, x_n)^t$ è il vettore colonna delle coordinate del vettore v rispetto a \mathcal{B} , allora le coordinate di $f(v)$ sono date da AX .

Dimostrazione. È una verifica, che omettiamo. \square

In altre parole, le coordinate del vettore immagine $f(v)$ si ottengono moltiplicando la matrice associata ad f per il vettore colonna delle coordinate di v .

Esempio Consideriamo il polinomio $p(x) = 3 - x + x^2$, che ha coordinate $\begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$ rispetto alla base canonica. Sappiamo che la matrice associata all'omomorfismo derivazione, rispetto alle basi canoniche, è $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. Allora le coordinate di $p'(x)$ sono date da

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

Dunque $p'(x) = -1 + 4x$.

Esempio Siano V e W due spazi vettoriali di dimensione 3 e 2, rispettivamente. Fissiamo la base $\mathcal{B} = (v_1, v_2, v_3)$ di V e la base $\mathcal{B}' = (w_1, w_2)$ di W , e si consideri l'unico omomorfismo

$f : V \rightarrow W$ definito da:

$$\begin{cases} f(v_1) = 2w_1 + w_2 \\ f(v_2) = -4w_1 - 2w_2 \\ f(v_3) = 2w_1 + w_2 \end{cases}$$

- a) Scrivere la matrice associata a f rispetto alle basi scelte.
- b) Determinare $f(v_1 - v_2 + 3v_3)$ e $f(2v_1 + v_2)$.
- c) Determinare l'insieme di tutti i vettori $v \in V$ tali che $f(v) = O$ (vettore nullo di W).

Soluzione. a) Direttamente dalla definizione, la matrice associata è:

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

- b) Il vettore $v_1 - v_2 + 3v_3$ ha coordinate $X = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ rispetto alla base \mathcal{B} . Dunque le coordinate del suo trasformato sono:

$$AX = \begin{pmatrix} 2 & -4 & 2 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \end{pmatrix}.$$

Ne segue che $f(v_1 - v_2 + 3v_3) = 10w_1 + 5w_2$.

Analogamente si trova che

$$f(2v_1 + v_2) = O.$$

- c) Occorre risolvere il sistema:

$$AX = O,$$

che ha ∞^1 soluzioni $\begin{pmatrix} 2s - t \\ s \\ t \end{pmatrix}$ con $s, t \in \mathbf{R}$. Dunque i vettori cercati sono del tipo:

$$(2s - t)v_1 + sv_2 + tv_3.$$

□

6 Nucleo

Definizione Sia $f : V \rightarrow V'$ un'applicazione lineare. Il nucleo di f è il sottoinsieme di V costituito da tutti i vettori la cui immagine è il vettore nullo di V' .

Il nucleo di f si denota con il simbolo $\text{Ker} f$. Dunque:

$$\text{Ker} f = \{v \in V : f(v) = O\}.$$

Il nucleo è importante perchè caratterizza l'iniettività dell'applicazione.

Proposizione a) $\text{Ker} f$ è un sottospazio di V .

b) f è iniettiva se e solo se $\text{Ker} f = \{O\}$.

Dimostrazione. a) Dimostriamo che il nucleo è un sottospazio. Sappiamo già che $f(O) = O$, dunque $O \in V$. La linearità della f implica facilmente che il nucleo è chiuso rispetto alla somma e al prodotto per uno scalare.

b) Supponiamo che f sia iniettiva e facciamo vedere che $\text{Ker} f = \{O\}$. Se $v \in \text{Ker} f$, allora $f(v) = O$. D'altra parte $f(O) = O$; quindi $f(v) = f(O)$. Poichè f è iniettiva per ipotesi, si deve avere $v = O$. Dunque il nucleo contiene solamente il vettore nullo.

Supponiamo ora che $\text{Ker} f = \{O\}$ e facciamo vedere che f è iniettiva, verificando che

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

Infatti:

$$\begin{aligned} f(u) = f(v) &\implies f(u) - f(v) = O \\ &\implies f(u - v) = O \\ &\implies u - v \in \text{Ker} f \\ &\implies u - v = O \\ &\implies u = v. \end{aligned}$$

□

In pratica, per trovare una base del nucleo basta risolvere l'equazione vettoriale

$$f(v) = O$$

che si traduce in un sistema lineare omogeneo. Vedremo esempi più avanti.

Esercizio Sia f un'applicazione lineare iniettiva da V in V' , e siano v_1, \dots, v_k vettori linearmente indipendenti di V . Dimostrare che i vettori trasformati

$$f(v_1), \dots, f(v_k)$$

sono linearmente indipendenti. Cioè: *un'applicazione lineare iniettiva trasforma vettori linearmente indipendenti in vettori linearmente indipendenti*. (Attenzione: f deve essere iniettiva, altrimenti l'affermazione non è sempre vera).

7 Immagine

Data un'applicazione lineare $f : V \rightarrow V'$ consideriamo ora la sua immagine

$$\text{Im}f = \{v' \in V' : v' = f(v) \text{ per qualche } v \in V\}.$$

Dunque $\text{Im}f$ è un sottoinsieme dello spazio di arrivo V' . Si dimostra facilmente che

- $\text{Im}f$ è un sottospazio di V' .

Ricordiamo che f si dice suriettiva se $\text{Im}f = V'$.

Proposizione Sia $f : V^n \rightarrow V'^m$ un'applicazione lineare. Allora:

- f è suriettiva se e solo se $\dim(\text{Im}f) = m$.
- Se i vettori v_1, \dots, v_k generano V , allora i vettori $f(v_1), \dots, f(v_k)$ generano $\text{Im}f$, cioè:

$$\text{Im}f = L[f(v_1), \dots, f(v_k)].$$

Dimostrazione. a) f è suriettiva se e solo se $\text{Im}f = V'$. Poiché $\text{Im}f$ è un sottospazio di V' , questo avviene se e solo se $\dim \text{Im}f = \dim V' = m$.

b) Basta far vedere che vale la doppia inclusione:

$$\begin{aligned} \text{Im}f &\subseteq L[f(v_1), \dots, f(v_k)]; \\ L[f(v_1), \dots, f(v_k)] &\subseteq \text{Im}f \end{aligned}$$

Per la prima inclusione, mostriamo che ogni vettore dell'immagine è una combinazione lineare dei vettori $f(v_1), \dots, f(v_k)$. Se $v' \in \text{Im}f$, allora $v' = f(v)$ per qualche $v \in V$. Scriviamo

$$v = a_1 v_1 + \dots + a_k v_k;$$

quindi

$$\begin{aligned} v' = f(v) &= f(a_1 v_1 + \dots + a_k v_k) \\ &= a_1 f(v_1) + \dots + a_k f(v_k) \end{aligned}$$

La seconda inclusione è lasciata per esercizio. \square

In particolare, se $f : V^n \rightarrow V'^m$ è lineare, e se (v_1, \dots, v_n) è una base di V , allora:

$$\text{Im}f = L[f(v_1), \dots, f(v_n)],$$

quindi l'immagine è generata dai trasformati dei vettori di una qualunque base.

La proposizione seguente permette di calcolare la dimensione dell'immagine tramite il rango di una matrice associata.

Proposizione Sia $f : V \rightarrow V'$ un'applicazione lineare, e sia A la matrice associata a f rispetto alle basi $\mathcal{B} = (v_1, \dots, v_n)$ di V e $\mathcal{B}' = (v'_1, \dots, v'_n)$ di V' . Allora

$$\dim \operatorname{Im} f = \operatorname{rk} A.$$

Dimostrazione. Ricordiamo che, se E è il sottospazio di W generato dai vettori u_1, \dots, u_k , allora la dimensione di E uguaglia il rango della matrice ottenuta incolonnando le coordinate di u_1, \dots, u_k rispetto ad una qualunque base di W . Poiché

$$\operatorname{Im} f = L[f(v_1), \dots, f(v_n)],$$

e le colonne della matrice associata A sono proprio le coordinate di $f(v_1), \dots, f(v_n)$, la proposizione segue immediatamente. \square

Nel caso in cui $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ è lineare, possiamo considerare la matrice canonica di f (cioè la matrice associata a f rispetto alle basi canoniche). Allora:

- a) L'immagine di f è generata dai vettori colonna di A .
- b) La dimensione dell'immagine di f uguaglia il rango di A .

8 Esempi

Esempio Determinare una base del nucleo e una base dell'immagine dell'applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ definita da

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + 2z \\ x + y + 4z \end{pmatrix}.$$

Soluzione. Per trovare il nucleo, basta porre:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

quindi

$$\begin{cases} x - y + 2z = 0 \\ x + y + 4z = 0 \end{cases}$$

Questo è un sistema lineare omogeneo di due equazioni in tre incognite, con insieme delle soluzioni

$$\text{Ker } f = \left\{ \begin{pmatrix} -3t \\ -t \\ t \end{pmatrix}, t \in \mathbf{R} \right\}.$$

Quindi una base di $\text{Ker } f$ è $\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}$ e la sua dimensione è 1, in particolare f non è iniettiva.

Fissata la base canonica (e_1, e_2, e_3) di \mathbf{R}^3 , sappiamo che l'immagine è generata dai vettori $f(e_1), f(e_2), f(e_3)$, quindi

$$\text{Im } f = L \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right].$$

Chiaramente i tre generatori sono linearmente dipendenti; togliendo l'ultimo abbiamo però l'indipendenza lineare dei primi due, e una base di $\text{Im } f$ è dunque

$$\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right).$$

La dimensione dell'immagine è 2, pari alla dimensione dello spazio di arrivo: ciò significa che $\text{Im } f = \mathbf{R}^2$ e quindi f è suriettiva.

Per la dimensione dell'immagine, bastava anche usare la matrice canonica di f :

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 1 & 4 \end{pmatrix},$$

che ha rango 2. Dunque $\dim \text{Im } f = 2$ e quindi $\text{Im } f = \mathbf{R}^2$.

Notiamo che, nell'esempio:

$$\dim \text{Ker } f + \dim \text{Im } f = 3$$

e 3 è proprio la dimensione dello spazio di partenza.

Esempio Sia $D : \mathbf{R}^4[x] \rightarrow \mathbf{R}^4[x]$ l'applicazione lineare che associa a un polinomio la sua derivata:

$$D(p(x)) = p'(x).$$

Determinare basi del nucleo e dell'immagine di D .

Soluzione. Per il nucleo, imponiamo $D(p(x)) = 0$. La derivata di un polinomio è nulla se e solo se il polinomio è una costante; dunque $\text{Ker} D$ è il sottospazio formato dai polinomi costanti. Tale sottospazio ha dimensione 1, ed è generato dal polinomio costante $E_0(x) = 1$. Consideriamo la base canonica di $\mathbf{R}^4[x]$:

$$(1, x, x^2, x^3).$$

Dunque $\text{Im} f$ è generata dai trasformati:

$$D(1) = 0, D(x) = 1, D(x^2) = 2x, D(x^3) = 3x^2.$$

Si vede facilmente che $\text{Im} f$ è il sottospazio formato dai polinomi di grado minore o uguale a 2. Dunque $\dim \text{Im} f = 3$. Si poteva anche procedere con una matrice associata.

Esempio Sia $D_2 : \mathbf{R}^4[x] \rightarrow \mathbf{R}^4[x]$ l'applicazione lineare che associa a un polinomio la sua derivata seconda:

$$D_2(p(x)) = p''(x).$$

Determinare basi del nucleo e dell'immagine di D .

Soluzione. Si vede facilmente che il nucleo è formato dai polinomi di grado minore o uguale a 1, e che l'immagine è uguale al nucleo.

9 Teorema della dimensione

Per un'applicazione lineare da uno spazio vettoriale V^n in un altro spazio vettoriale V' esiste un legame fra la dimensione del nucleo e quella dell'immagine. Tale relazione è data dal teorema seguente, noto come *teorema della dimensione*.

Teorema Sia $f : V^n \rightarrow V'$ un'applicazione lineare. Allora

$$\dim \text{Ker} f + \dim \text{Im} f = n$$

dove n è la dimensione dello spazio di partenza.

Dimostrazione. Fissiamo una base \mathcal{B} di V e una base \mathcal{B}' di V' e consideriamo la matrice A , associata a f rispetto a tali basi. Se v è un vettore di V , e $X \in \mathbf{R}^n$ è il vettore colonna delle coordinate di v rispetto a \mathcal{B} , sappiamo che le coordinate di $f(v)$ sono date dal vettore colonna $AX \in \mathbf{R}^m$. Dunque l'equazione che definisce il nucleo:

$$f(v) = 0$$

si traduce nel sistema lineare omogeneo, di m equazioni in n incognite

$$S : AX = 0.$$

Si ha dunque che

$$\dim \operatorname{Ker} f = \dim \operatorname{Sol}(S).$$

Dal teorema di Rouché-Capelli sappiamo che la dimensione di $\operatorname{Sol}(S)$ è uguale a $n - \operatorname{rk} A$.
Dunque:

$$\begin{aligned} \dim \operatorname{Ker} f &= n - \operatorname{rk} A \\ &= n - \dim \operatorname{Im} f \end{aligned}$$

e il teorema è dimostrato. \square

Esempio Si consideri la matrice

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

- a) Scrivere esplicitamente l'applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ rappresentata da A rispetto alle basi canoniche.
- b) Determinare una base e la dimensione di $\operatorname{Ker} f$.
- c) Determinare una base e la dimensione di $\operatorname{Im} f$.

Soluzione. a) Poiché A è la matrice canonica di f , abbiamo che

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ -x - 2y - z \\ 2x + 4y + 2z \end{pmatrix}$$

- b) Il nucleo si ottiene imponendo $f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Otteniamo un sistema lineare omogeneo la cui unica equazione significativa è

$$x + 2y + z = 0.$$

Risolvendo, otteniamo

$$\operatorname{Ker} f = \left\{ \begin{pmatrix} -2s - t \\ s \\ t \end{pmatrix} : t, s \in \mathbf{R} \right\}.$$

Poiché

$$\begin{pmatrix} -2s-t \\ s \\ t \end{pmatrix} = t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix},$$

otteniamo la base di $\text{Ker } f$:

$$\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right),$$

e dunque $\dim \text{Ker } f = 2$.

b) Dal teorema della dimensione abbiamo che $\dim \text{Im } f = 1$: dunque una base di $\text{Im } f$ è data da un suo qualunque vettore non nullo, ad esempio

$$f(e_1) = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

In effetti, sappiamo che $\text{Im } f$ è generata dalle colonne della sua matrice canonica A :

$$\text{Im } f = L \left[\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right],$$

e tali generatori sono tutti multiplo del primo, che è dunque una base.

Notiamo che f non è né iniettiva né suriettiva. \square

Esempio Determinare nucleo e immagine dell'applicazione lineare $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ rappresentata dalla matrice $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix}$ rispetto alle basi canoniche.

Soluzione. Il rango della matrice associata è 3, di conseguenza

$$\dim \text{Im } f = 3$$

e $\text{Im } f = \mathbf{R}^3$ (spazio di arrivo). Dunque f è suriettiva. Per il teorema della dimensione, si ha

$$\dim \text{Ker } f = 0,$$

dunque f è anche iniettiva. In conclusione, f è biiettiva. \square

Esempio Data l'applicazione lineare $f : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{Mat}(2 \times 2)$ definita da

$$f \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+d & b-c \\ b-c & a+d \end{pmatrix}$$

determinare basi di $\text{Ker } f$ e $\text{Im } f$.

Soluzione. Il nucleo si ottiene imponendo $f(A) = O$. Dunque

$$\begin{pmatrix} a+d & b-c \\ b-c & a+d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

che dà luogo al sistema

$$\begin{cases} a+d=0 \\ b-c=0 \end{cases}.$$

La matrice generica del nucleo è dunque

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

e una base del nucleo, la cui dimensione è pari a 2, è

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

Consideriamo la base canonica di $\mathbf{Mat}(2 \times 2)$, che scriveremo:

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sappiamo che $\text{Im } f$ è generata dalle matrici:

$$f(E_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, f(E_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, f(E_3) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, f(E_4) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Per il teorema della dimensione, la dimensione dell'immagine è pari a:

$$\dim \text{Im } f = \dim(\mathbf{Mat}(2 \times 2)) - \dim \text{Ker } f = 4 - 2 = 2,$$

per cui basta selezionare 2 vettori generatori linearmente indipendenti, ad esempio il primo e il secondo. Dunque una base dell'immagine è

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

Osserviamo che

$$\begin{cases} f(E_1) = E_1 + E_4 \\ f(E_2) = E_2 + E_3 \\ f(E_3) = -E_2 - E_3 \\ f(E_4) = E_1 + E_4 \end{cases}$$

Dunque la matrice associata a f rispetto alle basi canoniche è:

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

In effetti, $\dim \operatorname{Im} f = \operatorname{rk} A = 2$.

Esempio Siano V uno spazio vettoriale con base $\mathcal{B} = (v_1, v_2, v_3)$ e W un secondo spazio vettoriale con base $\mathcal{B}' = (w_1, w_2, w_3)$. Si consideri l'unica applicazione lineare $f : V \rightarrow W$ tale che:

$$\begin{cases} f(v_1) = w_1 + w_3 \\ f(v_2) = 2w_1 + w_2 + 3w_3 \\ f(v_3) = w_1 + w_2 + 2w_3 \end{cases}$$

- Scrivere la matrice associata a f rispetto alle basi $\mathcal{B}, \mathcal{B}'$.
- Determinare la dimensione di $\operatorname{Ker} f$.
- Determinare la dimensione di $\operatorname{Im} f$.

Soluzione. Matrice associata $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 3 & 2 \end{pmatrix}$. Si vede che $\operatorname{rk} A = 2$, da cui $\dim \operatorname{Im} f = 2$ con base $f(v_1), f(v_2)$. Dal teorema della dimensione otteniamo subito che $\dim \operatorname{Ker} f = 1$.
□

Proposizione Sia $f : V^n \rightarrow V^m$ lineare.

- Se f è iniettiva, allora $n \leq m$.
- Se f è suriettiva, allora $n \geq m$.
- Se f è biiettiva, allora $n = m$.

Dimostrazione. a) È una conseguenza diretta del teorema della dimensione. Infatti, se f è iniettiva la dimensione del nucleo è zero, dunque

$$\dim \operatorname{Im} f = n.$$

D'altra parte, l'immagine di f è un sottospazio di V'^m : dunque

$$\dim \operatorname{Im} f \leq m.$$

In conclusione abbiamo necessariamente

$$n \leq m.$$

L'affermazione b) si dimostra in modo analogo. c) è conseguenza di a) e b). \square

10 Isomorfismi

Un'applicazione lineare iniettiva e suriettiva (quindi biiettiva) si dice un *isomorfismo*. Due spazi vettoriali V, V' si dicono *isomorfi* se esiste almeno un isomorfismo $f : V \rightarrow V'$.

Esempio L'applicazione $T : \mathbf{Mat}(m \times n) \rightarrow \mathbf{Mat}(n \times m)$ definita da

$$T(A) = A^t$$

per ogni $A \in \mathbf{Mat}(m \times n)$ è lineare, iniettiva e suriettiva (come si verifica subito) dunque è un isomorfismo. Ne segue che $\mathbf{Mat}(m \times n)$ è isomorfo a $\mathbf{Mat}(n \times m)$.

La proposizione seguente mostra che spazi vettoriali isomorfi hanno la stessa struttura; un isomorfismo conserva infatti la proprietà di indipendenza lineare, e la proprietà di generare lo spazio.

Proposizione Sia $f : V \rightarrow V'$ un isomorfismo di spazi vettoriali.

- a) I vettori $v_1, \dots, v_k \in V$ sono linearmente indipendenti se e solo se i trasformati $f(v_1), \dots, f(v_k) \in V'$ sono linearmente indipendenti.
- b) I vettori v_1, \dots, v_k generano V se e solo se i trasformati $f(v_1), \dots, f(v_k)$ generano V' .
- c) f trasforma basi di V in basi di V' .
- d) $\dim V = \dim V'$.

Dimostrazione. a) Supponiamo che v_1, \dots, v_k siano vettori linearmente indipendenti di V , e che si abbia

$$a_1 f(v_1) + \dots + a_k f(v_k) = O.$$

Poiché f è lineare, otteniamo

$$f(a_1 v_1 + \dots + a_k v_k) = O,$$

e per definizione $a_1v_1 + \dots + a_kv_k \in \text{Ker } f$. Per ipotesi, f è iniettiva, dunque $\text{Ker } f = \{O\}$. Di conseguenza:

$$a_1v_1 + \dots + a_kv_k = O$$

e poiché v_1, \dots, v_k sono linearmente indipendenti otteniamo $a_1 = \dots = a_k = 0$. Dunque $f(v_1), \dots, f(v_k)$ sono linearmente indipendenti.

In modo analogo si dimostra che se $f(v_1), \dots, f(v_k)$ sono linearmente indipendenti allora v_1, \dots, v_k sono linearmente indipendenti.

b) Supponiamo che i vettori v_1, \dots, v_k generino V . Allora sappiamo che $f(v_1), \dots, f(v_k)$ generano $\text{Im } f$. Poiché f è suriettiva, si ha $\text{Im } f = V'$ dunque tali vettori generano anche V' .

Viceversa, supponiamo che $f(v_1), \dots, f(v_k)$ generino V' , e sia $v \in V$. Allora, poiché $f(v) \in V'$ possiamo scrivere

$$f(v) = a_1f(v_1) + \dots + a_kf(v_k),$$

con $a_1, \dots, a_k \in \mathbf{R}$. Poiché f è lineare:

$$f(v) = f(a_1v_1 + \dots + a_kv_k)$$

e, dall'iniettività di f , otteniamo

$$v = a_1v_1 + \dots + a_kv_k.$$

Ogni vettore di V è combinazione lineare di v_1, \dots, v_k , che dunque generano V .

c) Se (v_1, \dots, v_k) è una base di V , allora i vettori $f(v_1), \dots, f(v_k)$ sono linearmente indipendenti per la parte a) e generano V' per la parte b): dunque $f(v_1), \dots, f(v_k)$ formano una base di V' .

d) È immediata dalla parte c). \square

Infine, osserviamo che un isomorfismo da V in V' trasforma un sottospazio E di V in un sottospazio E' di V' , isomorfo a E .

10.1 Uso delle coordinate

Sia V^n un qualunque spazio vettoriale di dimensione n . Fissiamo una base $\mathcal{B} = (v_1, \dots, v_n)$ di V e consideriamo l'applicazione $\mathcal{F} : V^n \rightarrow \mathbf{R}^n$ definita da:

$$\mathcal{F}(v) = \text{coordinate di } v \text{ rispetto a } \mathcal{B}.$$

Esplicitamente, se v si esprime, nella base scelta, come combinazione lineare $v = a_1v_1 + \dots + a_kv_k$, allora

$$\mathcal{F}(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}.$$

Si verifica facilmente che \mathcal{F} è lineare; inoltre \mathcal{F} è iniettiva (vettori distinti hanno coordinate distinte) e suriettiva (per la verifica, applicare il teorema della dimensione a \mathcal{F} , sapendo che $\dim \text{Ker} \mathcal{F} = 0$). Dunque \mathcal{F} è un isomorfismo. In particolare abbiamo dimostrato il seguente risultato.

Teorema *Ogni spazio vettoriale di dimensione n è isomorfo a \mathbf{R}^n .*

Esempio Lo spazio vettoriale dei polinomi di grado minore di n , denotato con $\mathbf{R}^n[x]$, è isomorfo a \mathbf{R}^n . Un isomorfismo $\mathcal{F} : \mathbf{R}^n[x] \rightarrow \mathbf{R}^n$ si ottiene prendendo le coordinate rispetto alla base canonica di $\mathbf{R}^n[x]$:

$$\mathcal{F}(a_0 + a_1x + \dots + a_{n-1}x^{n-1}) = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}.$$

Esempio Lo spazio vettoriale $\mathbf{Mat}(p \times q)$ ha dimensione pq , ed è dunque isomorfo a \mathbf{R}^{pq} .

Possiamo anche dire che \mathbf{R}^n serve come prototipo di tutti gli spazi vettoriali di dimensione n . Usando le coordinate, possiamo trasferire un dato problema da uno spazio vettoriale V^n a \mathbf{R}^n , dove il problema può essere risolto con l'aiuto delle matrici, del determinante e del rango.

Un esempio è dato dalla seguente proposizione, già enunciata, ma non dimostrata, nella Parte 5.

Proposizione *Sia V^n uno spazio vettoriale con base \mathcal{B} . Allora i vettori $v_1, \dots, v_k \in V$ sono linearmente indipendenti se e solo se il rango della matrice di colonne $\mathcal{F}(v_1), \dots, \mathcal{F}(v_k)$ è uguale a k .*

Dimostrazione. \mathcal{F} è un isomorfismo; dunque, dalla proposizione del paragrafo precedente, i vettori $v_1, \dots, v_k \in V$ sono linearmente indipendenti se e solo se le rispettive coordinate $\mathcal{F}(v_1), \dots, \mathcal{F}(v_k)$ sono vettori linearmente indipendenti di \mathbf{R}^n . Per il criterio del rango, questo avviene se e solo se il rango della matrice di colonne $\mathcal{F}(v_1), \dots, \mathcal{F}(v_k)$ è uguale a k . \square

Parte 7. Autovettori e autovalori

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Endomorfismi, 1
- 2 Cambiamento di base, 3
- 3 Matrici simili, 6
- 4 Endomorfismi diagonalizzabili, 7
- 5 Autovettori e autovalori, 7
- 6 Il polinomio caratteristico, 11
- 7 Calcolo degli autospazi, 15
- 8 Primo criterio, 18
- 9 Secondo criterio, 23
- 10 Matrici diagonalizzabili, 30

1 Endomorfismi

- Se V è uno spazio vettoriale, un *endomorfismo* di V è semplicemente un' applicazione lineare $f : V \rightarrow V$. Un endomorfismo di V è spesso chiamato *operatore* di V .

Possiamo ad esempio definire l'*endomorfismo identità*, denotato con $I : V \rightarrow V$, che associa a ogni vettore di V il vettore stesso:

$$I(v) = v.$$

- Fissata una base \mathcal{B} di V diremo *matrice associata a f rispetto a \mathcal{B}* la matrice associata all'applicazione lineare f prendendo la base \mathcal{B} sia nello spazio di partenza che in quello di arrivo.

La matrice associata dipende dalla scelta della base; poiché uno spazio vettoriale non nullo ammette infinite basi diverse, uno stesso endomorfismo ammetterà (tranne rari casi) infinite matrici associate, tutte diverse tra loro. Studieremo il modo di trovare la base più conveniente, rispetto alla quale la matrice associata assuma una forma particolarmente semplice.

Esempio Consideriamo l'endomorfismo f di \mathbf{R}^3 definito da:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7x - 13y + 6z \\ 2x - 8y + 6z \\ 2x - 13y + 11z \end{pmatrix}.$$

La matrice associata a f rispetto alla base canonica è $A = \begin{pmatrix} 7 & -13 & 6 \\ 2 & -8 & 6 \\ 2 & -13 & 11 \end{pmatrix}$.

Cambiamo ora base, e consideriamo la base (v_1, v_2, v_3) di \mathbf{R}^3 , dove

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 13 \\ 2 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.$$

Risulta $f \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $f \begin{pmatrix} 13 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 65 \\ 10 \\ 0 \end{pmatrix}$, $f \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -15 \\ 0 \\ 5 \end{pmatrix}$. Dunque:

$$\begin{cases} f(v_1) = 0 \\ f(v_2) = 5v_2 \\ f(v_3) = 5v_3 \end{cases}$$

e la matrice associata è $A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, diagonale. È chiaro che la matrice A' è molto

più semplice; per studiare l'endomorfismo f la base (v_1, v_2, v_3) è più conveniente della base canonica. \square

Esempio Sia f l'endomorfismo di \mathbf{R}^2 definito da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x + y \end{pmatrix}.$$

Determinare le matrici A, A', A'' associate a f :

a) rispetto alla base canonica;

b) rispetto alla base $\mathcal{B}' = (v'_1, v'_2)$ dove $v'_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$;

c) rispetto alla base $\mathcal{B}'' = (v''_1, v''_2)$ dove $v''_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $v''_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

Soluzione. a) A è semplicemente la matrice canonica di f : $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

b) Si ha:

$$\begin{cases} f(v'_1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \cdot v'_1 + 0 \cdot v'_2 \\ f(v'_2) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 0 \cdot v'_1 + 2 \cdot v'_2 \end{cases}$$

Dunque $A' = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$.

c) Si ha:

$$\begin{cases} f(v''_1) = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = -3 \cdot v''_1 + 3 \cdot v''_2 \\ f(v''_2) = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = -5 \cdot v''_1 + 5 \cdot v''_2 \end{cases}$$

Dunque $A'' = \begin{pmatrix} -3 & -5 \\ 3 & 5 \end{pmatrix}$.

Le matrici associate sono quindi, rispettivamente

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, A' = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, A'' = \begin{pmatrix} -3 & -5 \\ 3 & 5 \end{pmatrix}.$$

Per vedere qual'è la relazione fra le diverse matrici associate dobbiamo prima studiare la relazione che intercorre fra due basi di uno stesso spazio vettoriale.

2 Cambiamento di base

Sia V uno spazio vettoriale e siano $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{B}' = (v'_1, \dots, v'_n)$ due basi di V . Ogni vettore della base \mathcal{B}' si esprimerà dunque come combinazione lineare dei vettori della base \mathcal{B} :

$$\begin{cases} v'_1 = a_{11}v_1 + a_{21}v_2 + \dots + a_{n1}v_n \\ v'_2 = a_{12}v_1 + a_{22}v_2 + \dots + a_{n2}v_n \\ \dots \\ v'_n = a_{1n}v_1 + a_{2n}v_2 + \dots + a_{nn}v_n \end{cases} \quad (1)$$

La matrice ottenuta incolonnando le coordinate:

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

è detta *matrice del cambiamento di base* (o matrice di passaggio) da \mathcal{B} a \mathcal{B}' . La i -esima colonna di M è dunque data dalle coordinate del vettore v'_i rispetto a \mathcal{B} , per ogni $i = 1, \dots, n$.

La matrice M è evidentemente $n \times n$.

Le relazioni in (1) si esprimono, in forma compatta:

$$(v'_1, \dots, v'_n) = (v_1, \dots, v_n)M \quad (2)$$

dove, a destra, si intende il prodotto del vettore riga (v_1, \dots, v_n) (le cui entrate sono vettori) per la matrice M . Si scriverà anche

$$\mathcal{B}' = \mathcal{B}M.$$

Proposizione a) *La matrice di un cambiamento di base è invertibile.*

b) *Viceversa, sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V^n , e $M = \{a_{ij}\}$ una matrice invertibile. Allora i vettori v'_1, \dots, v'_n definiti dalle relazioni in (1) formano una base di V^n .*

c) *Se $\mathcal{B}, \mathcal{B}'$ sono due basi, la matrice di passaggio da \mathcal{B}' a \mathcal{B} è l'inversa della matrice di passaggio da \mathcal{B} a \mathcal{B}' .*

Dimostrazione. a) Il rango della matrice di passaggio M è uguale alla dimensione del sottospazio di V^n generato dai vettori v'_1, \dots, v'_n . Poiché questi vettori per ipotesi formano una base di V^n , tale sottospazio è tutto V^n , quindi il rango vale n e la matrice è invertibile.

b) Per ipotesi, il rango della matrice M è n : dunque il sottospazio generato dai vettori v'_1, \dots, v'_n ha dimensione n e coincide con V^n . Ciò significa che v'_1, \dots, v'_n sono $n = \dim V^n$ vettori generatori e dunque formano una base.

c) L'affermazione si dimostra moltiplicando ambo i membri della relazione (2), a destra, per l'inversa M^{-1} . \square

Esempio Fissiamo $V = \mathbf{R}^2$, e siano $\mathcal{B} = \mathcal{BC} = (e_1, e_2)$ la base canonica e \mathcal{B}' la base (w_1, w_2) dove $w_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $w_2 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$. Poiché

$$\begin{cases} w_1 = e_1 + 3e_2 \\ w_2 = 4e_1 + 5e_2 \end{cases} \quad (3)$$

la matrice di passaggio è

$$M = \begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix}.$$

Notiamo che le relazioni (3) si esprimono in forma matriciale:

$$(w_1, w_2) = (e_1, e_2) \begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix},$$

dove a destra si è moltiplicato il vettore riga (e_1, e_2) (le cui entrate sono vettori) per la matrice $\begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix}$. \square

Dalla definizione segue facilmente che

Osservazione Sia $V = \mathbf{R}^n$. La matrice di passaggio dalla base canonica (e_1, \dots, e_n) alla base (w_1, \dots, w_n) si ottiene semplicemente incolonnando i vettori w_1, \dots, w_n .

Esempio Consideriamo la base $\mathcal{B} = (v_1, v_2, v_3)$ formata dai vettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

e la base $\mathcal{B}' = (w_1, w_2, w_3)$, dove

$$w_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}, w_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Determinare la matrice di passaggio da \mathcal{B} a \mathcal{B}' .

Soluzione. Abbiamo le relazioni

$$\begin{cases} w_1 = v_1 - v_3 \\ w_2 = 2v_1 + v_2 \\ w_3 = v_1 + 2v_2 - v_3 \end{cases}$$

e quindi

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

3 Matrici simili

Diremo che due matrici quadrate A, A' sono *simili* se esiste una matrice invertibile M tale che

$$A' = M^{-1}AM.$$

Risulta che matrici associate ad uno stesso endomorfismo (rispetto a basi diverse) sono simili.

Teorema *Sia f un endomorfismo di uno spazio vettoriale V , e siano \mathcal{B} e \mathcal{B}' due basi di V . Se A è la matrice associata a f rispetto a \mathcal{B} , e A' è la matrice associata a f rispetto a \mathcal{B}' , allora A e A' sono simili. Precisamente,*

$$A' = M^{-1}AM,$$

dove M è la matrice di passaggio da \mathcal{B} a \mathcal{B}' .

Dimostrazione. È una verifica diretta, che omettiamo. \square

Esempio In uno degli esempi precedenti, abbiamo visto che l'endomorfismo f di \mathbf{R}^2 definito da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x + y \end{pmatrix}.$$

si rappresenta con la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ rispetto alla base canonica \mathcal{BC} , e con la matrice $A' = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ rispetto alla base $\mathcal{B}' = (v'_1, v'_2)$ dove $v'_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Per il teorema, le due matrici sono simili. In effetti si ha

$$A' = M^{-1}AM$$

dove $M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ è la matrice di passaggio da \mathcal{BC} a \mathcal{B}' .

Osservazione Vale anche il viceversa del teorema precedente: date due matrici simili, diciamo A e A' , allora esse rappresentano uno stesso endomorfismo. Ad esempio, se f è l'endomorfismo di \mathbf{R}^n rappresentato da A rispetto alla base canonica, e se \mathcal{B}' è la base di \mathbf{R}^n tale che la matrice di passaggio dalla base canonica a \mathcal{B}' è M , allora A' rappresenta f nella base \mathcal{B}' .

4 Endomorfismi diagonalizzabili

• Un endomorfismo di uno spazio vettoriale V si dice *diagonalizzabile* se puo' essere rappresentato da una matrice diagonale; in altre parole, se esiste una base di V rispetto alla quale la matrice associata è diagonale.

Esempio Sia $f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ definito da

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9x - 12y + 3z \\ 4x - 7y + 3z \\ 4x - 12y + 8z \end{pmatrix}.$$

La matrice associata rispetto alla base canonica è $A = \begin{pmatrix} 9 & -12 & 3 \\ 4 & -7 & 3 \\ 4 & -12 & 8 \end{pmatrix}$: non è diagonale.

Pero' possiamo trovare una base piu' fortunata. Siano infatti

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix},$$

che formano una base di \mathbf{R}^3 . Un calcolo mostra che:

$$\begin{cases} f(v_1) = 0 \\ f(v_2) = 5v_2 \\ f(v_3) = 5v_3 \end{cases}.$$

Dunque la matrice associata a f rispetto alla base $\mathcal{B}' = (v_1, v_2, v_3)$ è $A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$,

diagonale. Per definizione, l'endomorfismo f è diagonalizzabile.

Notiamo che i vettori della base "buona" hanno tutti la seguente proprietà: il trasformato del vettore è un multiplo del vettore stesso. Tale proprietà caratterizza quelli che saranno chiamati *autovettori* di f .

Studieremo il seguente problema:

• dato un endomorfismo, stabilire se esso è diagonalizzabile, e trovare eventualmente una base rispetto alla quale la matrice associata è diagonalizzabile.

5 Autovettori e autovalori

5.1 Definizione

Definizione Sia f un endomorfismo di uno spazio vettoriale V .

a) Un vettore $v \neq O$ si dice autovettore di f associato all'autovalore $\lambda \in \mathbf{R}$ se

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

b) Uno scalare λ si dice autovalore di f se esiste un vettore $v \neq O$ tale che $f(v) = \lambda v$.

Esempio Sia f l'endomorfismo f di \mathbf{R}^2 definito da $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix}$, e siano $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Un calcolo mostra che

$$\begin{cases} f(v_1) = O = 0v_1 \\ f(v_2) = 2v_2 \end{cases}$$

Allora, per definizione, v_1 è un autovettore di f associato all'autovalore $\lambda = 0$, e v_2 è un autovettore associato all'autovalore $\lambda = 2$.

Notiamo che i due autovettori formano una base (v_1, v_2) di \mathbf{R}^2 , e che la matrice associata a tale base è diagonale $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ con elementi diagonali dati esattamente dagli autovalori.

Ricordiamo che, per definizione, un autovettore di un endomorfismo è, per definizione, *non nullo*. Al contrario, lo scalare 0 può essere un autovalore di f . In effetti, osserviamo che

- 0 è un autovalore di f se e solo se $\text{Ker } f \neq \{O\}$. Ogni vettore del nucleo, diverso dal vettore nullo, è un autovettore con autovalore 0.

5.2 Caratterizzazione degli endomorfismi diagonalizzabili

Supponiamo ora che ci sia una base di V , diciamo (v_1, \dots, v_n) , formata da autovettori di f . Allora esistono autovalori $\lambda_1, \dots, \lambda_n$ (non necessariamente tutti distinti) tali che

$$\begin{cases} f(v_1) = \lambda_1 v_1 \\ f(v_2) = \lambda_2 v_2 \\ \dots \\ f(v_n) = \lambda_n v_n. \end{cases}$$

Per definizione, la matrice associata a f rispetto a tale base di autovettori è diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

con elementi diagonali dati dagli autovalori. Viceversa, se la matrice associata a f rispetto a una data base (v_1, \dots, v_n) è diagonale, allora i vettori di tale base sono autovettori di f associati, rispettivamente, agli elementi diagonali $\lambda_1, \dots, \lambda_n$. In conclusione, abbiamo dimostrato il seguente risultato.

Teorema *Un endomorfismo f di uno spazio vettoriale V è diagonalizzabile se e solo se V ammette una base formata da autovettori di f .*

Osserviamo i seguenti esempi banali di endomorfismi diagonalizzabili.

Esempio L'endomorfismo nullo $O : V \rightarrow V$ ha matrice associata nulla rispetto a una qualunque base, dunque è banalmente diagonalizzabile. Gli autovalori sono tutti uguali a zero.

Esempio L'endomorfismo identità $I : V \rightarrow V$ è diagonalizzabile, poiché ha matrice associata data, appunto, dalla matrice identità (rispetto a una qualunque base). Gli autovalori sono tutti uguali a 1.

Osserviamo ora che *non tutti* gli endomorfismi sono diagonalizzabili; anzi, ci sono endomorfismi che non ammettono autovettori (e quindi non ammettono autovalori).

Esempio Consideriamo l'endomorfismo di \mathbf{R}^2 definito da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix},$$

con matrice canonica $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Supponiamo per assurdo che $\begin{pmatrix} x \\ y \end{pmatrix}$ sia un autovettore di f con autovalore λ :

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

con $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Allora si avrebbe

$$\begin{cases} \lambda x = y \\ \lambda y = -x \end{cases}$$

Moltiplichiamo la prima equazione per x , la seconda per y e sommiamo. Otteniamo la relazione

$$\lambda(x^2 + y^2) = 0,$$

da cui $\lambda = 0$ oppure $x^2 + y^2 = 0$. Ora, nessuno dei casi si può verificare, perché altrimenti $x = y = 0$. Dunque f non ha né autovettori, né autovalori. \square

5.3 Autospazio associato a un autovalore

È facile verificare che, se λ è un autovalore, il sottoinsieme di V

$$E(\lambda) = \{v \in V : f(v) = \lambda v\}$$

è un sottospazio di V , detto *autospazio associato a λ* . La sua dimensione è detta anche *molteplicità geometrica* di λ , e si denota con $MG(\lambda)$. Quindi:

$$MG(\lambda) = \dim E(\lambda).$$

Notiamo che $E(\lambda)$ è formato dal vettore nullo, e da tutti gli autovettori associati a λ . Se λ è un autovalore, allora per definizione esiste almeno un vettore non nullo nel sottospazio $E(\lambda)$. Dunque, se λ è un autovalore si ha sempre

$$MG(\lambda) \geq 1.$$

Per definizione, si ha inoltre:

$$E(0) = \text{Ker } f.$$

Esempio L'endomorfismo $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix}$ di \mathbf{R}^2 ha autovalori $\lambda_1 = 0$ e $\lambda_2 = 2$. L'autospazio $E(0)$ associato a 0 è il nucleo di f , di equazione $x + y = 0$, dunque

$$\dim E(0) = 1 \quad \text{con base} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

L'autospazio $E(2)$ associato all'autovalore 2 è definito dall'equazione

$$f \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix};$$

Con semplici calcoli, si vede che $E(2)$ è descritto dall'equazione $x - y = 0$. Dunque

$$\dim E(2) = 1 \quad \text{con base} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Entrambi gli autovalori hanno molteplicità geometrica 1. \square

Consideriamo ora il seguente problema:

- dato l'endomorfismo f , determinare tutti gli autovalori e gli autospazi di f .

Vedremo che, se f è un endomorfismo di uno spazio vettoriale V di dimensione n , allora gli autovalori sono le radici di un certo polinomio di grado n , detto il *polinomio caratteristico* di f .

6 Il polinomio caratteristico

In ciò che segue, $V = V^n$ è uno spazio vettoriale di dimensione n .

Proposizione *Sia f un endomorfismo di V^n e sia A una qualunque matrice associata a f . Allora $\lambda \in \mathbf{R}$ è un autovalore di f se e solo se*

$$\det(A - \lambda I) = 0,$$

dove I è la matrice identità.

Dimostrazione. Supponiamo che λ sia un autovalore di f : allora esiste un vettore $v \neq 0$ tale che $f(v) = \lambda v$. Fissiamo una qualunque base \mathcal{B} di V e consideriamo la matrice A , associata a f rispetto a \mathcal{B} . Sia $X \in \mathbf{R}^n$ il vettore colonna delle coordinate di v (notiamo che $X \neq 0$ per ipotesi). Allora sappiamo che AX è il vettore colonna delle coordinate di $f(v)$. Dunque

$$AX = \lambda X.$$

Poiché $X = IX$ l'equazione si scrive:

$$(A - \lambda I)X = 0. \tag{4}$$

Tale equazione equivale a un sistema lineare omogeneo di n equazioni in n incognite, con matrice dei coefficienti data da $A - \lambda I$. Per ipotesi, $X \neq 0$ è un autosoluzione del sistema: dunque $\text{rk}(A - \lambda I) < n$, cioè $\det(A - \lambda I) = 0$.

Viceversa, se $\lambda \in \mathbf{R}$ è tale che $\det(A - \lambda I) = 0$ possiamo invertire il ragionamento e trovare un'autosoluzione X del sistema (4): a X corrisponde un vettore non nullo $v \in V$ tale che $f(v) = \lambda v$ e quindi λ risulta un autovalore di f . \square

Consideriamo la funzione di x :

$$p_A(x) = \det(A - xI).$$

Si verifica facilmente che $p_A(x)$ è un polinomio di grado n nella variabile x , detto il *polinomio caratteristico* di A .

Abbiamo il seguente risultato.

Teorema *Sia f un endomorfismo di V^n , e sia A una matrice associata a f . Allora:*

- a) *Gli autovalori di f sono le radici del polinomio caratteristico di A .*
- b) *Se λ è un autovalore di f allora:*

$$MG(\lambda) = \dim E(\lambda) = n - \text{rk}(A - \lambda I).$$

Dimostrazione. La prima parte è già dimostrata. La seconda segue immediatamente dal

fatto che la dimensione dell'autospazio $E(\lambda)$ associato a λ uguaglia la dimensione del sottospazio delle soluzioni del sistema lineare omogeneo $(A - \lambda I)X = O$. Dal teorema di Rouché-Capelli, tale dimensione vale $n - \text{rk}(A - \lambda I)$. \square

È noto che un polinomio di grado n ammette al massimo n radici distinte. Poiché il polinomio caratteristico di una matrice $n \times n$ ha grado n , osserviamo che

- Un endomorfismo di uno spazio vettoriale V^n ammette al massimo n autovalori distinti.

Esempio Trovare gli autovalori dell'endomorfismo di \mathbf{R}^2 : $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2x + 6y \\ -2x + 5y \end{pmatrix}$.

Soluzione. Fissando la base canonica possiamo prendere come matrice associata la matrice canonica di f , cioè $A = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}$. Si ha:

$$A - xI = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} = \begin{pmatrix} -2-x & 6 \\ -2 & 5-x \end{pmatrix}.$$

Dunque

$$\begin{aligned} p_A(x) &= \begin{vmatrix} -2-x & 6 \\ -2 & 5-x \end{vmatrix} \\ &= x^2 - 3x + 2 \end{aligned}$$

Gli autovalori si ottengono risolvendo l'equazione caratteristica $x^2 - 3x + 2 = 0$. Troviamo le due soluzioni:

$$\lambda_1 = 1, \quad \lambda_2 = 2.$$

che saranno quindi gli autovalori di f . \square

Esempio Abbiamo già osservato che l'endomorfismo $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}$ di \mathbf{R}^2 non ammette autovalori. In effetti, la sua matrice canonica $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ha polinomio caratteristico

$$p_A(x) = \begin{vmatrix} -x & 1 \\ -1 & -x \end{vmatrix} = x^2 + 1,$$

che non ammette radici (reali). \square

Esercizio Dimostrare che il polinomio caratteristico di una matrice 2×2 si scrive

$$p_A(x) = x^2 - \operatorname{tr} A \cdot x + \det A.$$

dove $\operatorname{tr} A$ è la traccia di A (definita come la somma degli elementi diagonali di A).

Ad esempio, se $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ allora $\operatorname{tr} A = 5$ e $\det A = -2$: dunque

$$p_A(x) = x^2 - 5x - 2.$$

Esempio Il polinomio caratteristico della matrice $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{pmatrix}$ è:

$$p_A(x) = \begin{vmatrix} 1-x & 1 & 1 \\ 0 & 1-x & 1 \\ 0 & 0 & -3-x \end{vmatrix} = -(x-1)^2 \cdot (x+3)$$

e gli autovalori distinti sono: $\lambda_1 = 1, \lambda_2 = -3$. \square

Notiamo che la matrice dell'esempio precedente è triangolare superiore. In generale, se A è triangolare superiore (rispettivamente, inferiore) anche $A - xI$ è triangolare superiore (rispettivamente, inferiore). Si ottiene facilmente che

- gli autovalori (distinti) di una matrice triangolare (superiore o inferiore) sono gli elementi diagonali (distinti) della matrice.

6.1 Matrici simili hanno lo stesso polinomio caratteristico

Abbiamo visto che gli autovalori di un endomorfismo f sono le radici del polinomio caratteristico di una matrice associata a f . Sappiamo che le matrici associate ad un endomorfismo sono, in genere, diverse, poiché dipendono dalla scelta di una base. In questa sezione verificheremo che tutte le matrici associate ad uno stesso endomorfismo hanno lo stesso polinomio caratteristico. Iniziamo con un esempio.

Esempio Le matrici $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} -3 & -5 \\ 3 & 5 \end{pmatrix}$ sono tutte associate all'operatore di \mathbf{R}^2

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix}$$

rispetto alle basi: \mathcal{BC} (base canonica), $\mathcal{B} = \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \mathcal{B}' = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right)$. Un calcolo mostra che, in effetti, tutte e tre le matrici hanno polinomio caratteristico $x^2 - 2x$.

Il risultato generale è conseguenza della seguente proposizione.

Proposizione *Matrici simili hanno lo stesso polinomio caratteristico.*

Dimostrazione. Dobbiamo far vedere che, se A e A' sono due matrici simili, allora $p_A(x) = p_{A'}(x)$. Sia C una matrice invertibile tale $A' = C^{-1}AC$. Allora:

$$\begin{aligned} p_{A'}(x) &= \det(A' - xI) \\ &= \det(C^{-1}AC - xI) \\ &= \det(C^{-1}AC - C^{-1}(xI)C) \\ &= \det(C^{-1}(A - xI)C) \\ &= \det(C^{-1}) \cdot \det(A - xI) \cdot \det C \\ &= \frac{1}{\det C} \cdot \det(A - xI) \cdot \det C \\ &= \det(A - xI) \\ &= p_A(x). \end{aligned}$$

□

Corollario *Tutte le matrici associate ad uno stesso endomorfismo hanno lo stesso polinomio caratteristico.*

Dimostrazione. Supponiamo che A e A' siano due matrici associate allo stesso endomorfismo f . Allora sappiamo che A e A' sono simili, e dunque, per la proposizione precedente, esse hanno lo stesso polinomio caratteristico. □

Grazie alla precedente proposizione, possiamo definire il polinomio caratteristico di un endomorfismo f di V^n come il polinomio caratteristico di una *qualunque* matrice associata a f (tale polinomio è sempre lo stesso). Scriveremo dunque

$$p_f(x) = p_A(x),$$

dove A è una matrice associata a f .

Esempio Sia $D : \mathbf{R}^3[x] \rightarrow \mathbf{R}^3[x]$ l'endomorfismo "derivazione":

$$D(p(x)) = p'(x).$$

La matrice associata a D rispetto alla base canonica $(1, x, x^2)$ di $\mathbf{R}^3[x]$ è $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$.

Dunque

$$p_f(x) = p_A(x) = -x^3,$$

e l'unico autovalore di D è $\lambda = 0$.

7 Calcolo degli autospazi

Una volta calcolati gli autovalori $\lambda_1, \dots, \lambda_k$ di un endomorfismo f (tramite il polinomio caratteristico di una sua matrice associata A) sarà possibile determinare gli autospazi semplicemente risolvendo l'equazione vettoriale

$$f(v_j) = \lambda_j v_j$$

per ognuno degli autovalori trovati. Notiamo che le coordinate X di un qualunque vettore dell'autospazio $E(\lambda_j)$ sono le soluzioni del sistema lineare omogeneo

$$(A - \lambda_j I)X = O.$$

In questo modo è possibile stabilire se l'endomorfismo è diagonalizzabile oppure no. Iniziamo discutendo due esempi; daremo in seguito dei criteri generali.

Esempio Determinare gli autospazi dell'endomorfismo di \mathbf{R}^2 definito da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2x + 6y \\ -2x + 5y \end{pmatrix}$$

e stabilire se f è diagonalizzabile.

Soluzione. La matrice di f nella base canonica è $A = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}$. Quindi:

$$p_A(x) = x^2 - 3x + 2 = (x - 1)(x - 2).$$

Gli autovalori distinti di f sono $\lambda_1 = 1, \lambda_2 = 2$. Per trovare l'autospazio associato all'autovalore 1 dobbiamo risolvere l'equazione vettoriale $f(v) = v$; passando alle coordinate, l'equazione è equivalente al sistema

$$(A - I)X = O.$$

Poiché $A - I = \begin{pmatrix} -3 & 6 \\ -2 & 4 \end{pmatrix}$ il sistema si scrive

$$\begin{cases} -3x + 6y = 0 \\ -2x + 4y = 0. \end{cases}$$

Quindi

$$E(1) = \left\{ \begin{pmatrix} 2t \\ t \end{pmatrix}, t \in \mathbf{R} \right\}, \quad \text{con base } v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Per l'autovalore 2 si ha $A - 2I = \begin{pmatrix} -4 & 6 \\ -2 & 3 \end{pmatrix}$ e il sistema è

$$\begin{cases} -4x + 6y = 0 \\ -2x + 3y = 0. \end{cases}$$

Quindi

$$E(2) = \left\{ \begin{pmatrix} 3s \\ 2s \end{pmatrix}, s \in \mathbf{R} \right\}, \quad \text{con base } v_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Notiamo che i due autospazi hanno entrambi dimensione 1 e quindi gli autovalori hanno entrambi molteplicità geometrica 1:

$$MG(1) = MG(2) = 1.$$

Collezionando le basi di ciascuno dei due autospazi (prendendo cioè $\mathcal{B}_1 \cup \mathcal{B}_2$) otteniamo la coppia di vettori linearmente indipendenti

$$(v_1, v_2) = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right),$$

che è dunque una base di \mathbf{R}^2 formata da autovettori di f . L'endomorfismo f è diagonalizzabile, e la matrice associata a f rispetto a tale base è $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. \square

È importante osservare che f ha due autovalori distinti $\lambda_1 = 1, \lambda_2 = 2$, entrambi di molteplicità geometrica 1. Quindi la somma delle molteplicità geometriche degli autovalori è 2, pari alla dimensione dello spazio su cui opera f :

$$MG(1) + MG(2) = 2 = \dim \mathbf{R}^2.$$

Esempio Determinare gli autovalori e gli autospazi dell'endomorfismo f di \mathbf{R}^3 rappresentato dalla matrice $A = \begin{pmatrix} -1 & 1 & 5 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{pmatrix}$ rispetto alla base canonica. Stabilire se f è diagonalizzabile.

Soluzione. f si scrive esplicitamente:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x + y + 5z \\ -y + 3z \\ 2z \end{pmatrix}.$$

Poiché A è triangolare superiore, gli autovalori saranno gli elementi diagonali. f ha dunque due autovalori distinti $\lambda_1 = -1, \lambda_2 = 2$ e, di conseguenza, due autospazi: $E(-1), E(2)$. Cerchiamo una base di ciascuno di essi. Per $E(-1)$ dobbiamo risolvere l'equazione $f(v) = -v$, dunque il sistema

$$(A + I)X = O.$$

Esplicitamente $A + I = \begin{pmatrix} 0 & 1 & 5 \\ 0 & 0 & 3 \\ 0 & 0 & 3 \end{pmatrix}$ dunque il sistema è:

$$\begin{cases} y + 5z = 0 \\ 3z = 0 \\ 3z = 0 \end{cases}$$

e il suo insieme delle soluzioni ha dimensione 1 con base $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. In conclusione

$$\dim E(-1) = 1, \quad \text{con base } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Per verificare il risultato ricordiamo che, in generale, la dimensione dell'autospazio $E(\lambda)$ di un endomorfismo $f : V^n \rightarrow V^n$ è data da

$$MG(\lambda) = n - \text{rk}(A - \lambda I).$$

In questo caso $\lambda = -1$ e $\text{rk}(A + I) = 2$, dunque $MG(-1) = 3 - 2 = 1$.

Veniamo ora al secondo autospazio, $E(2)$. La sua dimensione è $MG(2) = 3 - \text{rk}(A - 2I)$. Poiché la matrice

$$A - 2I = \begin{pmatrix} -3 & 1 & 5 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

ha rango due, otteniamo $MG(2) = 1$. Una base di $E(2)$ si ottiene risolvendo il sistema $(A - 2I)X = 0$:

$$\begin{cases} -3x + y + 5z = 0 \\ -3y + 3z = 0 \end{cases}$$

che ammette ∞^1 soluzioni $\begin{pmatrix} 2t \\ t \\ t \end{pmatrix}$, con $t \in \mathbf{R}$. Ne segue che:

$$\dim E(2) = 1, \quad \text{con base } \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Uniamo ora le basi dei due autospazi: otteniamo i due autovettori

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

che sono linearmente indipendenti ma che *non possono* formare una base di \mathbf{R}^3 , poiché \mathbf{R}^3 ha dimensione 3. In conclusione, f non è diagonalizzabile. In effetti, per avere una base di autovettori almeno uno dei due autospazi avrebbe dovuto avere dimensione maggiore o uguale a 2. La non diagonalizzabilità di f è dunque conseguenza del fatto che la somma delle molteplicità geometriche degli autovalori è *minore* della dimensione dell'intero spazio:

$$MG(-1) + MG(2) = 2 < 3 = \dim \mathbf{R}^3.$$

□

Nella prossima sezione mostreremo che un endomorfismo di un qualunque spazio vettoriale V^n è diagonalizzabile se e solo se la somma delle molteplicità geometriche di tutti i suoi autovalori vale precisamente n .

8 Primo criterio

Osserviamo che, dati due autovalori distinti $\lambda_1 \neq \lambda_2$ di un endomorfismo f , si ha sempre

$$E(\lambda_1) \cap E(\lambda_2) = \{O\}.$$

Infatti, se $v \in E(\lambda_1) \cap E(\lambda_2)$, allora $f(v) = \lambda_1 v$ e $f(v) = \lambda_2 v$ da cui, uguagliando, otteniamo

$$(\lambda_1 - \lambda_2)v = O.$$

Per ipotesi, $\lambda_1 - \lambda_2 \neq 0$ dunque necessariamente $v = O$.

Proposizione *Sia f un endomorfismo di uno spazio vettoriale V^n , con autovalori distinti $\lambda_1, \dots, \lambda_h$ (si suppone $h \geq 1$).*

a) Sia \mathcal{B}_i una base dell'autospazio $E(\lambda_i)$, dove $i = 1, \dots, h$. Allora i vettori dell'insieme

$$\mathcal{B}_1 \cup \dots \cup \mathcal{B}_h$$

sono linearmente indipendenti.

b) La somma di tutte le molteplicità geometriche è minore o uguale a n :

$$\sum_{i=1}^h MG(\lambda_i) \leq n.$$

Dimostrazione. a) Per dare un'idea della dimostrazione esamineremo solo il caso in cui ci

siano due autovalori distinti $\lambda_1 \neq \lambda_2$. Il caso generale si dimostra per induzione. Fissate le basi $\mathcal{B}_1 = (u_1, \dots, u_k)$ di $E(\lambda_1)$ e $\mathcal{B}_2 = (v_1, \dots, v_l)$ di $E(\lambda_2)$, supponiamo che

$$a_1 u_1 + \dots + a_k u_k + b_1 v_1 + \dots + b_l v_l = O.$$

Ponendo $u = a_1 u_1 + \dots + a_k u_k$ e $v = b_1 v_1 + \dots + b_l v_l$ otteniamo

$$u + v = O.$$

Ora per ipotesi $u \in E(\lambda_1)$ e $v \in E(\lambda_2)$; poiché $u = -v$ vediamo che

$$u \in E(\lambda_1) \cap E(\lambda_2)$$

dunque $u = O$ e

$$a_1 u_1 + \dots + a_k u_k = O.$$

Poiché i vettori u_1, \dots, u_k sono linearmente indipendenti per ipotesi otteniamo infine $a_1 = \dots = a_k = 0$. Analogamente $v = O$ implica $b_1 = \dots = b_l = 0$. La conclusione è che i vettori dell'unione

$$\mathcal{B}_1 \cup \mathcal{B}_2 = \{u_1, \dots, u_k, v_1, \dots, v_l\}$$

sono linearmente indipendenti.

b) Siccome \mathcal{B}_i è una base di $E(\lambda_i)$, il numero dei vettori di \mathcal{B}_i è $\dim E(\lambda_i) = MG(\lambda_i)$, per ogni $i = 1, \dots, h$. Dunque i vettori dell'insieme $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_h$ sono, in numero, pari alla somma di tutte le molteplicità geometriche; poiché sono linearmente indipendenti per la parte a), tale somma non può superare la dimensione n . \square

La proposizione afferma che, unendo le basi di tutti gli autospazi, otteniamo sempre vettori linearmente indipendenti: se gli autovettori così ottenuti sono in numero sufficiente (cioè n) allora essi formeranno una base dell'intero spazio e l'endomorfismo risulterà diagonalizzabile. Enunciamo dunque il criterio seguente, che chiameremo *primo criterio di diagonalizzabilità*.

Teorema Un endomorfismo di uno spazio vettoriale V^n è diagonalizzabile se e solo se la somma delle molteplicità geometriche dei suoi autovalori è uguale a n . In altre parole, se e solo se

$$\sum_{i=1}^k MG(\lambda_i) = n,$$

dove $\lambda_1, \dots, \lambda_k$ sono gli autovalori distinti di f .

Dimostrazione. Supponiamo che $\sum_{i=1}^k MG(\lambda_i) = n$. I vettori di $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ sono in tutto $n = \dim V^n$ e sono linearmente indipendenti grazie alla proposizione precedente; dunque tali vettori formano una base di autovettori e l'endomorfismo è diagonalizzabile. Viceversa, se la somma delle molteplicità geometriche è diversa da n allora deve essere minore di n per la parte b) della proposizione. È evidente che in tal caso non esiste una base di autovettori. \square

Corollario Sia f un operatore di V^n , e supponiamo che f ammetta n autovalori distinti. Allora f è diagonalizzabile.

Dimostrazione. Poiché la molteplicità geometrica di un autovalore è maggiore o uguale a uno, e gli autovalori distinti sono n , la somma di tutte le molteplicità geometriche è almeno n ; per la parte b) della proposizione tale somma deve essere uguale a n e quindi f risulta diagonalizzabile. In particolare, l'esistenza di n autovalori distinti implica che tutti gli autospazi hanno dimensione 1. \square

• Il corollario dà una condizione *sufficiente*, ma non necessaria, per la diagonalizzabilità.

Per stabilire se un dato endomorfismo di V^n sia diagonalizzabile oppure no, possiamo procedere come segue.

1. Fissiamo una base e consideriamo una matrice associata A .
2. Troviamo gli autovalori distinti di f , diciamo $\lambda_1, \dots, \lambda_k$, risolvendo l'equazione caratteristica $p_A(x) = 0$.
3. Calcoliamo la molteplicità geometrica di ciascun autovalore, con la formula

$$MG(\lambda_i) = n - \text{rk}(A - \lambda_i I).$$

4. Se la somma di tutte le molteplicità geometriche è minore di n , l'endomorfismo f non è diagonalizzabile, e abbiamo finito.
5. Se tale somma vale n , allora f è diagonalizzabile. Per trovare una base di autovettori, troviamo una base \mathcal{B}_i di ciascun autospazio risolvendo il sistema $(A - \lambda_i I)X = O$. Uniamo

tutte le basi così trovate: l'insieme di autovettori $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ sarà la base di V^n cercata.

Ecco degli esempi.

8.1 Esempio

Consideriamo l'endomorfismo f di \mathbf{R}^3

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y - z \\ 2x + 2y - 2z \\ -x - y + z \end{pmatrix}$$

Scegliendo la base canonica, possiamo considerare la matrice associata $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ -1 & -1 & 1 \end{pmatrix}$

con polinomio caratteristico $p_A(x) = -x^3 + 4x^2$ che si fattorizza come

$$p_A(x) = -x^2(x - 4).$$

Abbiamo due autovalori distinti: $\lambda_1 = 0$ e $\lambda_2 = 4$. Un calcolo mostra che $\text{rk} A = 1$ dunque

$$MG(0) = 2,$$

mentre $\text{rk}(A - 4I) = 1$ dunque

$$MG(4) = 1.$$

La somma delle molteplicità geometriche è 3, pari alla dimensione, e f risulta diagonalizzabile. Una base di $E(0)$ (che, per inciso, è il nucleo di f) è data dai vettori

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Una base di $E(4)$ è data dal vettore

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Unendo le basi degli autospazi, otteniamo i tre vettori

$$(v_1, v_2, w_1)$$

che, grazie alla proposizione, sono linearmente indipendenti (non c'è bisogno di ulteriori verifiche) e formano la base di autovettori cercata.

Notiamo che la matrice associata a f rispetto alla base $\mathcal{B} = (v_1, v_2, w_1)$ è diagonale, con elementi diagonali dati dagli autovalori associati, rispettivamente, a v_1, v_2, w_1 :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

ed è simile alla matrice A , nel senso che $D = C^{-1}AC$. La matrice C è la matrice di passaggio dalla base canonica alla base \mathcal{B} , e dunque le colonne di C sono date dalla base di autovettori \mathcal{B} :

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

□

8.2 Esempio

Consideriamo l'endomorfismo f di \mathbf{R}^3 rappresentato da $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -3 \\ 0 & 2 & 0 \end{pmatrix}$ rispetto alla base canonica. Il polinomio caratteristico di A è

$$p_A(x) = (2 - x)(x^2 + 6)$$

che ammette l'unico autovalore $\lambda_1 = 2$ (infatti, il fattore $x^2 + 6$ è irriducibile). Si vede che $A - 2I$ ha rango 2, dunque $MG(2) = 1$. La somma delle molteplicità geometriche è 1, minore di 3, e f non è diagonalizzabile. □

8.3 Esempio

Consideriamo l'endomorfismo f di \mathbf{R}^3 rappresentato da $A = \begin{pmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ rispetto alla base canonica. Poiché A è triangolare superiore, gli autovalori di f saranno gli elementi diagonali di A . Quindi f ammette tre autovalori distinti

$$\lambda_1 = 2, \lambda_2 = 3, \lambda_3 = -1.$$

Poiché la dimensione è 3, il corollario assicura che f è diagonalizzabile.

8.4 Esempio

Si consideri l'endomorfismo f di \mathbf{R}^3 rappresentato da $A = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$ rispetto alla base canonica. Per quali valori di a, b l'endomorfismo risulta diagonalizzabile?

Soluzione. Abbiamo l'unico autovalore 1 con $MG(1) = 3 - \text{rk}(A - I)$ e quindi f è diagonalizzabile se e solo se $\text{rk}(A - I) = 0$, cioè se e solo se $A - I = 0$, che corrisponde a $a = b = 0$. \square

9 Secondo criterio

In questa sezione daremo un altro criterio necessario e sufficiente per la diagonalizzabilità.

9.1 Radici di un polinomio e molteplicità

Sia $p(x)$ un polinomio di grado n a coefficienti reali. Il numero $\lambda \in \mathbf{R}$ è una *radice* di $p(x)$ se

$$p(\lambda) = 0.$$

Non tutti i polinomi ammettono radici.

Esempio Il polinomio $p(x) = x^2 + 1$ non ammette radici. Più in generale, il polinomio di secondo grado $p(x) = ax^2 + bx + c$ ammette radici se e solo se il discriminante $\Delta = b^2 - 4ac \geq 0$. In tal caso le radici si ottengono dalla formula

$$\lambda = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

Quindi avremo due radici distinte se $\Delta > 0$, una sola radice se $\Delta = 0$ e nessuna radice se $\Delta < 0$.

È ben noto che λ è una radice di $p(x)$ se e solo se $p(x)$ è divisibile per $x - \lambda$, e si avrà

$$p(x) = (x - \lambda)q(x),$$

dove $q(x)$ è un polinomio di grado $n - 1$. Se $q(\lambda) \neq 0$, diremo che λ è una radice di molteplicità 1. Se $q(\lambda) = 0$ allora possiamo dividere $q(x)$ per $x - \lambda$ e avremo

$$p(x) = (x - \lambda)^2 r(x)$$

con $r(x)$ polinomio di grado $n - 2$. Continuando in questo modo, arriveremo dopo $k \leq n$ passi alla decomposizione

$$p(x) = (x - \lambda)^k r(x),$$

con $r(x)$ polinomio tale che $r(\lambda) \neq 0$. Diremo allora che λ è una radice con *molteplicità* k . Dunque

- λ è una radice di molteplicità k se $p(x)$ è divisibile per $(x - \lambda)^k$ ma non per $(x - \lambda)^{k+1}$.

Il polinomio $p(x)$ si dice *totalmente riducibile* se si spezza come un prodotto di polinomi di primo grado del tipo $x - \lambda_j$, eventualmente moltiplicato per una costante $c \neq 0$. Se $\lambda_1, \dots, \lambda_k$ sono le radici *distinte*, allora $p(x)$ è totalmente riducibile se si scrive

$$p(x) = c(x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \cdots (x - \lambda_k)^{m_k},$$

con $m_1 + \cdots + m_k = n$. È chiaro dalla definizione che, per ogni j , l'esponente m_j è la molteplicità della radice λ_j .

Esempio Consideriamo il polinomio $p(x) = 2x^5 - 2x^4 - 4x^3$. Allora:

$$\begin{aligned} p(x) &= 2x^3(x^2 - x - 2) \\ &= 2x^3(x - 2)(x + 1). \end{aligned}$$

Dunque $p(x)$ è totalmente riducibile, con radici $0, 2, -1$: la prima ha molteplicità 3, e le altre hanno molteplicità 1. La somma delle molteplicità è 5, pari al grado di $p(x)$.

Esempio Il polinomio $p(x) = x^4 - x^2 - 6$ non è totalmente riducibile. Infatti

$$\begin{aligned} p(x) &= (x^2 - 3)(x^2 + 2) \\ &= (x - \sqrt{3})(x + \sqrt{3})(x^2 + 2). \end{aligned}$$

Il fattore $x^2 + 2$ non ammette radici, e non può essere ulteriormente decomposto. $p(x)$ ammette le radici $\sqrt{3}, -\sqrt{3}$, entrambe di molteplicità 1. Dunque la somma delle molteplicità delle radici è 2, minore di 4 (il grado di $p(x)$).

Osserviamo che

- la somma delle molteplicità delle radici di un polinomio di grado n è sempre minore o uguale a n ; tale somma vale n se e solo se il polinomio è totalmente riducibile.

9.2 Secondo criterio di diagonalizzabilità

Sia f un endomorfismo di V^n e A una matrice associata. Sappiamo che gli autovalori di f sono le radici di $p_A(x)$, il polinomio caratteristico di A .

- Definiamo *molteplicità algebrica* dell'autovalore λ la molteplicità di λ quale radice di $p_A(x)$. Essa si denota con il simbolo

$$MA(\lambda).$$

Ricordiamo che un polinomio è totalmente riducibile se e solo se la somma delle molteplicità delle sue radici è uguale al suo grado. Dunque:

- Il polinomio caratteristico di f è totalmente riducibile se e solo se la somma delle molteplicità algebriche dei suoi autovalori vale n .

Un'autovalore λ dà luogo a due molteplicità: la molteplicità algebrica $MA(\lambda)$ e la molteplicità geometrica $MG(\lambda)$ (definita, ricordiamo, come la dimensione dell'autospazio associato $E(\lambda)$).

Esempio Supponiamo che una matrice associata a f sia $A = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Allora

$$p_A(x) = (x - 2)^3(x - 3).$$

Abbiamo due autovalori: $\lambda_1 = 2$, di molteplicità algebrica 3, e $\lambda_2 = 3$, di molteplicità algebrica 1. Si verifica che

$$MG(2) = 4 - \text{rk}(A - 2I) = 4 - 3 = 1$$

e inoltre $MG(3) = 1$. Quindi $MA(2) > MG(2)$ mentre $MA(3) = MG(3)$.

Proposizione Se λ è un autovalore si ha sempre

$$MA(\lambda) \geq MG(\lambda) \geq 1.$$

Dimostrazione. La disuguaglianza $MG(\lambda) \geq 1$ è stata già osservata. La dimostrazione della disuguaglianza $MA(\lambda) \geq MG(\lambda)$ è omessa. \square

Veniamo ora al risultato più importante di questa sezione.

Teorema Un endomorfismo di V^n è diagonalizzabile se e solo se

- 1) il polinomio caratteristico di f è totalmente riducibile;
- 2) per ogni autovalore λ si ha $MA(\lambda) = MG(\lambda)$.

Dimostrazione. Supponiamo che 1) e 2) siano verificate, e siano $\lambda_1, \dots, \lambda_k$ gli autovalori distinti di f . Allora, poiché $p_f(x)$ è totalmente riducibile, abbiamo

$$\sum_{j=1}^k MA(\lambda_j) = n.$$

D'altra parte, abbiamo $MA(\lambda_j) = MG(\lambda_j)$ per ogni j , dunque

$$\sum_{j=1}^k MG(\lambda_j) = n$$

e f risulta diagonalizzabile grazie al primo criterio, dimostrato nella sezione precedente.

Viceversa, se f è diagonalizzabile si avrà $\sum_{j=1}^k MG(\lambda_j) = n$. Dunque

$$n \geq \sum_{j=1}^k MA(\lambda_j) \geq \sum_{j=1}^k MG(\lambda_j) = n,$$

dove la prima disuguaglianza scende dal fatto che il polinomio caratteristico ha grado n , e la seconda è vera poiché $MA(\lambda_j) \geq MG(\lambda_j)$ per ogni j . Ne segue che le due disuguaglianze devono essere uguaglianze, quindi

$$\sum_{j=1}^k MA(\lambda_j) = n,$$

e $p_f(x)$ è totalmente riducibile. Inoltre, necessariamente $MA(\lambda_j) = MG(\lambda_j)$ per ogni j . \square

- L'autovalore λ si dice *semplice* se $MA(\lambda) = 1$, e si dice *multiplo* se $MA(\lambda) > 1$.

Dalla proposizione, vediamo che se λ è semplice allora $MA(\lambda) = MG(\lambda) = 1$. Dunque, è sufficiente determinare l'uguaglianza delle molteplicità solo per gli autovalori multipli.

Concretamente, abbiamo il seguente algoritmo alternativo per stabilire se un dato endomorfismo f è diagonalizzabile oppure no.

- 1) Consideriamo una matrice associata A e calcoliamo $p_A(x)$.
- 2) Se $p_A(x)$ non è totalmente riducibile allora f non è diagonalizzabile e abbiamo finito.
- 3) Se $p_A(x)$ è totalmente riducibile, consideriamo gli autovalori multipli, diciamo $\lambda_1, \dots, \lambda_h$. Se le molteplicità algebriche e geometriche di ciascuno di questi autovalori coincidono allora f è diagonalizzabile, altrimenti no.

Esempio Consideriamo l'endomorfismo f di \mathbf{R}^4 rappresentato dalla matrice

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

rispetto alla base canonica (vedi esempio precedente). Si ha $p_A(x) = (x-2)^3(x-3)$, che è totalmente riducibile, con un solo autovalore multiplo $\lambda_1 = 2$ di molteplicità algebrica 3 (l'altro autovalore $\lambda_2 = 3$ è semplice). È dunque sufficiente calcolare la molteplicità geometrica dell'autovalore 2. Si ha $MG(2) = 1$, dunque $MA(2) > MG(2)$ e l'endomorfismo f non è diagonalizzabile.

Esempio Consideriamo l'endomorfismo f di \mathbf{R}^4 rappresentato dalla matrice

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

rispetto alla base canonica. Un calcolo mostra che

$$p_A(x) = (x-2)^2(x^2+x+1).$$

Il fattore quadratico x^2+x+1 non ammette radici, dunque $p_A(x)$ non è totalmente riducibile. Di conseguenza f non è diagonalizzabile.

Esempio Consideriamo l'endomorfismo f di \mathbf{R}^4 rappresentato dalla matrice

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

rispetto alla base canonica. Un calcolo mostra che

$$p_A(x) = (x^2 - 5x + 6)(x^2 - 2x - 3) = (x+1)(x-2)(x-3)^2$$

Il polinomio caratteristico è totalmente riducibile, con autovalori $-1, 2, 3$ di cui solo il terzo è multiplo:

$$MA(3) = 2.$$

È sufficiente dunque calcolare $MG(3)$. Si ha:

$$\operatorname{rk}(A - 3I) = \operatorname{rk} \begin{pmatrix} -2 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & -3 \end{pmatrix} = 2,$$

dunque

$$MG(3) = 4 - \operatorname{rk}(A - 3I) = 2.$$

In conclusione, $MA(3) = MG(3) = 2$ per l'unico autovalore multiplo, dunque f è diagonalizzabile.

- È chiaro a questo punto che le molteplicità geometriche degli autovalori sono:

$$MG(-1) = MG(2) = 1 \quad \text{e} \quad MG(3) = 2.$$

Dunque f ammette una base di autovettori

$$\mathcal{B} = (v_1, v_2, v_3, v_4)$$

con v_1 associato all'autovalore -1 , v_2 associato all'autovalore 2 , e v_3, v_4 entrambi associati all'autovalore 3 . La matrice associata a f rispetto a tale base è

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

Di conseguenza, A è simile a D : se M è la matrice di colonne v_1, v_2, v_3, v_4 allora

$$D = M^{-1}AM.$$

Per esercizio, calcolare esplicitamente una base di autovettori e la matrice di passaggio M .

9.3 Esempi su spazi di matrici

In questa sezione daremo esempi di endomorfismi di $\mathbf{Mat}(2 \times 2)$. La base canonica di $\mathbf{Mat}(2 \times 2)$ sarà scritta (E_1, E_2, E_3, E_4) con

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Esempio Consideriamo l'endomorfismo $T : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{Mat}(2 \times 2)$ definito da $T(A) = A + A^t$. Dimostrare che T è diagonalizzabile.

Soluzione. Esplicitamente $T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a & b+c \\ b+c & 2d \end{pmatrix}$. Dunque la matrice associata a T nella base canonica è:

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

Il polinomio caratteristico è $p_A(x) = x(x-2)^3$, con autovalori distinti $\lambda_1 = 0, \lambda_2 = 2$. Si ha $\text{rk} A = 3$ quindi $MG(0) = 1$, $\text{rk}(A - 2I) = 1$ quindi $MG(2) = 3$. Si ha:

$$MG(0) + MG(2) = 4 = \dim \mathbf{Mat}(2 \times 2),$$

dunque T è diagonalizzabile. Troviamo ora basi di ciascun autospazio. Una base di $E(0)$ è $E_2 - E_3$. Una base di $E(2)$ è $(E_1, E_2 + E_3, E_4)$. Dunque una base di $\mathbf{Mat}(2 \times 2)$ formata da autovettori è:

$$(E_2 - E_3, E_1, E_2 + E_3, E_4) = \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Esempio Consideriamo l'endomorfismo $T : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{Mat}(2 \times 2)$ definito da $T(A) = AN$, dove $N = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Stabilire se T è diagonalizzabile.

Soluzione. Esplicitamente $T \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x-y & x-y \\ z-w & z-w \end{pmatrix}$. La matrice associata nella base canonica è

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Un calcolo mostra che $p_A(x) = x^4$. Abbiamo dunque un solo autovalore $\lambda_1 = 0$ con $MA(0) = 4$. Poichè $\text{rk} A = 2$ si ha $MG(0) = 2 < 4$, dunque T non è diagonalizzabile.

Esempio Sia $T : \mathbf{Mat}(2 \times 2) \rightarrow \mathbf{Mat}(2 \times 2)$ definito da $T \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} -2x + 6y & -2x + 5y \\ -2z + w & w \end{pmatrix}$. Stabilire se T è diagonalizzabile.

Soluzione. Matrice di T nella base canonica:

$$A = \begin{pmatrix} -2 & 6 & 0 & 0 \\ -2 & 5 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Polinomio caratteristico $p_A(x) = (x-1)^2(x+2)(x-2)$, totalmente riducibile. Abbiamo tre autovalori distinti $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -2$, di cui solo il primo è multiplo: $MA(1) = 2$. Si ha $\text{rk}(A - I) = 2$ dunque $MG(1) = 2 = MA(1)$: per il secondo criterio, T è diagonalizzabile. Si può verificare che una base di autovettori è:

$$\left(\begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right),$$

in cui i primi due autovettori sono associati a 1, il terzo autovettore è associato a 2 e il quarto a -2 . In questa base la matrice associata è:

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

10 Matrici diagonalizzabili

Se A è una matrice quadrata $n \times n$, diremo che il vettore colonna $v \in \mathbf{R}^n$, con $v \neq O$, è un *autovettore* di A se

$$Av = \lambda v,$$

per un certo $\lambda \in \mathbf{R}$, detto *autovalore* relativo a v . In altre parole:

- v è un autovettore di A se e solo se v è un autovettore dell'endomorfismo di \mathbf{R}^n rappresentato da A rispetto alla base canonica.
- Diremo che la matrice A è *diagonalizzabile* se A è simile ad una matrice diagonale; se cioè possiamo trovare una matrice diagonale D e una matrice invertibile M tali che

$$D = M^{-1}AM. \quad (5)$$

Diagonalizzare una matrice (quando ciò è possibile) significa trovare matrici D e M che verificano la relazione (5).

Il teorema seguente ci dice quando è possibile diagonalizzare una matrice.

Teorema Sia A una matrice quadrata, e f l'operatore di \mathbf{R}^n rappresentato da A rispetto alla base canonica. Allora:

- A è diagonalizzabile se e solo se f è diagonalizzabile.
- Se $\mathcal{B} = (v_1, \dots, v_n)$ è una base di autovettori di f , associati rispettivamente agli autovalori $\lambda_1, \dots, \lambda_n$ (non necessariamente distinti) e se M è la matrice di colonne v_1, \dots, v_n , allora si ha:

$$M^{-1}AM = D$$

dove D è la matrice diagonale di elementi diagonali $\lambda_1, \dots, \lambda_n$, rispettivamente.

In altre parole, A è simile alla matrice i cui elementi diagonali sono gli autovalori di f , e M è la matrice le cui colonne formano una base di autovettori di f (presi nello stesso ordine degli autovalori).

Dimostrazione. a) Se f è diagonalizzabile, allora esiste una matrice associata diagonale D : sappiamo che matrici associate allo stesso endomorfismo sono simili, dunque A è simile a D . Il viceversa si dimostra in modo analogo.

b) Sappiamo che nella base di autovettori \mathcal{B} l'endomorfismo f si rappresenta con la matrice diagonale D avente elementi diagonali $\lambda_1, \dots, \lambda_n$, e che

$$D = M^{-1}AM$$

dove M è la matrice di passaggio dalla base canonica alla base di autovettori \mathcal{B} . Per definizione di matrice di passaggio, le colonne di M sono proprio i vettori di \mathcal{B} , e dunque si ha (b). \square

10.1 Esempio

Diagonalizzare, se possibile, la matrice $A = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}$.

Soluzione. Gli autovalori distinti di A sono $\lambda_1 = 1, \lambda_2 = 2$, sono due, dunque A è diagonalizzabile. Base di autovettori:

$$(v_1, v_2) = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right).$$

Dunque $D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ e $M = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$ e risulta $D = M^{-1}AM$.

10.2 Serie di esempi

Consideriamo le seguenti matrici:

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 2 & 2 & 3 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Vediamo quali fra queste sono diagonalizzabili, applicando il secondo criterio.

Matrice A_1 . Polinomio caratteristico $-(x-1)^2(x+3)$, totalmente riducibile. Autovalori distinti $1, -3$, di cui solo il primo è multiplo, con $MA(1) = 2$. Si ha $MG(3) = 1 < MA(1)$. Quindi A_1 non è diagonalizzabile.

Matrice A_2 . Polinomio caratteristico $-(x-1)^2(x-3)$, totalmente riducibile. Autovalori distinti $1, 3$, di cui solo il primo è multiplo, con $MA(1) = 2$. Si ha $MG(1) = 2$ quindi A_2 è diagonalizzabile.

Matrice A_3 . Polinomio caratteristico $-(x-1)(x-2)(x+1)$. Autovalori distinti $1, 2, -1$: sono tre, tutti semplici, quindi A_3 è diagonalizzabile.

Matrice A_4 . Polinomio caratteristico $-(x-2)(x^2+1)$: il fattore quadratico x^2+1 non ha radici, dunque il polinomio caratteristico non è totalmente riducibile e A_4 non è diagonalizzabile.

Possiamo diagonalizzare solo A_2 e A_3 .

Diagonalizzazione di A_2 . Autovalori $1, 3$. Cerchiamo una base degli autospazi. Si ha

$$A_2 - I = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 2 & 2 & 2 \end{pmatrix}$$

e l'autospazio $E(1)$ si ottiene risolvendo il sistema $(A_2 - I)X = O$. Un calcolo mostra che una base di $E(1)$ è $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Si ha poi

$$A_2 - 3I = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -3 & -1 \\ 2 & 2 & 0 \end{pmatrix}$$

per cui una base di $E(3)$ è $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$. Unendo le basi dei due autospazi otteniamo la base di autovettori:

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right) \quad \text{associati, nell'ordine, a } 1, 1, 3.$$

Dunque

$$M = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Possiamo verificare che $D = M^{-1}AM$ ovvero che $MD = AM$.

Diagonalizzazione di A_3 . Autovalori $1, 2, -1$. Si ha:

$$\begin{aligned} A_3 - I &= \begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \end{pmatrix}, \quad \text{base di } E(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \\ A_3 - 2I &= \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -3 \end{pmatrix}, \quad \text{base di } E(2) = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}. \\ A_3 + I &= \begin{pmatrix} 2 & 2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{base di } E(-1) = \begin{pmatrix} -5/2 \\ 1 \\ 3 \end{pmatrix}. \end{aligned}$$

Quindi una base di autovettori è:

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -5/2 \\ 1 \\ 3 \end{pmatrix} \right) \quad \text{associati, nell'ordine, a } 1, 2, -1.$$

Si ha dunque:

$$M = \begin{pmatrix} 1 & 2 & -\frac{5}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Parte 8. Prodotto scalare, teorema spettrale

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Prodotto scalare in \mathbf{R}^n , 1
- 2 Basi ortonormali, 4
- 3 Algoritmo di Gram-Schmidt, 7
- 4 Matrici ortogonali, 12
- 5 Complemento ortogonale di un sottospazio, 13
- 6 Endomorfismi simmetrici, 17
- 7 Teorema spettrale, 20

1 Prodotto scalare in \mathbf{R}^n

1.1 Definizione

Dati i vettori $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ di \mathbf{R}^n , definiamo *prodotto scalare di u e v* il numero reale:

$$\langle u, v \rangle = x_1 y_1 + \cdots + x_n y_n.$$

Il risultato del prodotto scalare è dunque un numero.

Esempio Se $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $v = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ allora $\langle u, v \rangle = 3$.

Notiamo che il prodotto scalare di due vettori può risultare nullo anche se nessuno dei due fattori è il vettore nullo.

Esempio Se $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $v = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ allora $\langle u, v \rangle = 0$.

- I vettori u e v si dicono *ortogonali* se il loro prodotto scalare è nullo: $\langle u, v \rangle = 0$.
Notazione:

$$u \perp v.$$

Dunque i vettori dell'esempio precedente sono ortogonali. È evidente che, se O è il vettore nullo, si ha $\langle v, O \rangle = 0$ per ogni $v \in \mathbf{R}^n$: dunque il vettore nullo è ortogonale a tutti i vettori.

La denominazione di *vettori ortogonali* legata alla condizione $\langle u, v \rangle = 0$ (che è puramente algebrica) sarà giustificata quando studieremo la geometria analitica, e introdurremo i vettori geometrici del piano e dello spazio. Infatti, l'introduzione del prodotto scalare permette di definire, in modo puramente algebrico, la *norma* di un vettore (che va interpretata come la distanza del vettore stesso dal vettore nullo) e l'*angolo* fra due vettori non nulli. Per il momento, ci proponiamo di studiare le proprietà algebriche dell'operazione di prodotto scalare.

Proposizione Siano u, v, w vettori arbitrari di \mathbf{R}^n e sia $k \in \mathbf{R}$ un qualunque scalare. Allora si hanno le seguenti proprietà.

- 1) $\langle u, v \rangle = \langle v, u \rangle$.
- 2) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.
- 3) $\langle ku, v \rangle = k\langle u, v \rangle$.
- 4) $\langle u, u \rangle \geq 0$.
- 5) $\langle u, u \rangle = 0$ se e solo se $u = O$.

La 1) dice che il prodotto scalare è commutativo. Le proprietà 2), 3) esprimono la cosiddetta proprietà di *bilinearità*. Le proprietà 4) e 5) esprimono il fatto che il prodotto scalare è *definito positivo*.

Dimostrazione. La dimostrazione di 1), 2), 3) si riduce a una semplice verifica. Osserviamo

che, se $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ allora

$$\langle u, u \rangle = x_1^2 + \cdots + x_n^2,$$

che è un numero sempre positivo o nullo: questo dimostra la 4). Se $\langle u, u \rangle = 0$ evidentemente $x_1 = \cdots = x_n = 0$, e quindi $u = O$. \square

Dalle proprietà di bilinearità osserviamo che il prodotto scalare si comporta in modo naturale rispetto alle combinazioni lineari. Per ogni scelta dei vettori $v_1, \dots, v_k, u, w \in \mathbf{R}^n$ e degli scalari $a_1, \dots, a_k \in \mathbf{R}$ si ha:

$$\langle a_1 v_1 + \cdots + a_k v_k, w \rangle = a_1 \langle v_1, w \rangle + \cdots + a_k \langle v_k, w \rangle.$$

Di conseguenza, poiché il prodotto scalare è commutativo, si ha anche

$$\langle u, a_1 v_1 + \cdots + a_k v_k \rangle = a_1 \langle u, v_1 \rangle + \cdots + a_k \langle u, v_k \rangle$$

1.2 Norma e disuguaglianza di Schwarz

Per definizione, la *norma* di un vettore $u \in \mathbf{R}^n$ è il numero positivo o nullo

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Esplicitamente $\|u\| = \sqrt{x_1^2 + \cdots + x_n^2}$ ovvero

$$\|u\|^2 = x_1^2 + \cdots + x_n^2.$$

In particolare, $\|u\| \geq 0$ e si ha l'uguaglianza solo quando $u = O$: la norma di un vettore non nullo è sempre positiva.

Esempio Se $u = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ allora $\|u\| = \sqrt{1 + 9 + 4} = \sqrt{14}$.

Teorema (Disuguaglianza di Schwarz). *Dati $u, v \in \mathbf{R}^n$ si ha sempre:*

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

Inoltre, vale l'uguaglianza se e solo se u e v sono linearmente dipendenti.

Dimostrazione. Omessa. \square

Esempio Dati n numeri reali a_1, \dots, a_n si ha sempre:

$$(a_1 + \cdots + a_n)^2 \leq n(a_1^2 + \cdots + a_n^2).$$

Infatti, basta applicare la disuguaglianza di Schwarz ai vettori $u = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ e $v = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

Notiamo che si ha l'uguaglianza solo quando a_1, \dots, a_n sono tutti uguali tra loro.

1.3 Angolo tra due vettori

Supponiamo che u e v siano due vettori non nulli. Per la disuguaglianza di Schwarz, si ha

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

dunque esiste un unico valore di $\theta \in [0, \pi]$ tale che

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

Per definizione, θ è detto l'*angolo* tra u e v .

Esempio Dati $u = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$ si ha:

$$\|u\| = \sqrt{6}, \quad \|v\| = \sqrt{6}, \quad \langle u, v \rangle = 3.$$

Dunque $\cos \theta = \frac{1}{2}$ cioè $\theta = \frac{\pi}{3}$.

2 Basi ortonormali

2.1 Ortogonalità e indipendenza lineare

Proposizione Siano v_1, \dots, v_k vettori non nulli di \mathbf{R}^n , a due a due ortogonali. Allora v_1, \dots, v_k sono linearmente indipendenti. In particolare, $k \leq n$.

Dimostrazione. Supponiamo che

$$a_1 v_1 + \dots + a_k v_k = O. \tag{1}$$

Prendendo il prodotto scalare dei due membri della (1) per v_1 otteniamo

$$\begin{aligned} 0 &= \langle a_1 v_1 + a_2 v_2 + \dots + a_k v_k, v_1 \rangle \\ &= a_1 \langle v_1, v_1 \rangle + a_2 \langle v_2, v_1 \rangle + \dots + a_k \langle v_k, v_1 \rangle \\ &= a_1 \|v_1\|^2 \end{aligned}$$

perché per ipotesi $\langle v_j, v_1 \rangle = 0$ per ogni $j = 2, \dots, k$. Per ipotesi, v_1 è non nullo, dunque $\|v_1\|^2 > 0$ e ne segue che $a_1 = 0$. Prendendo successivamente il prodotto scalare dei due

membri della (1), ordinatamente per v_2, \dots, v_k , si dimostra in modo analogo che $a_j = 0$ per ogni j . \square

- n vettori non nulli, a due a due ortogonali formano una base di \mathbf{R}^n (che sarà chiamata *base ortogonale*).

Esempio I vettori $v_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ formano una base ortogonale di \mathbf{R}^2 perchè $\langle v_1, v_2 \rangle = 0$.

I vettori $w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ sono non nulli e a due a due ortogonali:

$$\langle w_1, w_2 \rangle = \langle w_1, w_3 \rangle = \langle w_2, w_3 \rangle = 0.$$

Dunque (w_1, w_2, w_3) è una base ortogonale di \mathbf{R}^3 .

Esempio La matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

ha rango 4. Infatti i suoi vettori colonna sono a due a due ortogonali, e quindi sono linearmente indipendenti.

- Il numero *massimo* di vettori di \mathbf{R}^n , non nulli e ortogonali a due a due, è n .

In modo analogo, possiamo definire la nozione di *base ortogonale* di un qualunque sottospazio E di \mathbf{R}^n : se $\dim E = k$ allora i vettori v_1, \dots, v_k formano una base ortogonale di E se sono non nulli e $\langle v_i, v_j \rangle = 0$ per ogni $i \neq j$.

Esempio Il sottospazio $E : x + y + z = 0$ di \mathbf{R}^3 ha dimensione 2. I due vettori $v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ appartengono a E e sono ortogonali tra loro, dunque formano una base ortogonale di E .

2.2 Basi ortonormali

Diremo che una base (v_1, \dots, v_k) di un sottospazio E di \mathbf{R}^n è una *base ortonormale* di E se:

$$\langle v_i, v_j \rangle = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j. \end{cases}$$

Dunque una base ortonormale è formata da vettori a due a due ortogonali, tutti di norma unitaria. Una base ortonormale è, in particolare, anche ortogonale.

Esempio La base canonica di \mathbf{R}^n è una base ortonormale.

Fare i conti con le basi ortonormali è più semplice. Ad esempio, trovare le coordinate di un vettore rispetto a una base implica, normalmente, la risoluzione di un certo numero di sistemi lineari. Se la base è ortonormale, è sufficiente calcolare un certo numero di prodotti scalari.

Proposizione Sia $\mathcal{B} = (v_1, \dots, v_k)$ una base ortonormale di un sottospazio E di \mathbf{R}^n . Allora le coordinate del vettore $v \in E$ rispetto a \mathcal{B} sono date da

$$\begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_k \rangle \end{pmatrix},$$

e sono dette coefficienti di Fourier di v .

Dimostrazione. Se $v \in E$ possiamo scrivere

$$v = a_1 v_1 + \dots + a_k v_k$$

e per definizione le coordinate di v sono a_1, \dots, a_k . Ora, prendendo il prodotto scalare dei due membri successivamente per v_1, \dots, v_k , otteniamo facilmente

$$a_j = \langle v, v_j \rangle$$

per ogni $j = 1, \dots, k$. \square

Esempio I vettori:

$$v_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, v_3 = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, v_4 = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix},$$

sono a due a due ortogonali e hanno tutti norma 1. Dunque tali vettori formano una

base ortonormale \mathcal{B} di \mathbf{R}^4 . Calcoliamo le coordinate del vettore $v = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ rispetto a \mathcal{B} . I

coefficienti di Fourier sono

$$\langle v, v_1 \rangle = 5, \langle v, v_2 \rangle = -2, \langle v, v_3 \rangle = -1, \langle v, v_4 \rangle = 0.$$

Dunque v ha coordinate

$$\begin{pmatrix} 5 \\ -2 \\ -1 \\ 0 \end{pmatrix},$$

rispetto a \mathcal{B} . In altre parole $v = 5v_1 - 2v_2 - v_3$. \square

3 Algoritmo di Gram-Schmidt

Lo scopo di questa sezione è quello di dimostrare che ogni sottospazio di \mathbf{R}^n ammette almeno una base ortonormale.

3.1 Vettore normalizzato

Proposizione 1) Dato un vettore v e uno scalare $a \in \mathbf{R}$ si ha: $\|av\| = |a|\|v\|$.

2) Se $v \neq 0$ il vettore

$$u = \frac{1}{\|v\|}v$$

ha norma 1.

Dimostrazione. Si ha, dalle proprietà del prodotto scalare:

$$\|av\|^2 = \langle av, av \rangle = a^2 \langle v, v \rangle = a^2 \|v\|^2,$$

e la 1) segue prendendo la radice quadrata ad ambo i membri. La 2) segue immediatamente dalla 1) prendendo $a = \frac{1}{\|v\|}$. \square

Il vettore $u = \frac{1}{\|v\|}v$ si dice *normalizzato* di v . Normalizzare un vettore significa semplicemente dividere il vettore per la propria norma.

Esempio Il vettore $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ ha norma $\sqrt{14}$. Il suo normalizzato è dunque

$$u = \frac{1}{\|v\|}v = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

e ha norma 1.

Corollario Se (v_1, \dots, v_k) è una base ortogonale del sottospazio E , allora i vettori normalizzati

$$u_1 = \frac{1}{\|v_1\|}v_1, \quad \dots \quad, u_k = \frac{1}{\|v_k\|}v_k$$

formano una base ortonormale di E .

Dimostrazione. I vettori u_1, \dots, u_k hanno tutti norma 1, ed evidentemente appartengono a E . Essi sono a due a due ortogonali, poiché

$$\langle u_i, u_j \rangle = \frac{1}{\|v_i\|} \frac{1}{\|v_j\|} \langle v_i, v_j \rangle = 0$$

per ogni $i \neq j$. \square

Esempio Il sottospazio di \mathbf{R}^3 definito dall'equazione $E : x + y + z = 0$ ha dimensione 2 e ha una base ortogonale formata dai vettori

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Per ottenere una base ortonormale di E è sufficiente normalizzare i vettori v_1, v_2 . Si ottiene la base ortonormale

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

3.2 Procedimento di ortonormalizzazione

L'algoritmo di Gram-Schmidt è un procedimento che, applicato ad una base di un sottospazio di \mathbf{R}^n , permette di ottenere una base ortogonale del sottospazio stesso; normalizzando i vettori di tale base, otterremo una base ortonormale. Descriviamo l'algoritmo in dettaglio.

Sia E un sottospazio di \mathbf{R}^n e sia (v_1, \dots, v_k) una sua base. Dunque $\dim E = k$.

Notiamo che se $k = 1$ la base è formata dal solo vettore v_1 . È sufficiente dunque normalizzare v_1 per ottenere la base ortonormale cercata.

1) Supponiamo che la dimensione di E sia 2, e sia (v_1, v_2) una base di E . Introduciamo nuovi vettori (w_1, w_2) nel modo seguente:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - aw_1 \end{cases}$$

con $a \in \mathbf{R}$ da determinare in modo opportuno. Notiamo che i vettori w_1, w_2 appartengono a E ; inoltre w_2 non è nullo (altrimenti v_1 e v_2 sarebbero linearmente dipendenti). Ora scegliamo il coefficiente a in modo tale che w_2 risulti ortogonale a w_1 . È facile vedere che ciò accade se solo se:

$$a = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle}.$$

Dunque, con tale scelta, otteniamo la base ortogonale (w_1, w_2) di E .

2) Supponiamo ora che $\dim E = 3$, con base (v_1, v_2, v_3) e poniamo:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - aw_1 \\ w_3 = v_3 - bw_1 - cw_2 \end{cases}$$

dove $a = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle}$ è stato già determinato, cosicché $\langle w_1, w_2 \rangle = 0$. Imponendo le condizioni

$$\langle w_3, w_1 \rangle = \langle w_3, w_2 \rangle = 0,$$

otteniamo:

$$b = \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle}, \quad c = \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle}.$$

Con tali scelte di a, b, c otteniamo quindi la base ortogonale (w_1, w_2, w_3) di E e quindi, normalizzando, una base ortonormale (notiamo che w_3 non è nullo, altrimenti v_1, v_2, v_3 sarebbero linearmente dipendenti).

Procedendo per induzione, possiamo enunciare il seguente teorema.

Teorema (Algoritmo di Gram-Schmidt) *Sia (v_1, \dots, v_k) una base del sottospazio E di \mathbf{R}^n . Si introducano i vettori:*

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21}w_1 \\ w_3 = v_3 - a_{31}w_1 - a_{32}w_2 \\ \dots \\ w_k = v_k - a_{k1}w_1 - a_{k2}w_2 - \dots - a_{k,k-1}w_{k-1} \end{cases}$$

dove si è posto:

$$a_{ij} = \frac{\langle v_i, w_j \rangle}{\langle w_j, w_j \rangle}.$$

Allora (w_1, \dots, w_k) è una base ortogonale di E , e quindi i vettori normalizzati:

$$u_1 = \frac{1}{\|w_1\|} w_1, \dots, u_k = \frac{1}{\|w_k\|} w_k,$$

formano una base ortonormale di E .

Esempio Trovare una base ortonormale del sottospazio E di \mathbf{R}^3 di equazione:

$$E : x - y - 2z = 0.$$

Soluzione. Determiniamo una base di E , e poi applichiamo l'algoritmo di Gram-Schmidt per ottenere una base ortonormale. Base di E :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

L'algoritmo consiste di due passi:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21}w_1 \end{cases}.$$

Si ha $w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, dunque:

$$a_{21} = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} = \frac{2}{2} = 1.$$

Allora:

$$\begin{cases} w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ w_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \end{cases}.$$

Si verifica che in effetti $\langle w_1, w_2 \rangle = 0$. Una base ortonormale di E è dunque:

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

- Ovviamente la base ortonormale ottenuta dipende dalla base di partenza. Per esercizio, vedere quale base ortonormale si ottiene scambiando i vettori della base di partenza.

Esempio Trovare una base ortonormale del sottospazio di \mathbf{R}^4 generato dai vettori:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}.$$

Soluzione. I tre vettori formano una base di E . Applichiamo l'algoritmo di Gram-Schmidt alla terna v_1, v_2, v_3 :

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21}w_1 \\ w_3 = v_3 - a_{31}w_1 - a_{32}w_2 \end{cases}$$

Abbiamo $w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, e quindi $a_{21} = 1$. Dunque

$$w_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Ora:

$$\begin{cases} a_{31} = \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} = \frac{4}{2} = 2 \\ a_{32} = \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} = \frac{-3}{3} = -1 \end{cases}$$

dunque:

$$w_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}.$$

Otteniamo la base ortogonale:

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

e, normalizzando, la base ortonormale:

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, u_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

4 Matrici ortogonali

Abbiamo visto che la matrice M di passaggio fra due basi $\mathcal{B}, \mathcal{B}'$ di uno spazio vettoriale è invertibile. Se le basi $\mathcal{B}, \mathcal{B}'$ sono ortonormali, la matrice di passaggio avrà delle proprietà particolari.

Definizione Una matrice quadrata M si dice ortogonale se verifica $MM^t = I$. Quindi una matrice ortogonale M è invertibile e

$$M^{-1} = M^t,$$

cioè l'inversa coincide con la trasposta.

Esempio La matrice $M = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$ è ortogonale.

Esempio La matrice $M = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \end{pmatrix}$ è ortogonale.

In entrambi i casi si verifica infatti che $MM^t = I$.

Osserviamo che, se $MM^t = I$ allora, prendendo il determinante e applicando il teorema di Binet, si $(\det M)^2 = 1$. Dunque

- se M è una matrice ortogonale allora $\det M = 1$ oppure $\det M = -1$.

Il teorema seguente fornisce le proprietà importanti di una matrice ortogonale.

Teorema a) La matrice di passaggio fra due basi ortonormali di \mathbf{R}^n (o di un suo sottospazio) è ortogonale.

b) Una matrice $A \in \mathbf{Mat}(n \times n)$ è ortogonale se e solo se le colonne di A formano una base ortonormale di \mathbf{R}^n .

Dimostrazione. La dimostrazione si riduce a una verifica, che omettiamo. \square

Osserviamo che le colonne delle matrici ortogonali dei due esempi precedenti formano, effettivamente, una base ortonormale di \mathbf{R}^2 (primo esempio), e di \mathbf{R}^3 (secondo esempio). Dalla parte b) del teorema abbiamo anche

- Incolonnando i vettori di una base ortonormale di \mathbf{R}^n otteniamo una matrice ortogonale $n \times n$.

Infine, si può dimostrare che le matrici ortogonali di \mathbf{R}^2 sono di due tipi:

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

con $\theta \in \mathbf{R}$, oppure

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix},$$

con $\alpha \in \mathbf{R}$. Le matrici del primo tipo hanno determinante 1, mentre quelle del secondo tipo hanno determinante -1 .

5 Complemento ortogonale di un sottospazio

Sia u_1 un vettore fissato di \mathbf{R}^n e si consideri il sottoinsieme

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = 0\},$$

formato da tutti i vettori ortogonali a u_1 . Per le proprietà di bilinearità del prodotto scalare, E risulta allora un sottospazio di \mathbf{R}^n . Più in generale, fissati k vettori di \mathbf{R}^n , diciamo u_1, \dots, u_k , l'insieme:

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \dots = \langle v, u_k \rangle = 0\},$$

formato dai vettori di \mathbf{R}^n ortogonali a u_1, \dots, u_k è un sottospazio di \mathbf{R}^n .

Esempio Sia $u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Trovare una base di $E = \{v \in \mathbf{R}^3 : \langle v, u_1 \rangle = 0\}$.

Soluzione. Sia $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ il vettore generico di \mathbf{R}^3 . Imponendo l'ortogonalità al vettore u_1 otteniamo l'unica condizione

$$x + y - z = 0.$$

Dunque E è il sottospazio delle soluzioni dell'equazione, e una sua base è, ad esempio, $\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$. E ha dimensione 2. \square

Esempio Dati i vettori $u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$ si consideri il sottospazio

$$F = \{v \in \mathbf{R}^3 : \langle v, u_1 \rangle = \langle v, u_2 \rangle = 0\}.$$

- a) Trovare una base di F e calcolare la sua dimensione.
- b) Trovare un vettore di F avente norma 1.

Soluzione. a) Imponendo al vettore generico $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ l'ortogonalità a u_1 e u_2 vediamo che F è descritto dalle equazioni

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

La matrice dei coefficienti del sistema che definisce F è $A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \end{pmatrix}$. Notiamo che le righe di A sono proprio i vettori u_1, u_2 (più precisamente, u_1^t, u_2^t): siccome u_1, u_2 sono linearmente indipendenti il rango vale 2 e l'insieme delle soluzioni F ha dimensione:

$$\dim F = 3 - \text{rk} A = 1.$$

Una base si ottiene risolvendo il sistema. Si ottiene ad esempio la base $\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$.

b) Un vettore di E di norma 1 si ottiene normalizzando il vettore della base trovata, dunque $w = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$. Un altro vettore possibile è $-w = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$. Verificare che non ce ne sono altri. \square

Generalizzando, otteniamo il seguente risultato.

Proposizione Se i vettori $u_1, \dots, u_k \in \mathbf{R}^n$ sono linearmente indipendenti, allora

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \dots = \langle v, u_k \rangle = 0\},$$

è un sottospazio di \mathbf{R}^n di dimensione $n - k$.

Dimostrazione. Abbiamo già osservato che E è un sottospazio. Sia $v = (x_1, \dots, x_n)^t$ il vettore generico di \mathbf{R}^n . Imponendo l'ortogonalità di v a ciascuno dei vettori u_1, \dots, u_k otteniamo un sistema lineare omogeneo di k equazioni nelle n incognite x_1, \dots, x_n . Dunque E ha dimensione $n - \text{rk}A$, dove A è la matrice dei coefficienti. Ora, si verifica che le righe di A sono i vettori trasposti di u_1, \dots, u_k . Poiché per ipotesi u_1, \dots, u_k sono linearmente indipendenti il rango di A vale k e dunque

$$\dim E = n - k.$$

□

5.1 Complemento ortogonale di un sottospazio

Sia E un sottospazio di \mathbf{R}^n . Definiamo *complemento ortogonale di E* l'insieme E^\perp costituito dai vettori di \mathbf{R}^n ortogonali a *tutti* i vettori di E :

$$E^\perp = \{v \in \mathbf{R}^n : \langle v, w \rangle = 0 \text{ per ogni } w \in E\}.$$

Dalle proprietà del prodotto scalare risulta che E^\perp è chiuso rispetto alla somma e al prodotto per uno scalare, dunque è un sottospazio di \mathbf{R}^n .

Risulta che $v \in E^\perp$ se e solo se v è ortogonale a tutti i vettori di una base di E . Infatti:

Proposizione *Sia (v_1, \dots, v_k) una base di E . Allora $v \in E^\perp$ se e solo se $\langle v, v_i \rangle = 0$ per ogni $i = 1, \dots, k$.*

Dimostrazione. Supponiamo che $\langle v, v_i \rangle = 0$ per ogni $i = 1, \dots, k$. Se w è un qualunque vettore di E , allora w è combinazione lineare dei vettori della base: $w = a_1 v_1 + \dots + a_k v_k$. Quindi

$$\langle v, w \rangle = a_1 \langle v, v_1 \rangle + \dots + a_k \langle v, v_k \rangle = 0,$$

che dimostra che v è ortogonale a w . Siccome $w \in E$ è arbitrario, $v \in E^\perp$. Il viceversa è immediato. □

Esempio Determinare una base di E^\perp , complemento ortogonale del sottospazio E di \mathbf{R}^4

generato dai vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

Soluzione. Imponiamo al vettore generico $v = (x, y, z, w)^t \in \mathbf{R}^4$ l'ortogonalità ai vettori della base (v_1, v_2) di E , ottenendo il sistema omogeneo:

$$\begin{cases} x + y + z + w = 0 \\ x + z = 0 \end{cases}.$$

Risolvendo il sistema, otteniamo la base di E^\perp : $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$. \square

Le proprietà importanti del complemento ortogonale sono espresse nel seguente teorema.

Teorema *Sia E un sottospazio di \mathbf{R}^n e E^\perp il suo complemento ortogonale. Allora*

- a) $E \cap E^\perp = \{O\}$.
- b) $\dim E^\perp = n - \dim E$.
- c) $\mathbf{R}^n = E \oplus E^\perp$.

Dimostrazione. a) Se $v \in E \cap E^\perp$ allora v è ortogonale a tutti i vettori di E ; in particolare v è ortogonale a sé stesso, e dunque $\langle v, v \rangle = 0$. Ma l'unico vettore con tale proprietà è il vettore nullo.

b) Sia $\dim E = k$ e sia (u_1, \dots, u_k) una base di E . Sappiamo che $v \in E^\perp$ se e solo se v è ortogonale ai vettori di una base di E : dunque

$$E^\perp = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \dots = \langle v, u_k \rangle = 0\}.$$

Siccome u_1, \dots, u_k sono linearmente indipendenti, dalla proposizione del paragrafo precedente otteniamo che $\dim E^\perp = n - k = n - \dim E$.

c) Applichiamo la formula di Grassmann ai sottospazi E, E^\perp :

$$\dim(E + E^\perp) + \dim(E \cap E^\perp) = \dim E + \dim E^\perp.$$

Da a) e b) concludiamo che $\dim(E + E^\perp) = n$. Dunque $E + E^\perp = \mathbf{R}^n$. Poiché $E \cap E^\perp = \{O\}$ la somma è diretta: $\mathbf{R}^n = E \oplus E^\perp$. \square

5.2 Proiezione ortogonale su un sottospazio

Dal teorema precedente abbiamo che un vettore $v \in \mathbf{R}^n$ si spezza, in modo unico, come somma di un vettore $w \in E$ e di un vettore $w^\perp \in E^\perp$:

$$v = w + w^\perp.$$

In particolare, w e w^\perp sono ortogonali.

- Il vettore w è detto la *proiezione ortogonale* di v sul sottospazio E . Denoteremo w con il simbolo $P_E(v)$.

Esempio È dato il sottospazio di \mathbf{R}^3 descritto dall'equazione $x + y - 2z = 0$. Il vettore $v = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \in \mathbf{R}^3$ si spezza

$$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 3/2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix},$$

dove il primo vettore appartiene a E e il secondo a E^\perp . Quindi $P_E(v) = \begin{pmatrix} 1/2 \\ 3/2 \\ 1 \end{pmatrix}$.

In generale, se (u_1, \dots, u_k) è una base *ortonormale* di E allora la proiezione ortogonale si calcola con la formula

$$P_E(v) = \langle v, u_1 \rangle u_1 + \dots + \langle v, u_k \rangle u_k.$$

6 Endomorfismi simmetrici

In questa sezione studieremo una classe importante di endomorfismi di \mathbf{R}^n : gli endomorfismi detti *simmetrici*. Tali endomorfismi sono caratterizzati dalla proprietà di ammettere una base *ortonormale* di autovettori, e sono legati in modo naturale alle matrici simmetriche. In particolare, risulterà che ogni matrice simmetrica è diagonalizzabile.

Definizione Un endomorfismo di \mathbf{R}^n si dice *simmetrico* se la sua matrice associata rispetto alla base canonica è simmetrica.

Esempio Sia $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definito da $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x + 3y \\ 3x + 5y \end{pmatrix}$. La matrice canonica di f è:

$$A = \begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}.$$

Siccome A è simmetrica, f è simmetrico.

Esempio L'endomorfismo $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3y \end{pmatrix}$ ha matrice canonica $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ dunque non è simmetrico.

Teorema *Le seguenti condizioni sono equivalenti.*

- a) f è un endomorfismo simmetrico di \mathbf{R}^n .
- b) La matrice associata a f rispetto ad una qualunque base ortonormale di \mathbf{R}^n è simmetrica.
- c) Per ogni coppia di vettori $u, v \in \mathbf{R}^n$ si ha

$$\langle f(u), v \rangle = \langle u, f(v) \rangle.$$

Dimostrazione. a) \implies b) Supponiamo che f sia simmetrico, e sia A la sua matrice canonica. Per definizione, A è simmetrica. Se \mathcal{B} è una base ortonormale di \mathbf{R}^n , e A' è la matrice associata a f rispetto a tale base, allora sappiamo che

$$A' = M^{-1}AM,$$

dove M è la matrice di passaggio dalla base canonica \mathcal{BC} alla base \mathcal{B} . Poichè tali basi sono entrambe ortonormali, si ha che M è ortogonale, quindi $M^{-1} = M^t$. Dunque $A' = M^tAM$, ed è sufficiente dimostrare che M^tAM è simmetrica. Ma questo è immediato:

$$(M^tAM)^t = M^tA(M^t)^t = M^tAM.$$

b) \implies c) Si ha la seguente identità, valida per ogni matrice A e per ogni scelta di u, v , vettori colonna di \mathbf{R}^n :

$$\langle Au, v \rangle = \langle u, A^t v \rangle.$$

L'identità si verifica con un calcolo diretto, e fornisce un legame tra il prodotto scalare e la trasposta di una matrice. Supponiamo che la matrice A , associata ad f rispetto alla base canonica, sia simmetrica. Ora sappiamo che f si scrive

$$f(v) = Av.$$

Poichè A è simmetrica, si ha $A = A^t$ e dall'identità precedente:

$$\langle f(u), v \rangle = \langle Au, v \rangle = \langle u, Av \rangle = \langle u, f(v) \rangle.$$

c) \implies a) Premettiamo che, se A è una matrice $n \times n$ e e_1, \dots, e_n sono i vettori della base canonica di \mathbf{R}^n , un calcolo mostra che

$$\langle Ae_i, e_j \rangle = a_{ji},$$

dove a_{ji} è l'elemento di posto (j, i) della matrice A .

Per ipotesi, si ha la proprietà c). Dunque, se A è la matrice canonica di f , l'identità

$$\langle Au, v \rangle = \langle u, Av \rangle$$

risulta vera per ogni scelta dei vettori colonna u, v . Prendendo $u = e_i$ e $v = e_j$ otteniamo

$$a_{ji} = a_{ij}$$

per ogni i, j , dunque la matrice canonica di f è simmetrica e f risulta diagonalizzabile. \square

Isoliamo la seguente proprietà degli autospazi di un endomorfismo simmetrico.

Proposizione *Gli autospazi di un endomorfismo simmetrico sono ortogonali fra loro. In altre parole, se λ_1 e λ_2 sono autovalori distinti di f , e se $u \in E(\lambda_1)$ e $v \in E(\lambda_2)$ allora:*

$$\langle u, v \rangle = 0.$$

Dimostrazione. Siccome per ipotesi $f(u) = \lambda_1 u$ e $f(v) = \lambda_2 v$:

$$\langle f(u), v \rangle = \lambda_1 \langle u, v \rangle.$$

D'altra parte, per la c) del teorema:

$$\langle f(u), v \rangle = \langle u, f(v) \rangle = \lambda_2 \langle u, v \rangle.$$

Uguagliando otteniamo $\lambda_1 \langle u, v \rangle = \lambda_2 \langle u, v \rangle$ cioè

$$(\lambda_1 - \lambda_2) \langle u, v \rangle = 0,$$

e poiché $\lambda_1 - \lambda_2 \neq 0$ si ha necessariamente $\langle u, v \rangle = 0$. \square

Esempio Sia $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definito da $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x + 3y \\ 3x + 5y \end{pmatrix}$. Verifichiamo che gli autospazi di f sono ortogonali. Matrice canonica $A = \begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}$ con polinomio caratteristico $x^2 - 2x - 24$. e abbiamo due autovalori distinti: $\lambda_1 = -4$ e $\lambda_2 = 6$ e due autospazi $E(-4), E(6)$, entrambi di dimensione 1. Si trova che $E(-4)$ ha equazione $x + 3y = 0$ con base $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$, e $E(6)$ ha equazione $3x - y = 0$ con base $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Effettivamente, gli autospazi sono ortogonali tra loro, la coppia $\begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ è una base ortogonale di autovettori, e la coppia

$$\frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

è una base ortonormale di \mathbf{R}^2 formata da autovettori di f . \square

Esempio L'endomorfismo $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3y \end{pmatrix}$ ha matrice canonica $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ dunque non è simmetrico. Si osserva che f ha autovalori $\lambda_1 = 1, \lambda_2 = 3$ e autospazi:

$$E(1) : y = 0, \quad E(3) : x - y = 0.$$

Si vede subito che gli autospazi non sono ortogonali. Risulta che f è diagonalizzabile, con base di autovettori $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ma *non è possibile* trovare una base ortonormale di autovettori (se ortonormalizziamo la base, non otteniamo più autovettori).

Esempio Sia E un sottospazio di \mathbf{R}^n . L'endomorfismo P_E che associa al vettore $v \in \mathbf{R}^n$ la sua proiezione ortogonale sul sottospazio E è simmetrico.

Infatti, se fissiamo una base ortonormale (u_1, \dots, u_k) di E allora la proiezione ortogonale è data da

$$P_E(v) = \langle v, u_1 \rangle u_1 + \dots + \langle v, u_k \rangle u_k.$$

Se w è un secondo vettore di \mathbf{R}^n si ha

$$\langle P_E(v), w \rangle = \langle v, u_1 \rangle \langle u_1, w \rangle + \dots + \langle v, u_k \rangle \langle u_k, w \rangle.$$

Poiché il secondo membro rimane uguale scambiando v con w , si ha $\langle P_E(v), w \rangle = \langle P_E(w), v \rangle = \langle v, P_E(w) \rangle$ e P_E è simmetrico.

7 Teorema spettrale

Veniamo al seguente importante teorema, di cui omettiamo la dimostrazione.

Teorema spettrale. *Sia f un endomorfismo simmetrico di \mathbf{R}^n . Allora f è diagonalizzabile; inoltre esiste una base ortonormale di \mathbf{R}^n costituita da autovettori di f .*

Anche il viceversa è vero, ed è facile da dimostrare:

Teorema *Sia f un endomorfismo di \mathbf{R}^n , e supponiamo che esista una base ortonormale di \mathbf{R}^n formata da autovettori di f . Allora f è simmetrico.*

Dimostrazione. La matrice associata alla base di autovettori (che è ortonormale per ipotesi) è diagonale, dunque simmetrica, e quindi f è simmetrico per il teorema della sezione precedente. \square

Dunque, la classe degli endomorfismi di \mathbf{R}^n che ammettono una base ortonormale di autovettori coincide con la classe degli endomorfismi simmetrici.

Notiamo anche il fatto seguente.

Corollario *Ogni matrice simmetrica è diagonalizzabile, ed è ortogonalmente simile ad una matrice diagonale. Cioè, esiste una matrice diagonale D e una matrice ortogonale M tali che:*

$$D = M^{-1}AM = M^tAM.$$

Dimostrazione. Sia f l'endomorfismo di \mathbf{R}^n rappresentato da A rispetto alla base canonica. Poiché A è simmetrica, anche f è simmetrico. Per il teorema spettrale, possiamo trovare una base ortonormale \mathcal{B} formata da autovettori di f . In questa base, f si rappresenta con una matrice diagonale D ; inoltre si ha

$$D = M^{-1}AM,$$

dove M è la matrice di passaggio dalla base canonica alla base \mathcal{B} . Poiché tali basi sono entrambe ortonormali, la matrice M è ortogonale, quindi $M^{-1} = M^t$. \square

Diamo ora il procedimento per determinare una base ortonormale di autovettori di un endomorfismo simmetrico.

1. Calcoliamo il polinomio caratteristico e quindi troviamo gli autovalori di f , diciamo $\lambda_1, \dots, \lambda_k$.
2. Con l'algoritmo di Gram-Schmidt, troviamo una base *ortonormale* di ciascun autospazio, diciamo $\mathcal{B}_1, \dots, \mathcal{B}_k$.
3. Uniamo le basi ortonormali così trovate per ottenere la base $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ di \mathbf{R}^n . L'insieme di vettori così ottenuto formerà una base ortonormale di autovettori.

Infatti, ogni vettore di \mathcal{B} ha chiaramente norma 1. Inoltre, se prendiamo due vettori appartenenti alla stessa base \mathcal{B}_i questi sono ortogonali per costruzione; se prendiamo due vettori appartenenti a basi diverse, questi appartengono ad autospazi diversi e quindi sono ortogonali grazie alla proposizione della sezione precedente. I vettori di \mathcal{B} sono a due a due ortogonali e di norma 1, dunque \mathcal{B} è una base ortonormale.

Infine, per diagonalizzare una matrice simmetrica A , procediamo così:

1. Troviamo una base ortonormale \mathcal{B} formata da autovettori dell'endomorfismo di \mathbf{R}^n rappresentato da A rispetto alla base canonica.
2. Incolonniamo la base \mathcal{B} per ottenere una matrice ortogonale M .
3. Scriviamo la matrice diagonale D , i cui elementi diagonali sono gli autovalori di f , presi nello stesso ordine dei corrispondenti autovettori di \mathcal{B} .
4. Risulterà allora $D = M^tAM$.

7.1 Esempio

Sia f l'operatore di \mathbf{R}^2 rappresentato da $\begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}$ rispetto alla base canonica. Abbiamo già trovato una base ortonormale di autovettori:

$$\frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

associati rispettivamente a -4 e 6 . Quindi se prendiamo

$$M = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} -4 & 0 \\ 0 & 6 \end{pmatrix}.$$

si avrà $D = M^t A M$.

7.2 Esempio

Sia f l'operatore di \mathbf{R}^3 rappresentato da $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ rispetto alla base canonica. f è simmetrico. Il polinomio caratteristico è $-x^3 + 3x^2$ e gli autovalori sono $0, 3$. $E(0)$ è il nucleo, di equazione $x + y + z = 0$ e base $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Applicando l'algoritmo di Gram-Schmidt, otteniamo la base ortonormale di $E(0)$:

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, w_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

$E(3)$ ha base $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; si osserva che $E(3)$ è ortogonale a $E(0)$. Otteniamo la base ortonormale di $E(3)$:

$$w_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ne segue che una base ortonormale di autovettori è (w_1, w_2, w_3) cioè:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ponendo

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

si ha $D = M^t A M$.

7.3 Esempio

Sia f l'endomorfismo di \mathbf{R}^4 rappresentato, rispetto alla base canonica, dalla matrice

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

Poiché A è simmetrica, f è un endomorfismo simmetrico. Un calcolo mostra che $p_A(x) = (x-1)^2(x-3)^2$, dunque f ammette due autovalori distinti: $\lambda_1 = 1$ e $\lambda_2 = 3$, entrambi di molteplicità algebrica 2. Già sappiamo che f è diagonalizzabile, dunque la molteplicità geometrica di entrambi gli autovalori sarà 2.

Descriviamo gli autospazi. $A - I = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$ dunque $E(1)$ ha equazioni:

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

e quindi

$$E(1) = \left\{ \begin{pmatrix} t \\ s \\ -t \\ s \end{pmatrix} \in \mathbf{R}^4 : t, s \in \mathbf{R} \right\}.$$

Procedendo in modo analogo, si ha:

$$E(3) = \left\{ \begin{pmatrix} t' \\ s' \\ t' \\ -s' \end{pmatrix} \in \mathbf{R}^4 : t', s' \in \mathbf{R} \right\}.$$

Osserviamo che i due autospazi sono fra loro ortogonali, nel senso che:

$$\left\langle \begin{pmatrix} t \\ s \\ -t \\ s \end{pmatrix}, \begin{pmatrix} t' \\ s' \\ t' \\ -s' \end{pmatrix} \right\rangle = tt' + ss' - tt' - ss' = 0$$

per ogni $t, s, t', s' \in \mathbf{R}$.

Passiamo ora a costruire una base ortonormale di autovettori di f . Una base di $E(1)$ è data dalla coppia $((1, 0, -1, 0)^t, (0, 1, 0, 1)^t)$: i due vettori sono ortogonali, dunque una base ortonormale di $V(1)$ è:

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad w_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

In modo analogo, dalla base $((1, 0, 1, 0)^t, (0, 1, 0, -1)^t)$ di $E(3)$ otteniamo la base ortonormale di $E(3)$:

$$w_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad w_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Dunque

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = (w_1, w_2, w_3, w_4)$$

è una base ortonormale di \mathbf{R}^4 , costituita da autovettori di f .

La matrice A è diagonalizzabile; se M è la matrice ottenuta incolonnando la base ortonormale di autovettori descritta in precedenza, cioè

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

allora M è una matrice *ortogonale* che diagonalizza A , nel senso che

$$M^t A M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Parte 9. Geometria del piano

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Vettori geometrici del piano, 1
- 2 Lo spazio vettoriale V_O^2 , 3
- 3 Sistemi di riferimento, 8
- 4 Equazioni parametriche di una retta, 13
- 5 Equazione cartesiana di una retta, 16
- 6 Parallelismo e intersezione di rette, 18
- 7 Il prodotto scalare in V_O^2 , 22
- 8 Distanze e angoli, 24
- 9 Perpendicolarità di rette, 26

1 Vettori geometrici del piano

1.1 Definizione

Un *vettore del piano* è una coppia ordinata di punti (A, B) . Denoteremo il vettore con il simbolo

$$\overrightarrow{AB}$$

e disegneremo il vettore con una freccia che unisce (nell'ordine) il punto A con il punto B :

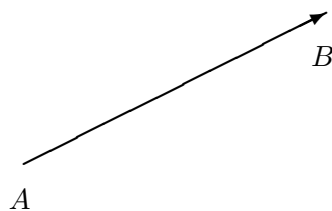


Figura 1: Vettore applicato in A

Il punto A è detto il *punto di applicazione* mentre il punto B è detto il *vertice* del vettore. Il vettore \overrightarrow{AA} è detto *vettore nullo*.

Un vettore non nullo è individuato da:

- un punto di applicazione (il punto A),
- una *direzione* (quella della retta per A e B),
- un *verso* (quello da A a B),
- un *modulo* (la lunghezza del segmento AB).

Il modulo si indica $\|\overrightarrow{AB}\|$; ovviamente il modulo di \overrightarrow{AB} uguaglia la distanza di A da B .

Vettori che hanno la stessa direzione sono contenuti in rette parallele, e si dicono *paralleli*. Per convenzione, il vettore nullo è parallelo a tutti i vettori.

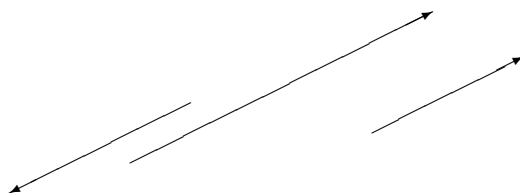


Figura 2: Vettori paralleli

Vettori che hanno stessa direzione, verso e modulo si dicono *equipollenti*. Vettori equipollenti differiscono solamente per il punto di applicazione.

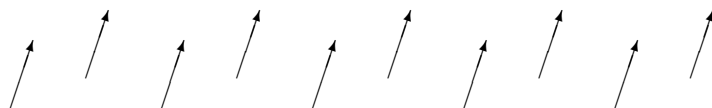


Figura 3: Vettori equipollenti

1.2 Traslazioni

È evidente dal V postulato di Euclide che, dato un vettore $\vec{v} = \overrightarrow{AB}$ e un altro punto A' , esiste un'unico punto B' tale che $\overrightarrow{A'B'}$ è equipollente a \vec{v} :

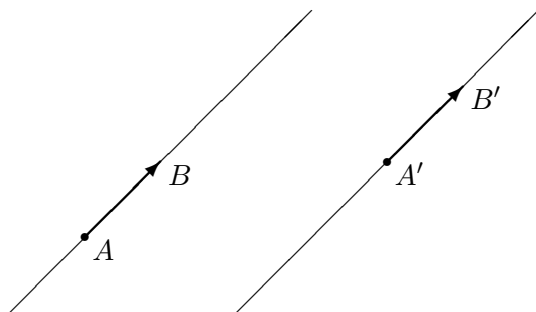


Figura 4: Traslato di un vettore

Il vettore $\overrightarrow{A'B'}$ è detto il *traslato* di \overrightarrow{AB} in A' . È evidente che allora il quadrilatero $AA'B'B$ è un parallelogramma.

- Se A' appartiene alla retta r per A e B , allora B' è l'unico punto della retta r tale che

$$\begin{cases} d(A', B') = d(A, B), \\ \overrightarrow{AB} \text{ e } \overrightarrow{A'B'} \text{ hanno lo stesso verso.} \end{cases}$$

Un vettore \vec{v} definisce così una *classe di vettori equipollenti*: ne ho uno per ogni punto del piano.

2 Lo spazio vettoriale V_O^2

2.1 Vettori applicati in un punto

Supponiamo ora di fissare un punto O del piano, detto *origine*, e consideriamo l'insieme di tutti i vettori applicati in O , denotato con V_O^2 :

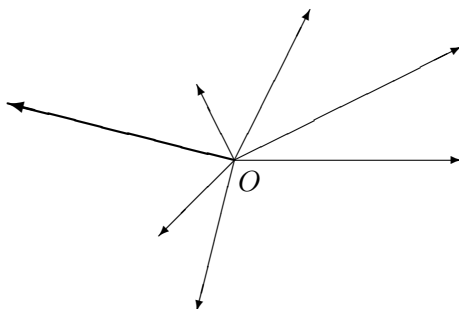


Figura 5: Vettori applicati in O

Vogliamo introdurre in V_O^2 operazioni di somma e di prodotto per uno scalare in modo da ottenere uno spazio vettoriale.

2.2 Somma di due vettori

Siano $\vec{u} = \overrightarrow{OA}$ e $\vec{v} = \overrightarrow{OB}$ due vettori applicati in O . Definiamo $\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC}$ dove

C = vertice dell'unico vettore applicato in B ed equipollente a \overrightarrow{OA} .

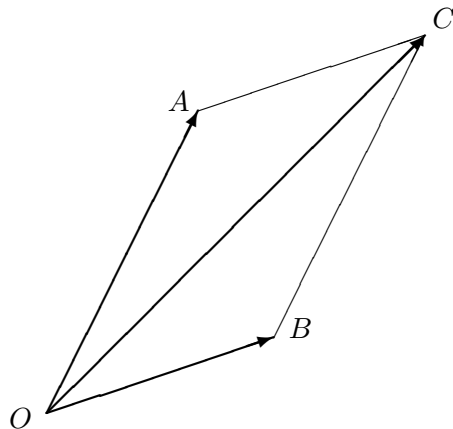
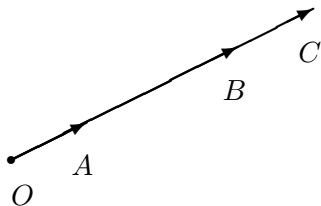


Figura 6: Regola del parallelogramma

Dalla figura è evidente che, se B non è allineato con O e A allora:

C = quarto vertice del parallelogramma su \overrightarrow{OA} e \overrightarrow{OB} .

Ma la definizione funziona anche quando O, A, B sono allineati. Nella figura seguente, si ha $\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC}$:



La somma si effettua dunque con quella che viene comunemente chiamata *regola del parallelogramma*.

2.3 Il prodotto per uno scalare

Dato $\vec{v} = \overrightarrow{OP}$ vogliamo definire il vettore $k\vec{v}$ per ogni $k \in \mathbf{R}$.

- Se $\vec{v} = \vec{O}$ definiamo $k\vec{v} = \vec{O}$; se $k = 0$ allora $k\vec{v} = \vec{O}$.

Sia ora $k \neq 0$ e $\vec{v} \neq \vec{O}$. Allora il vettore $k\vec{v}$ ha:

- la stessa direzione di \vec{v} ,
- lo stesso verso di \vec{v} se $k > 0$, verso opposto a quello di \vec{v} se $k < 0$
- modulo dato da $|k| \|\vec{v}\|$.

In pratica, $k\vec{v}$ si ottiene dilatando \vec{v} per il fattore $|k|$, e cambiando di verso se $k < 0$:



2.4 Assiomi di spazio vettoriale

Le operazioni appena introdotte verificano gli assiomi di spazio vettoriale. In particolare:

- La somma è *associativa*: $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.

Questa proprietà si può dimostrare con i risultati noti di geometria euclidea elementare.

- La somma è *commutativa*: $\vec{v} + \vec{w} = \vec{w} + \vec{v}$.

- Notiamo che: se $\vec{O} = \overrightarrow{OO}$ indica il vettore nullo di V_O^2 , allora

$$\vec{v} + \vec{O} = \vec{v}$$

per ogni $\vec{v} \in V_O^2$.

- Ogni vettore ha il suo *vettore opposto*. Infatti, se $\vec{v} = \overrightarrow{OA}$, sia A' il simmetrico di A rispetto all'origine: è evidente che $\overrightarrow{OA} + \overrightarrow{OA'} = \vec{O}$, quindi $-\vec{v} = \overrightarrow{OA'}$

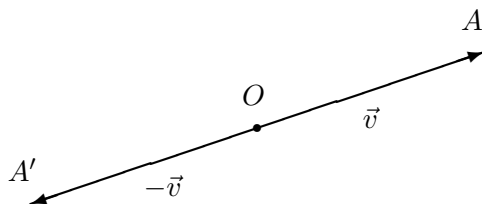


Figura 7: Vettori opposti

Le altre proprietà di spazio vettoriale sono verificate (come si dimostra facilmente), dunque:

Proposizione *L'insieme V_O^2 dei vettori applicati nell'origine, con la somma e il prodotto per uno scalare appena introdotte, è uno spazio vettoriale.*

2.5 Proprietà di V_O^2

La seguente proprietà è vera in *ogni* spazio vettoriale:

- Un vettore \vec{v} è linearmente indipendente se e solo se $\vec{v} \neq \vec{O}$.

Notiamo che se \vec{u} e \vec{v} sono vettori applicati nell'origine: $\vec{u} = \overrightarrow{OA}$, $\vec{v} = \overrightarrow{OB}$ allora \vec{u} e \vec{v} sono paralleli se e solo se i punti O, A, B sono allineati:

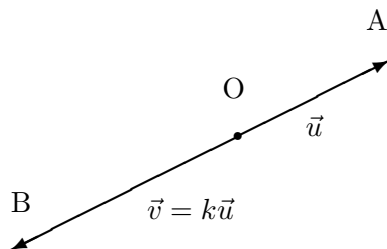


Figura 8: Vettori paralleli di V_O^2

Proposizione Due vettori \vec{u}, \vec{v} di V_O^2 sono linearmente dipendenti se e solo se sono paralleli.

Dimostrazione. Infatti sappiamo che, in uno spazio vettoriale qualunque, due vettori sono linearmente dipendenti se e solo se uno di essi è un multiplo dell'altro. È ora evidente che, ad esempio, $\vec{v} = k\vec{u}$ se e solo se \vec{u} e \vec{v} sono paralleli. \square

Nella figura, \overrightarrow{OA} e \overrightarrow{OB} sono linearmente indipendenti, mentre \overrightarrow{OC} e \overrightarrow{OD} sono linearmente dipendenti:



Figura 9: Vettori linearmente indipendenti e linearmente dipendenti

Dimostriamo ora:

Teorema La dimensione di V_O^2 è pari a 2. Le basi di V_O^2 sono tutte e sole le coppie costituite da vettori non paralleli.

Dimostrazione. Consideriamo due vettori \vec{e}_1, \vec{e}_2 non paralleli, dunque linearmente indipendenti. È sufficiente dimostrare che ogni vettore applicato in O , diciamo \overrightarrow{OP} , è una combinazione lineare di \vec{e}_1, \vec{e}_2 : in questo modo $\{\vec{e}_1, \vec{e}_2\}$ è anche un insieme di generatori, e quindi una base. Che $\{\vec{e}_1, \vec{e}_2\}$ sia un insieme di generatori è evidente dalla figura; il lettore è invitato a specificare i dettagli della costruzione geometrica che dimostra questo fatto.

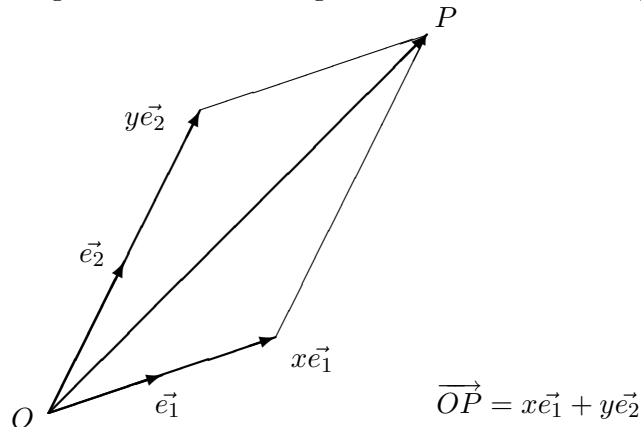


Figura 10: Base di V_O^2

Ne segue che \vec{e}_1, \vec{e}_2 sono anche generatori e formano una base. \square

Dal teorema segue in particolare che tre vettori di V_O^2 sono sempre linearmente dipendenti. In effetti, è facile dimostrare tale proprietà direttamente dalle definizioni di somma e prodotto per uno scalare, aiutandosi con una figura opportuna.

Fra tutte le basi di V_O^2 distingueremo una classe particolare di basi: le basi ortonormali.

- Una base (\vec{e}_1, \vec{e}_2) si dice *base ortonormale* se i vettori \vec{e}_1, \vec{e}_2 sono ortogonali e hanno modulo 1.

3 Sistemi di riferimento

3.1 Sistema di riferimento affine

Un *sistema di riferimento affine* nel piano consiste nella scelta di un punto O detto *origine* e di una base (\vec{e}_1, \vec{e}_2) di V_O^2 .

Un sistema di riferimento permette di descrivere i punti del piano con una coppia di numeri, le *coordinate* del punto. Precisamente, dato un punto P , consideriamo il vettore \overrightarrow{OP} , ed esprimiamo tale vettore come combinazione lineare dei vettori della base scelta:

$$\overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2.$$

Diremo allora che P ha coordinate (x, y) e identificheremo il punto P con le sue coordinate:

$$P = (x, y).$$

L'origine ha coordinate $(0, 0)$. La retta su cui giace \vec{e}_1 è detta *asse x*, mentre la retta su cui giace \vec{e}_2 è detta *asse y*. Gli assi sono orientati in modo concorde con i vettori della base.

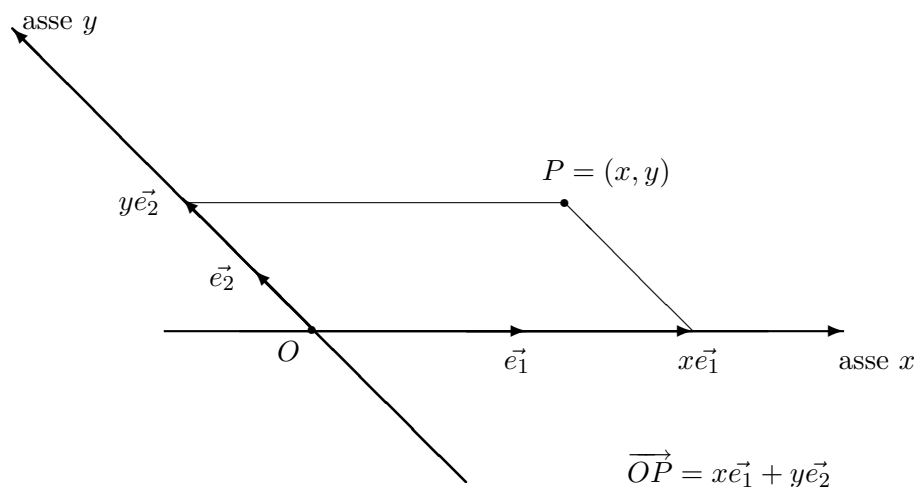


Figura 11: Sistema di riferimento affine e coordinate di un punto

3.2 Sistema di riferimento cartesiano

Se la base (\vec{e}_1, \vec{e}_2) scelta è ortonormale, nel senso che i vettori \vec{e}_1, \vec{e}_2 sono ortogonali e di modulo 1, diremo allora che il sistema di riferimento è *cartesiano*, e le coordinate di un punto P sono dette *coordinate cartesiane*.

In un sistema di riferimento cartesiano gli assi coordinati sono ortogonali fra loro. Inoltre, dato un punto P del piano, detto P_1 (risp. P_2) il piede della perpendicolare condotta da P_1 all'asse x (risp. asse y), si ha che, se $P = (x, y)$ allora

$$\begin{cases} x = \text{ascissa di } P = \text{distanza (con segno) di } P_1 \text{ dall'origine} \\ y = \text{ordinata di } P = \text{distanza (con segno) di } P_2 \text{ dall'origine} \end{cases}$$

L'ascissa x è positiva se P_1 segue l'origine rispetto al verso dell'asse x (orientato in modo concorde al verso del vettore \vec{e}_1). L'ascissa è nulla se il punto appartiene all'asse y ed è negativa se P_1 precede l'origine. Simili considerazioni valgono riguardo all'ordinata. Nella figura che segue, il punto P ha coordinate positive.

Per definizione, si ha $\overrightarrow{OP} = \overrightarrow{OP_1} + \overrightarrow{OP_2} = x\vec{e}_1 + y\vec{e}_2$.

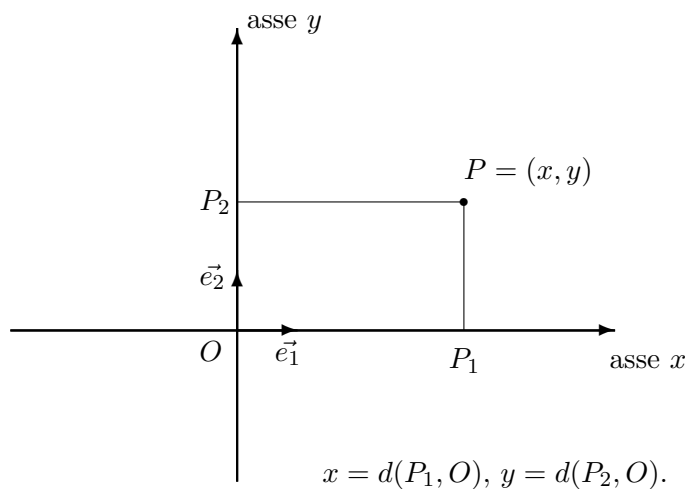


Figura 12: Coordinate cartesiane

D'ora in poi sceglieremo sempre un sistema di riferimento cartesiano.

3.3 Coordinate di un vettore

Teorema Se il vettore \overrightarrow{OP} ha coordinate (x, y) e il vettore $\overrightarrow{OP'}$ ha coordinate (x', y') allora:

- a) Il vettore $k\overrightarrow{OP}$ ha coordinate $k(x, y) = (kx, ky)$.
- b) Il vettore $\overrightarrow{OP} + \overrightarrow{OP'}$ ha coordinate $(x, y) + (x', y') = (x + x', y + y')$.

Dimostrazione. È sufficiente osservare che, per definizione, $\overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2$ e $\overrightarrow{OP'} = x'\vec{e}_1 + y'\vec{e}_2$ e applicare le proprietà di spazio vettoriale. \square

Il teorema si enuncia dicendo anche che, se $F : V_O^2 \rightarrow \mathbf{R}^2$ è l'applicazione che associa a un vettore le sue coordinate rispetto alla base scelta, allora F è un'applicazione lineare; tale applicazione è inoltre biiettiva, dunque è *isomorfismo di spazi vettoriali*. (In realtà questo fatto è vero in ogni spazio vettoriale).

Spesso identificheremo anche un vettore con le sue coordinate. Dunque la scrittura

$$\overrightarrow{OP} = (x, y)$$

significherà

$$\overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2.$$

Vogliamo ora attribuire coordinate a un vettore \overrightarrow{AB} applicato in un punto qualunque A del piano.

Definizione *Le coordinate del vettore \overrightarrow{AB} sono, per definizione, le coordinate del vettore \overrightarrow{OC} , traslato di \overrightarrow{AB} nell'origine.*

(Vedi Figura 12).

Proposizione *Le coordinate del vettore \overrightarrow{AB} sono date dalla differenza fra le coordinate del vertice e quelle del punto di applicazione. In altre parole, se $A = (x, y)$ e $B = (x', y')$ allora le coordinate del vettore \overrightarrow{AB} sono*

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

- Spesso scriveremo, in breve, che le coordinate di \overrightarrow{AB} sono $B - A$.

Dimostrazione. Se \overrightarrow{OC} è il traslato di \overrightarrow{AB} nell'origine, dalla legge del parallelogramma abbiamo che $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{OC}$ dunque:

$$\overrightarrow{OC} = \overrightarrow{OB} - \overrightarrow{OA}.$$

Le coordinate di \overrightarrow{AB} sono, per definizione, quelle di \overrightarrow{OC} : prendendo le coordinate di ambo i membri della relazione vettoriale precedente, vediamo che le coordinate di \overrightarrow{OC} valgono $(x', y') - (x, y) = (x' - x, y' - y)$. \square

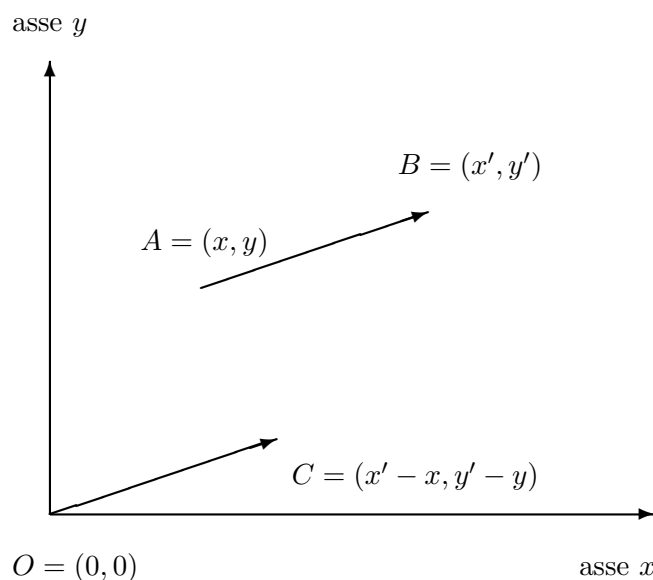


Figura 13: Traslato nell'origine

È evidente che vettori equipollenti hanno lo stesso traslato nell'origine, dunque hanno coordinate uguali. È vero anche il viceversa. In conclusione

- due vettori del piano sono equipollenti se e solo se hanno coordinate uguali.

Un vettore del piano è determinato dalle sue coordinate, e dal punto di applicazione.

Esempio Trovare il vertice dell'unico vettore del piano di coordinate $(-1, 4)$ e applicato nel punto $A = (2, 3)$.

Soluzione. Se $B = (x, y)$ è il vertice del vettore cercato, si dovrà avere

$$B - A = (-1, 4),$$

da cui $B = (2, 3) + (-1, 4) = (1, 7)$. \square

Esempio Sono dati i punti $A = (1, 1)$, $B = (2, 3)$, $C = (3, -1)$. Determinare le coordinate del punto D , quarto vertice del parallelogramma costruito su AB e AC .

Soluzione. Risolviamo con l'algebra dei vettori. Dalla legge del parallelogramma il punto D è tale che:

$$\overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AD}.$$

Prendendo le coordinate di ambo i membri, avremo quindi un'uguaglianza tra coppie di numeri: $B - A + C - A = D - A$, e dunque:

$$D = B + C - A = (2, 3) + (3, -1) - (1, 1) = (4, 1),$$

Quindi $D = (4, 1)$. \square

4 Equazioni parametriche di una retta

D'ora in poi fissiamo nel piano un sistema di coordinate cartesiane, con origine O e base ortonormale di riferimento (\vec{e}_1, \vec{e}_2) .

4.1 Vettore direttore, parametri direttori

Data una retta r del piano, un vettore non nullo $\vec{v} = \overrightarrow{OQ}$ si dice *vettore direttore* di r se \vec{v} è parallelo a r .

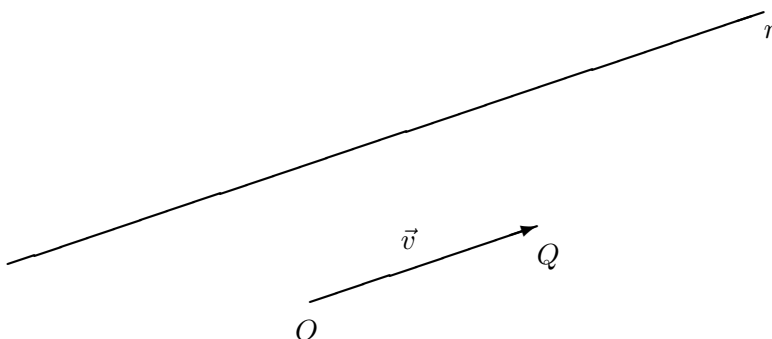


Figura 14: Vettore direttore di r

Definizione Le coordinate di un qualunque vettore direttore di r sono dette *parametri direttori* di r .

È chiaro che \overrightarrow{OQ} è un vettore direttore di r se e solo se Q appartiene alla retta per l'origine parallela a r : quindi una stessa retta ha infiniti vettori direttori, tutti paralleli tra loro. Poiché vettori paralleli hanno coordinate proporzionali, si avrà che

- Una retta ha infiniti parametri direttori, tutti proporzionali tra loro.

Diremo anche che i parametri direttori di una retta sono definiti *a meno di un fattore di proporzionalità*.

Proposizione Data una retta r del piano, sia P_0 un punto di r e $\vec{v} = \overrightarrow{OQ}$ un vettore direttore di r . Allora il punto P appartiene alla retta r se e solo se esiste $t \in \mathbf{R}$ tale che:

$$\overrightarrow{OP} = \overrightarrow{OP_0} + t\vec{v}.$$

Dimostrazione. Sia r' la retta per l'origine parallela a r . Con riferimento alla figura sottostante, consideriamo il punto Q' su r' tale che $\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{OQ'}$. Ora $\overrightarrow{OQ'}$ è parallelo a \overrightarrow{OQ} , dunque esiste $t \in \mathbf{R}$ tale che $\overrightarrow{OQ'} = t\overrightarrow{OQ}$ e si ha $\overrightarrow{OP} = \overrightarrow{OP_0} + t\vec{v}$.

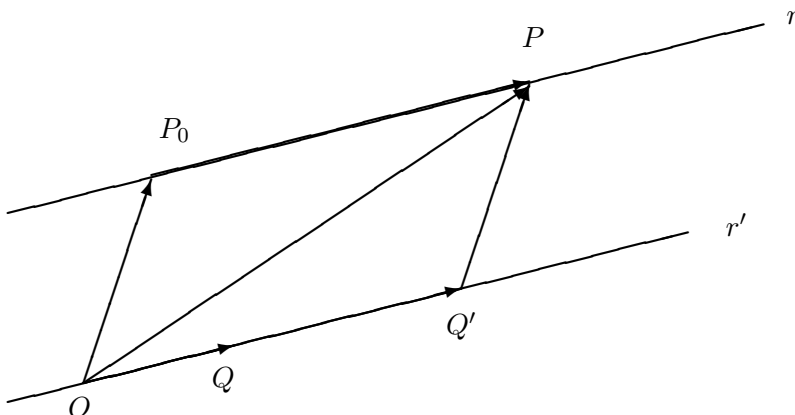


Figura 15: Equazione vettoriale di una retta

□

4.2 Equazioni parametriche

Una retta del piano è individuata da un suo punto e da una direzione. La direzione è specificata da un vettore direttore di r , che a sua volta è individuato dalle sue coordinate (parametri direttori di r). Dunque sarà possibile determinare una retta mediante

- un suo punto P_0 ,
- i suoi parametri direttori (l, m) .

Proposizione Sia r una retta, $P_0 = (x_0, y_0)$ un suo punto, e $\vec{v} = \overrightarrow{OQ}$ un suo vettore direttore, di coordinate (l, m) . Allora r può essere descritta dalle seguenti equazioni, dette equazioni parametriche della retta r :

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \end{cases}$$

Le equazioni vanno interpretate nel seguente modo: il punto (x, y) appartiene alla retta r se e solo se esiste $t \in \mathbf{R}$ tale che le due equazioni sono entrambe verificate. Al variare di $t \in \mathbf{R}$ otteniamo dall'equazione le coordinate tutti i punti di r .

Dimostrazione. Dalla equazione vettoriale della retta dimostrata nella proposizione precedente, abbiamo che il punto P appartiene alla retta r se e solo se

$$\overrightarrow{OP} = \overrightarrow{OP_0} + t\overrightarrow{OQ}$$

per qualche $t \in \mathbf{R}$. Dette (x, y) le coordinate di P , abbiamo la seguente uguaglianza tra vettori di \mathbf{R}^2 :

$$(x, y) = (x_0, y_0) + t(l, m),$$

e la tesi è dimostrata. \square

Esempio a) Scrivere le equazioni parametriche della retta r passante per il punto $P_0 = (1, -2)$ e parallela al vettore $\vec{v} = \vec{e}_1 + 3\vec{e}_2$.

b) Stabilire se il punto $A = (3, 2)$ appartiene alla retta.

Soluzione. a) Il vettore \vec{v} è un vettore direttore della retta, dunque i parametri direttori di r sono le coordinate di \vec{v} , cioè: $(l, m) = (1, 3)$. Otteniamo le seguenti equazioni parametriche:

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

b) Il punto A appartiene a r se e solo se esiste $t \in \mathbf{R}$ tale che:

$$\begin{cases} 3 = 1 + t \\ 2 = -2 + 3t \end{cases}$$

Si vede subito che le due equazioni sono incompatibili, dunque $A \notin r$. \square

Esempio Le equazioni $\begin{cases} x = 2 + 3t \\ y = 9t \end{cases}$ rappresentano la retta r' per il punto $Q_0 = (2, 0)$ di parametri direttori $(3, 9)$.

4.3 Retta passante per due punti

Proposizione Siano $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ due punti distinti del piano, e sia r la retta passante per P_1 e P_2 .

a) I parametri direttori di r sono proporzionali alla coppia $P_2 - P_1$, cioè a:

$$\begin{cases} l = x_2 - x_1 \\ m = y_2 - y_1 \end{cases}$$

b) *Equazioni parametriche di r sono date da*

$$\begin{cases} x = x_1 + (x_2 - x_1)t \\ y = y_1 + (y_2 - y_1)t \end{cases}$$

Dimostrazione. a) Il vettore $\overrightarrow{P_1P_2}$ è parallelo a r , e il suo traslato nell'origine \vec{v} sarà

dunque un vettore direttore di r . È sufficiente osservare che \vec{v} ha coordinate uguali a quelle di $\overrightarrow{P_1P_2}$, cioè $P_2 - P_1 = (x_2 - x_1, y_2 - y_1)$.

b) Basta osservare che r passa per P_1 e che, per la a), i parametri direttori sono proporzionali alla coppia $(x_2 - x_1, y_2 - y_1)$. \square

Esempio Scrivere le equazioni parametriche della retta passante per i punti $P_1 = (1, 3)$ e $P_2 = (5, -1)$.

Soluzione. I parametri direttori di r sono proporzionali a $P_2 - P_1 = (4, -4)$; la retta passa per $P_1 = (1, 3)$ dunque:

$$r : \begin{cases} x = 1 + 4t \\ y = 3 - 4t \end{cases}$$

Come già osservato, le equazioni parametriche *non* sono uniche. Ad esempio, potevamo prendere come punto iniziale il punto $P_2 = (5, -1)$ e come parametri direttori la coppia $(-1, 1)$, proporzionale alla precedente. La stessa retta r si può quindi rappresentare anche nel seguente modo:

$$r : \begin{cases} x = 5 - t \\ y = -1 + t \end{cases}$$

Notiamo che, eliminando il parametro t , otteniamo in entrambi i casi l'equazione nelle incognite x, y :

$$x + y - 4 = 0,$$

detta *equazione cartesiana* della retta r . Approfondiremo questo aspetto nella prossima sezione. \square

5 Equazione cartesiana di una retta

In questa sezione vedremo come rappresentare una retta con una singola equazione in due variabili.

5.1 Condizione di allineamento di tre punti

Sappiamo che, nel piano, due punti sono sempre allineati (esiste cioè una retta che li contiene entrambi). In generale, però, tre punti possono risultare non allineati.

Proposizione *I punti $P_1 = (x_1, y_1), P_2 = (x_2, y_2), P_3 = (x_3, y_3)$ sono allineati se e solo se*

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} = 0.$$

Dimostrazione. I tre punti sono allineati se e solo se i vettori $\overrightarrow{P_1P_2}$ e $\overrightarrow{P_1P_3}$ applicati nel punto P_1 sono paralleli, dunque linearmente dipendenti. Questo avviene se e solo se le rispettive coordinate: $P_2 - P_1 = (x_2 - x_1, y_2 - y_1)$ e $P_3 - P_1 = (x_3 - x_1, y_3 - y_1)$ sono vettori linearmente dipendenti di \mathbf{R}^2 , da cui la condizione. Più in generale osserviamo (senza dimostrazione) la seguente interpretazione geometrica: il *modulo* del determinante

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}$$

uguaglia l'area del parallelogramma costruito sui vettori $\overrightarrow{P_1P_2}$ e $\overrightarrow{P_1P_3}$: tale area è nulla esattamente quando i punti sono allineati. \square

Esempio Stabilire se i punti $A = (1, 1), B = (2, 2), C = (2, -3)$ sono allineati oppure no.

Soluzione. Basta esaminare il determinante $\begin{vmatrix} 1 & 1 \\ 1 & -4 \end{vmatrix}$: vale -5 , dunque è diverso da zero e i punti non sono allineati. Infatti, l'area del parallelogramma costruito sui vettori \overrightarrow{AB} e \overrightarrow{AC} vale $|-5| = 5$. \square

5.2 Equazione cartesiana di una retta

Abbiamo la seguente

Proposizione a) *Ogni retta del piano si rappresenta con un'equazione lineare in due variabili, del tipo:*

$$ax + by + c = 0, \tag{1}$$

con a, b non entrambi nulli, detta equazione cartesiana della retta. Questo significa che il punto $P = (x, y)$ appartiene alla retta se e solo se le sue coordinate verificano l'equazione (1).

b) L'equazione cartesiana della retta per i due punti distinti $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ è:

$$\begin{vmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix} = 0$$

Dimostrazione. Scegliamo due punti distinti della retta, diciamo $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$. Allora il punto generico $P = (x, y)$ del piano appartiene a r se e solo esso risulta allineato con P_1, P_2 . Dalla condizione di allineamento otteniamo dunque (dopo uno scambio di righe, che non altera l'equazione)

$$\begin{vmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix} = 0.$$

Sviluppando esplicitamente il determinante otteniamo un'equazione del tipo descritto nella parte a). Si noti in particolare che $a = (y_2 - y_1)$, $b = -(x_2 - x_1)$ non possono essere entrambi nulli altrimenti $P_1 = P_2$. \square

Esempio L'equazione cartesiana della retta per i punti $A = (2, 1)$, $B = (1, -2)$ è

$$\begin{vmatrix} x - 2 & y - 1 \\ -1 & -3 \end{vmatrix} = 0, \quad \text{cioè } 3x - y - 5 = 0.$$

5.3 Forme particolari

Vediamo alcuni casi particolari dell'equazione (1).

- Se $c = 0$ la retta passa per l'origine.

Infatti l'equazione diventa $ax + by = 0$ e $O = (0, 0)$ è evidentemente una soluzione.

- Se $a = 0$ la retta è parallela all'asse x .

Infatti si ha $by + c = 0$, cioè $y = -\frac{c}{b}$: i punti della retta hanno ordinata costante, uguale a $-\frac{c}{b}$ e la retta è parallela all'asse x .

- Se $b = 0$ la retta è parallela all'asse y .

6 Parallelismo e intersezione di rette

Osserviamo che due rette del piano r, r' sono parallele (cioè hanno la stessa direzione) se coincidono ($r = r'$) o non hanno punti comuni. In particolare, r e r' sono non parallele se e solo se hanno unico punto di intersezione.

Proposizione Siano $r : ax + by + c = 0$ e $r' : a'x + b'y + c' = 0$ due rette del piano. Allora:

a) r e r' sono parallele se e solo se:

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = 0,$$

quindi se e solo se i rispettivi coefficienti sono proporzionali: $(a', b') = k(a, b)$ per qualche $k \in \mathbf{R}$;

b) r e r' si incontrano in un unico punto se e solo se

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \neq 0$$

Dimostrazione. È sufficiente dimostrare b). Le rette si incontrano in un unico punto se e solo se il sistema lineare

$$\begin{cases} ax + by + c = 0 \\ a'x + b'y + c' = 0 \end{cases}$$

ammette un'unica soluzione. Per il teorema di Cramer, questo avviene se e solo se il determinante della matrice dei coefficienti è non nullo:

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \neq 0,$$

da cui la condizione. \square

Esempio Le rette $r : x + \frac{1}{2}y + 5 = 0$ e $r' : 2x + y - 3 = 0$ hanno coefficienti proporzionali, quindi sono parallele. Osserviamo che l'equazione di r è equivalente a $2x + y + 10 = 0$ e quindi le rette sono distinte.

Esempio Le rette $r : x + \frac{1}{2}y + 5 = 0$ e $r' : 2x + y + 10 = 0$ sono parallele e coincidenti.

Esempio Le rette $r : x + y - 3 = 0$ e $r' : 2x + y + 1 = 0$ non sono parallele, quindi hanno un unico punto in comune, ottenuto risolvendo il sistema

$$\begin{cases} x + y - 3 = 0 \\ 2x + y + 1 = 0 \end{cases}$$

Troviamo il punto $(-4, 7)$.

6.1 Parametri direttori

Abbiamo visto due possibili rappresentazioni di una retta del piano:

- tramite equazioni parametriche,
- tramite un'equazione cartesiana.

Il passaggio dalle equazioni parametriche all'equazione cartesiana avviene eliminando il parametro.

Esempio La retta di equazioni parametriche $r : \begin{cases} x = 1 - 2t \\ y = 2 + t \end{cases}$ ha equazione cartesiana $x + 2y + 5 = 0$ ottenuta sostituendo $t = y - 2$ nella prima equazione. Notiamo che r ha parametri direttori proporzionali a $(-2, 1)$.

Il passaggio dall'equazione cartesiana alle equazioni parametriche avviene risolvendo l'equazione.

Esempio La retta di equazione cartesiana $2x + y + 5 = 0$ ha equazioni parametriche $\begin{cases} x = t \\ y = -1 - 2t \end{cases}$.

In particolare, ha parametri direttori proporzionali a $(1, -2)$.

- In generale, i parametri direttori della retta r di equazione $ax + by + c = 0$ sono proporzionali a $(b, -a)$; in altre parole, r è parallela al vettore direttore $b\vec{e}_1 - a\vec{e}_2$.

6.2 Fasci di rette

6.3 Fascio di rette parallele

Fissata una retta r , il *fascio di rette parallele a r* è la totalità delle rette del piano parallele a r .

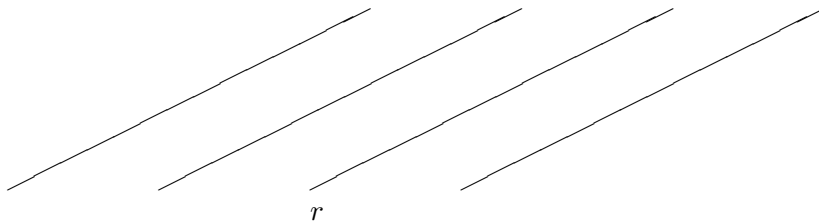


Figura 16: Fascio di rette parallele

Esempio È data la retta $r : x + 3y - 5 = 0$. Allora l'equazione della retta generica parallela a r si scrive

$$r' : x + 3y + k = 0,$$

con $k \in \mathbf{R}$, detta anche *equazione del fascio di rette parallele a r* .

Infatti, la retta $r' : ax + by + c = 0$ è parallela a r se e solo se $(a, b) = \lambda(1, 3)$, dunque: $r' : \lambda x + 3\lambda y + c = 0$, ma poiché $\lambda \neq 0$ possiamo dividere per λ e abbiamo $r' : x + 3y + k = 0$, dove abbiamo posto $k = \frac{c}{\lambda}$.

Esempio È data la retta $r : x + 3y - 5 = 0$. Determinare l'equazione cartesiana della retta r' parallela a r e passante per $(1, -1)$.

Soluzione. Abbiamo visto che l'equazione della retta generica parallela a r è

$$r' : x + 3y + k = 0.$$

Sostituendo le coordinate $(1, -1)$ nell'equazione otteniamo $k = 2$, dunque la retta cercata ha equazione

$$x + 3y + 2 = 0.$$

□

In generale, data la retta $r : ax + by + c = 0$ allora la retta generica parallela a r ha equazione $ax + by + k = 0$ con k parametro reale. In particolare, le rette del piano parallele a una data retta sono ∞^1 .

In conclusione, per trovare la retta r' parallela ad una data retta r e passante per P , possiamo procedere in questo modo:

1. scrivere l'equazione del fascio di rette parallele a r ,
2. imporre il passaggio per il punto P .

6.4 Fascio di rette passanti per un punto

Fissato un punto $P_0 = (x_0, y_0)$ il *fascio di rette passanti per P_0* è la totalità delle rette che passano per P_0 .

Si vede facilmente che l'equazione della retta generica di tale fascio si scrive:

$$a(x - x_0) + b(y - y_0) = 0,$$

Esempio È data la retta $r : x + 3y - 5 = 0$. Determinare l'equazione della retta r' parallela a r e passante per il punto $P_0 = (1, -1)$.

Soluzione. È l'esercizio del precedente paragrafo, che risolviamo in modo diverso. Scriviamo l'equazione del fascio di rette passanti per P_0 :

$$a(x - 1) + b(y + 1) = 0,$$

ovvero $ax + by - a + b = 0$. Ora imponiamo il parallelismo con r : basta prendere $(a, b) = (1, 3)$ da cui l'equazione

$$x + 3y + 2 = 0.$$

□

7 Il prodotto scalare in V_O^2

7.1 Definizione

Il *prodotto scalare* è un'operazione che associa a due vettori di V_O^2 un numero reale, denotato con $\langle \vec{v}, \vec{w} \rangle$. Precisamente, se \vec{v}, \vec{w} sono vettori applicati in O , definiamo:

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

dove θ è l'angolo convesso (cioè tale che $0 \leq \theta \leq \pi$) definito da \vec{v} e \vec{w} .

- Parlare di angolo ha senso solo quando sia \vec{v} che \vec{w} sono non nulli. Ma se uno dei due vettori è nullo poniamo per definizione $\langle \vec{v}, \vec{w} \rangle = 0$.

Abbiamo immediatamente:

Proposizione a) *Due vettori sono ortogonali se e solo se il loro prodotto scalare è nullo.*

b) *Si ha $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$.*

Dunque il prodotto scalare permette di misurare il modulo (lunghezza) di un vettore, e l'angolo fra due vettori. La proposizione seguente sarà utile per il calcolo esplicito del prodotto scalare.

Proposizione *Il prodotto scalare ha le seguenti proprietà. In ciò che segue, \vec{v}, \vec{w}, \dots sono vettori arbitrari applicati in O e $k \in \mathbf{R}$.*

a) $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$.

b) $\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle$.

c) $\langle \vec{v}, k\vec{w} \rangle = k\langle \vec{v}, \vec{w} \rangle$.

d) $\langle \vec{v}, \vec{v} \rangle \geq 0$.

e) $\langle \vec{v}, \vec{v} \rangle = 0$ se e solo se $\vec{v} = \vec{O}$.

La proprietà a) dice che il prodotto scalare è commutativo. b) e c) esprimono le proprietà di *bilinearità*. Le proprietà d) e e) sono dette di *positività*.

Dimostrazione. Le proprietà sono di facile dimostrazione, ad eccezione della b), che richiede un po' di lavoro in più, e di cui omettiamo la dimostrazione. \square

Osserviamo che, da a) e b), otteniamo:

$$\begin{aligned}\langle \vec{v}, a\vec{w}_1 + b\vec{w}_2 \rangle &= a\langle \vec{v}, \vec{w}_1 \rangle + b\langle \vec{v}, \vec{w}_2 \rangle \\ \langle a\vec{v}_1 + b\vec{v}_2, \vec{w} \rangle &= a\langle \vec{v}_1, \vec{w} \rangle + b\langle \vec{v}_2, \vec{w} \rangle\end{aligned}$$

7.2 Calcolo in un riferimento cartesiano

Ricordiamo che assegnare un riferimento cartesiano equivale a fissare un'origine O e una base ortonormale di V_O^2 , diciamo (\vec{e}_1, \vec{e}_2) . Allora ogni vettore si esprime

$$\vec{v} = \overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2.$$

La coppia (x, y) dà luogo alle coordinate di \vec{v} nel riferimento scelto, e scriveremo $\vec{v} = (x, y)$. Ora, conoscendo le coordinate, possiamo calcolare il prodotto scalare con facilità.

Proposizione Siano $\vec{v} = (x, y)$ e $\vec{w} = (x', y')$. Allora

$$\langle \vec{v}, \vec{w} \rangle = xx' + yy'.$$

In particolare, $\|\vec{v}\| = \sqrt{x^2 + y^2}$.

Dimostrazione. I prodotti scalari degli elementi della base ortonormale (\vec{e}_1, \vec{e}_2) si calcolano facilmente:

$$\begin{cases} \langle \vec{e}_1, \vec{e}_1 \rangle = \langle \vec{e}_2, \vec{e}_2 \rangle = 1 \\ \langle \vec{e}_1, \vec{e}_2 \rangle = \langle \vec{e}_2, \vec{e}_1 \rangle = 0 \end{cases}$$

Per ipotesi, $\vec{v} = x\vec{e}_1 + y\vec{e}_2$, $\vec{w} = x'\vec{e}_1 + y'\vec{e}_2$, dunque, utilizzando le proprietà algebriche enunciate in precedenza:

$$\begin{aligned}\langle \vec{v}, \vec{w} \rangle &= \langle x\vec{e}_1 + y\vec{e}_2, x'\vec{e}_1 + y'\vec{e}_2 \rangle \\ &= xx'\langle \vec{e}_1, \vec{e}_1 \rangle + xy'\langle \vec{e}_1, \vec{e}_2 \rangle + yx'\langle \vec{e}_2, \vec{e}_1 \rangle + yy'\langle \vec{e}_2, \vec{e}_2 \rangle \\ &= xx' + yy'.\end{aligned}$$

\square

Esempio Calcolare il prodotto scalare tra $\vec{v} = 3\vec{e}_1 - \vec{e}_2$ e $\vec{w} = -\vec{e}_1 + 5\vec{e}_2$.

Soluzione. Si ha $\vec{v} = (3, -1)$, $\vec{w} = (-1, 5)$ quindi $\langle \vec{v}, \vec{w} \rangle = -8$. \square

Ricordiamo che, dato un vettore $\vec{v} = \overrightarrow{AB}$ del piano applicato in un qualunque punto A , le coordinate di \vec{v} sono le coordinate del traslato di \vec{v} nell'origine, diciamo \vec{v}_0 . Poiché la traslazione di due vettori conserva sia i moduli che l'angolo compreso fra di essi, il prodotto scalare fra \vec{v} e \vec{w} è uguale al prodotto scalare dei rispettivi vettori traslati \vec{v}_0 e \vec{w}_0 . Dunque abbiamo

Proposizione *Siano \vec{v} e \vec{w} vettori del piano applicati in un qualunque punto A del piano, di coordinate rispettive (x, y) e (x', y') . Allora:*

$$\langle \vec{v}, \vec{w} \rangle = xx' + yy'.$$

In particolare, $\|\vec{v}\| = \sqrt{x^2 + y^2}$.

Ricordiamo la definizione di prodotto scalare fra due vettori (riga) di \mathbf{R}^2 :

$$\langle (x, y), (x', y') \rangle = xx' + yy'.$$

Allora si ha

Proposizione *Il prodotto scalare di due vettori del piano (applicati in un qualunque punto) uguaglia il prodotto scalare delle rispettive coordinate.*

8 Distanze e angoli

8.1 Formula della distanza

Se $A = (x_1, y_1)$ e $B = (x_2, y_2)$ sono due punti del piano, allora la distanza di A da B è data dalla formula:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Infatti, la distanza di A da B uguaglia il modulo del vettore $\vec{v} = \overrightarrow{AB}$ che ha coordinate

$$B - A = (x_2 - x_1, y_2 - y_1),$$

e quindi la formula segue dalla proposizione precedente.

Esempio Se $A = (2, -1)$ e $B = (1, 3)$ allora

$$d(A, B) = \sqrt{(1 - 2)^2 + (3 + 1)^2} = \sqrt{17}$$

8.2 Angoli

Dalla formula che definisce il prodotto scalare, vediamo subito che il coseno dell'angolo compreso fra i vettori non nulli \vec{v} e \vec{w} è dato da:

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

Esempio Calcolare il perimetro e il coseno degli angoli interni del triangolo di vertici $A = (1, 1), B = (2, 3), C = (3, 0)$.

Soluzione. Dalla formula della distanza otteniamo $d(A, B) = \sqrt{5} = d(A, C)$ mentre $d(B, C) = \sqrt{10}$. Il perimetro vale dunque $2\sqrt{5} + \sqrt{10}$. Notiamo che il triangolo è isoscele sulla base BC . Calcoliamo il coseno dell'angolo in A , diciamo θ_1 . Se $\vec{v} = \overrightarrow{AB}$ e $\vec{w} = \overrightarrow{AC}$ allora

$$\cos \theta_1 = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \|\vec{w}\|}$$

Ora \vec{v} ha coordinate $B - A = (1, 2)$ mentre \vec{w} ha coordinate $C - A = (2, -1)$. Dunque

$$\langle \vec{v}, \vec{w} \rangle = \langle (1, 2), (2, -1) \rangle = 0.$$

Quindi i due vettori sono ortogonali e l'angolo è $\theta_1 = \pi/2$. Il triangolo è allora isoscele rettangolo. È a questo punto evidente che gli altri due angoli sono entrambi uguali a $\pi/4$. Verifichiamo calcolando l'angolo al vertice B . Il suo coseno vale:

$$\cos \theta_2 = \frac{\langle \overrightarrow{BA}, \overrightarrow{BC} \rangle}{\|\overrightarrow{BA}\| \|\overrightarrow{BC}\|}.$$

Ora \overrightarrow{BA} ha coordinate $A - B = (-1, -2)$ mentre \overrightarrow{BC} ha coordinate $C - B = (1, -3)$. Quindi

$$\langle \overrightarrow{BA}, \overrightarrow{BC} \rangle = \langle (-1, -2), (1, -3) \rangle = 5.$$

D'altra parte $\|\overrightarrow{BA}\| = \sqrt{5}$ mentre $\|\overrightarrow{BC}\| = \sqrt{10}$. Dunque:

$$\cos \theta_2 = \frac{5}{\sqrt{5}\sqrt{10}} = \sqrt{\frac{1}{2}},$$

e $\theta_2 = \pi/4$. \square

9 Perpendicolarità di rette

Consideriamo due rette $r : ax + by + c = 0$, $r' : a'x + b'y + c' = 0$. Vogliamo enunciare una condizione che caratterizzi la perpendicolarità delle due rette. Ora sappiamo che i parametri direttori di r sono dati dalla coppia $(b, -a)$, che rappresenta le coordinate di un vettore parallelo alla retta: precisamente il vettore $\vec{v} = b\vec{e}_1 - a\vec{e}_2$. Analogamente, il vettore $\vec{w} = b'\vec{e}_1 - a'\vec{e}_2$ è parallelo alla retta r' . Dunque le rette saranno perpendicolari se e solo se i rispettivi vettori direttori sono ortogonali:

$$\langle \vec{v}, \vec{w} \rangle = 0,$$

quindi se e solo se $0 = \langle (b, -a), (b', -a') \rangle = bb' + aa'$. Abbiamo dimostrato la seguente:

Proposizione *Le rette $r : ax + by + c = 0$, $r' : a'x + b'y + c' = 0$ sono perpendicolari se e solo se:*

$$aa' + bb' = 0.$$

Esempio Nel piano sono dati la retta $r : 2x - y + 2 = 0$ e il punto $P = (2, 0)$. Determinare l'equazione della retta passante per P e ortogonale alla retta r .

Soluzione. Scriviamo l'equazione del fascio di rette perpendicolari a r e imponiamo il passaggio per P . Si ha che la retta:

$$r' : x + 2y + k = 0$$

è senz'altro ortogonale a r per ogni k . Sostituendo le coordinate di $P = (2, 0)$ nell'equazione otteniamo $k = -2$. Dunque la retta cercata ha equazione $x + 2y - 2 = 0$. \square

Osserviamo che se $r : ax + by + c = 0$ allora la coppia (a, b) rappresenta le coordinate di un vettore perpendicolare alla retta.

9.1 Proiezione ortogonale di un punto su una retta

Dato un punto P e una retta r , possiamo condurre la retta r' passante per P perpendicolare a r . La retta r' incontra r in un punto H , detto la *proiezione ortogonale* di P su r .

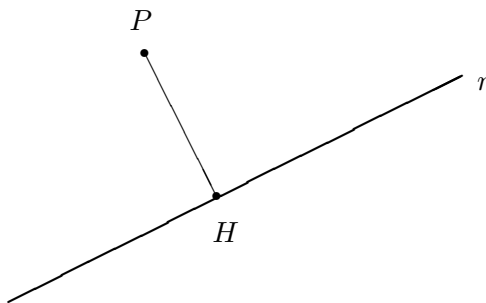


Figura 17: Proiezione ortogonale di un punto su una retta

Esempio Determinare la proiezione ortogonale di $P_0 = (2, 0)$ sulla retta $2x - y + 2 = 0$.

Soluzione. La retta per P_0 ortogonale a r ha equazione $x + 2y - 2 = 0$ (vedi esempio precedente). Dunque $H = r \cap r'$ è dato da $(-\frac{2}{5}, \frac{6}{5})$. \square

9.2 Distanza di un punto da una retta

Dato un punto P_0 e una retta r , la *distanza di P da r* è definita come il valore minimo che assume la distanza di P_0 da un punto di r : tale distanza si indica con il simbolo $d(P_0, r)$. Ragionando con un opportuno triangolo retto, si dimostra facilmente che la distanza di P_0 da r è data dalla distanza di P_0 da H , dove H è la proiezione ortogonale di P_0 su r :

$$d(P_0, r) = d(P_0, H).$$

Esempio Calcolare la distanza di $P_0 = (2, 0)$ dalla retta $2x - y + 2 = 0$.

Soluzione. La proiezione ortogonale, calcolata nell'esempio precedente, è $H = (-\frac{2}{5}, \frac{6}{5})$. Quindi

$$d(P_0, r) = d(P_0, H) = \frac{6}{\sqrt{5}}.$$

\square

Esiste una formula esplicita: se $P_0 = (x_0, y_0)$ e $r : ax + by + c = 0$ allora:

$$d(P_0, r) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}},$$

detta *formula della distanza di un punto da una retta*. Ne omettiamo la dimostrazione.

9.3 Punto medio e asse di un segmento

Dati due punti $A = (x_1, y_1)$ e $B = (x_2, y_2)$ il punto medio del segmento AB è il punto di coordinate

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right).$$

Si definisce *asse* del segmento AB l'insieme dei punti del piano equidistanti da A e B . Si osserva che l'asse del segmento AB è la retta r ortogonale al segmento AB , passante per il punto medio M di AB .

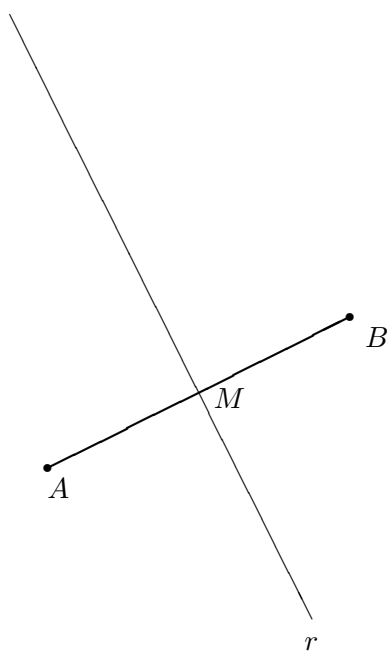


Figura 18: Asse del segmento AB

Si dirà anche che B è il *simmetrico di A rispetto alla retta r*.

Esempio Trovare l'equazione dell'asse del segmento AB , dove $A = (-1, 3)$, $B = (3, 5)$.

Soluzione. Il punto medio di AB ha coordinate $M = (1, 4)$. La retta generica perpendicolare al segmento AB ha equazione

$$r : 2x + y + k = 0,$$

dove $k \in \mathbf{R}$ (spiegare perché). Imponendo il passaggio per il punto M , otteniamo $k = 6$ e l'equazione dell'asse di AB è

$$r : 2x + y - 6 = 0.$$

□

9.4 Metodo del punto mobile

Esempio Nel piano sono dati la retta $r : 2x - y + 2 = 0$, il suo punto $A = (-1, 0)$ e il punto $Q = (1, 1)$.

- Determinare tutti i punti P sulla retta r tali che il triangolo di vertici A, P, Q abbia area 1.
- Determinare tutti i punti P' sulla retta r tali che il triangolo di vertici A, P', Q sia isoscele sulla base AP' .

Soluzione. Equazioni parametriche di r :

$$r : \begin{cases} x = t \\ y = 2t + 2 \end{cases}$$

Il punto generico (o punto *mobile*) su r ha dunque coordinate $P = (t, 2t + 2)$.

- L'area del triangolo di vertici A, P, Q vale $\frac{1}{2}d(A, P)d(Q, r)$. Ora:

$$d(A, P) = \sqrt{(t+1)^2 + (2t+2)^2} = \sqrt{5(t+1)^2},$$

mentre la distanza di Q da r vale:

$$d(Q, r) = \frac{3}{\sqrt{5}}.$$

Dunque dobbiamo avere

$$\frac{1}{2}\sqrt{5(t+1)^2} \cdot \frac{3}{\sqrt{5}} = 1,$$

e, risolvendo otteniamo due valori di t , dati da $t = -\frac{1}{3}$ e $t = -\frac{5}{3}$. I punti cercati sono dunque:

$$P_1 = \left(-\frac{1}{3}, \frac{4}{3}\right), \quad P_2 = \left(-\frac{5}{3}, -\frac{4}{3}\right).$$

- Dobbiamo imporre la seguente condizione al punto $P' = (t, 2t + 2)$:

$$d(P', Q) = d(A, Q) = \sqrt{5},$$

che si traduce nell'equazione:

$$5t^2 + 2t - 3 = 0.$$

Tale equazione ammette le soluzioni $t = -1, t = \frac{3}{5}$. Il primo valore dà luogo al punto $(-1, 0)$ (il punto A) che dobbiamo scartare. Il secondo valore dà luogo al punto

$$P' = (\frac{3}{5}, \frac{16}{5}),$$

che è l'unica soluzione del problema. \square

Parte 10. Geometria dello spazio I

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Lo spazio vettoriale V_O^3 , 1
- 2 Dipendenza e indipendenza lineare in V_O^3 , 2
- 3 Sistema di riferimento cartesiano, 5
- 4 Equazioni parametriche di una retta, 7
- 5 Equazione cartesiana di un piano, 11
- 6 Intersezione e parallelismo di due piani, 14
- 7 Equazioni cartesiane di una retta, 15
- 8 Parallelismo di una retta e un piano, 17

1 Lo spazio vettoriale V_O^3

1.1 Vettori dello spazio

Definizione Un vettore è una coppia ordinata (A, B) di punti dello spazio, che si denota con \overrightarrow{AB} .

A è detto il *punto di applicazione* e B è detto il *vertice* del vettore. Si estendono ai vettori dello spazio le definizioni già viste per i vettori del piano: direzione, verso e modulo.

Due vettori dello spazio si dicono *equipollenti* se hanno stessa direzione, stesso verso e stesso modulo.

Possiamo traslare vettori nel modo usuale:

- dati un vettore \overrightarrow{AB} e un punto A' , esiste un unico punto B' tale che $\overrightarrow{A'B'}$ è equipollente ad \overrightarrow{AB} . Il vettore $\overrightarrow{A'B'}$ si dice *traslato di \overrightarrow{AB} in A'* .

1.2 Lo spazio vettoriale V_O^3

Fissiamo un punto dello spazio O , detto *origine*, e consideriamo l'insieme dei vettori applicati in O . Tale insieme si denota con V_O^3 . Quindi

$$V_O^3 = \{\overrightarrow{OP} : P \text{ è un punto dello spazio}\}.$$

Esattamente come nel caso dei vettori del piano, possiamo definire:

- la somma di due vettori (con la regola del parallelogramma),
- il prodotto di un vettore per uno scalare.

Risulta allora che tali operazioni verificano gli assiomi di spazio vettoriale. In conclusione,

Proposizione V_O^3 , con le operazioni appena introdotte, è uno spazio vettoriale.

2 Dipendenza e indipendenza lineare in V_O^3

In questa sezione daremo un'interpretazione geometrica della dipendenza e indipendenza lineare di vettori di V_O^3 , e dimostreremo che V_O^3 ha dimensione 3. Richiamiamo in primo luogo alcuni fatti ben noti.

2.1 Alcuni fatti elementari

I concetti di *retta* e *piano* sono dati a priori.

- Diremo che i punti P_1, \dots, P_n sono *allineati* se appartengono ad una stessa retta.
- Diremo che i punti P_1, \dots, P_n sono *complanari* se appartengono ad uno stesso piano.

Abbiamo le seguenti proprietà.

- a) Per due punti distinti passa una e una sola retta.
- b) Per tre punti non allineati passa uno e un solo piano.

In particolare:

- c) due punti sono sempre allineati,
- d) tre punti sono sempre complanari.

Inoltre:

- e) per un punto dello spazio passano infinite rette,
- f) per due punti dello spazio passano infiniti piani.

Infine

- g) se un piano contiene due punti distinti, allora contiene l'intera retta per i due punti.

È chiaro che tre (o più) punti possono essere allineati oppure no, e quattro (o più) punti possono essere complanari oppure no.

2.2 Vettori allineati, vettori complanari

Analogamente al caso del piano, diremo che i vettori \overrightarrow{OA} e \overrightarrow{OB} sono *allineati* (o paralleli) se i punti O, A, B sono allineati.

Proposizione a) Due vettori \vec{v}, \vec{w} di V_O^3 sono linearmente dipendenti se e solo se sono allineati.

b) Se i vettori \vec{v}, \vec{w} non sono allineati, allora esiste un unico piano π contenente sia \vec{v} che \vec{w} .

Dimostrazione. a) è immediata dalla definizione di prodotto per uno scalare.

b) Se $\vec{v} = \overrightarrow{OA}$ e $\vec{w} = \overrightarrow{OB}$ non sono allineati allora i punti O, A, B non sono allineati : quindi esiste un unico piano π_0 passante per O, A, B . È evidente che π_0 contiene sia \vec{v} che \vec{w} . \square

Proposizione Supponiamo che π sia un piano dello spazio contenente l'origine, e consideriamo l'insieme di tutti i vettori applicati in O , con vertice in un punto di π :

$$E = \{\overrightarrow{OP} : P \in \pi\}.$$

Allora E è un sottospazio di V_O^3 di dimensione 2, che si identifica con V_O^2 .

Dimostrazione. La proposizione è più o meno ovvia: comunque, verifichiamo le proprietà di chiusura. È chiaro che il vettore nullo appartiene a E . Se $\vec{v} = \overrightarrow{OP}$ e $\vec{w} = \overrightarrow{OQ}$ appartengono a E allora per ipotesi $P, Q \in \pi$. Il vettore somma si scrive $\vec{v} + \vec{w} = \overrightarrow{OR}$ dove R è il vertice del parallelogramma sui lati OP, OQ . Poiché $O, P, Q \in \pi$, anche $R \in \pi$. Dunque $\vec{v} + \vec{w} \in E$ ed E è chiuso rispetto alla somma. La chiusura rispetto al prodotto per uno scalare è ovvia. Dunque E è un sottospazio. Da quanto detto è evidente che E si identifica con lo spazio vettoriale V_O^2 : quindi E ha dimensione 2. \square

Diremo che i vettori $\vec{u} = \overrightarrow{OA}, \vec{v} = \overrightarrow{OB}, \vec{w} = \overrightarrow{OC}$ sono *complanari* se i punti O, A, B, C sono complanari. In tal caso i vettori $\vec{u}, \vec{v}, \vec{w}$ sono tutti contenuti in uno stesso piano.

Teorema Tre vettori di V_O^3 sono linearmente dipendenti se e solo se sono complanari.

Dimostrazione. Supponiamo che $\vec{v}_1, \vec{v}_2, \vec{v}_3$ siano linearmente dipendenti. Allora uno di essi è combinazione lineare degli altri, e possiamo supporre che

$$\vec{v}_3 = a\vec{v}_1 + b\vec{v}_2.$$

Ora, se \vec{v}_1, \vec{v}_2 sono allineati, allora anche $\vec{v}_1, \vec{v}_2, \vec{v}_3$ sono allineati, e sono in particolare complanari. Se \vec{v}_1, \vec{v}_2 non sono allineati, allora esiste un unico piano π contenente entrambi

i vettori. Dunque $\vec{v}_1, \vec{v}_2 \in E$, dove $E = \{\overrightarrow{OP} : P \in \pi\}$. Poiché E è un sottospazio di V_O^3 , esso contiene tutte le combinazioni lineari di \vec{v}_1, \vec{v}_2 : quindi contiene anche v_3 , e $\vec{v}_1, \vec{v}_2, \vec{v}_3$ appartengono tutti al piano π .

Viceversa, supponiamo che $\vec{v}_1, \vec{v}_2, \vec{v}_3$ siano complanari, tutti contenuti in un piano π . Allora $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in E$, dove E è il sottospazio di V_O^3 formato dai vettori con vertice sul piano π . Per la proposizione, E ha dimensione 2 dunque $\vec{v}_1, \vec{v}_2, \vec{v}_3$ sono linearmente dipendenti. \square

Dimostreremo ora che V_O^3 ha dimensione 3. Osserviamo innanzitutto che nello spazio esiste sempre una terna di vettori $\vec{e}_1, \vec{e}_2, \vec{e}_3$, tutti di modulo unitario, e a due a due ortogonali (diremo allora che la terna $\vec{e}_1, \vec{e}_2, \vec{e}_3$ è *ortonormale*). Infatti, fissiamo un piano π per l'origine, e consideriamo una base ortonormale (\vec{e}_1, \vec{e}_2) di π . Prendiamo ora un vettore \vec{e}_3 di modulo unitario sulla retta per l'origine perpendicolare a π : è evidente che la terna $\vec{e}_1, \vec{e}_2, \vec{e}_3$ è ortonormale.

Proposizione a) *Lo spazio vettoriale V_O^3 ha dimensione 3.*

b) *Una terna di vettori di V_O^3 è una base se e solo se i vettori che la compongono non sono complanari.*

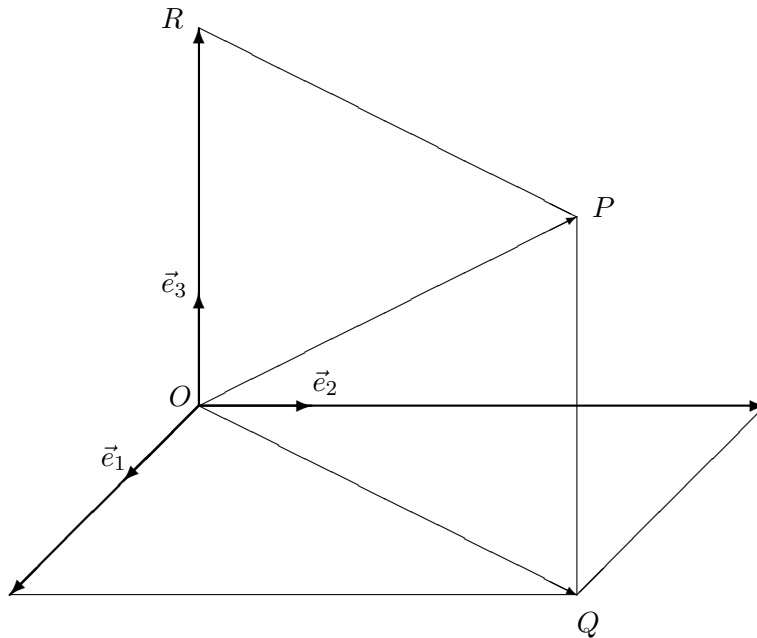
Dimostrazione. a) Per dimostrare che la dimensione di V_O^3 è tre basta trovare una base formata da tre vettori. Fissiamo una terna ortonormale $\vec{e}_1, \vec{e}_2, \vec{e}_3$. È chiaro che questi vettori non sono complanari, dunque sono linearmente indipendenti. Dimostriamo che $\vec{e}_1, \vec{e}_2, \vec{e}_3$ formano una base: per fare ciò, basta dimostrare che essi generano V_O^3 .

Dato un vettore $\vec{v} = \overrightarrow{OP}$, consideriamo il punto Q , piede della perpendicolare condotta da P al piano π contenente \vec{e}_1, \vec{e}_2 (vedi Figura 1). Se \overrightarrow{OR} è il traslato di \overrightarrow{QP} nell'origine, allora, per la regola del parallelogramma:

$$\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{OR},$$

inoltre \overrightarrow{OQ} e \overrightarrow{OR} sono ortogonali fra loro. Ora è chiaro che \overrightarrow{OQ} sta sul piano contenente \vec{e}_1, \vec{e}_2 , dunque è combinazione lineare \vec{e}_1, \vec{e}_2 , e \overrightarrow{OR} sta sulla retta contenente \vec{e}_3 , dunque è un multiplo di \vec{e}_3 . Di conseguenza, \overrightarrow{OP} sarà combinazione lineare di $\vec{e}_1, \vec{e}_2, \vec{e}_3$.

b) Dalle proprietà generali degli spazi vettoriali, e dalla parte a), sappiamo che tre vettori di V_O^3 formano una base se e solo se sono linearmente indipendenti, quindi, per il teorema, se e solo se non sono complanari. \square



$$\begin{aligned}\overrightarrow{OP} &= \overrightarrow{OQ} + \overrightarrow{OR} \\ &= x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3\end{aligned}$$

Figura 1: $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ è una base di V_O^3

3 Sistema di riferimento cartesiano

Un *sistema di riferimento cartesiano* nello spazio consiste nella scelta di un punto O , detto *origine*, e di una *base ortonormale* di V_O^3 . Dato un punto P , possiamo scrivere in modo unico

$$\overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3$$

e le coordinate del punto P saranno, per definizione, le coordinate di \overrightarrow{OP} . Scriveremo semplicemente

$$P = (x, y, z).$$

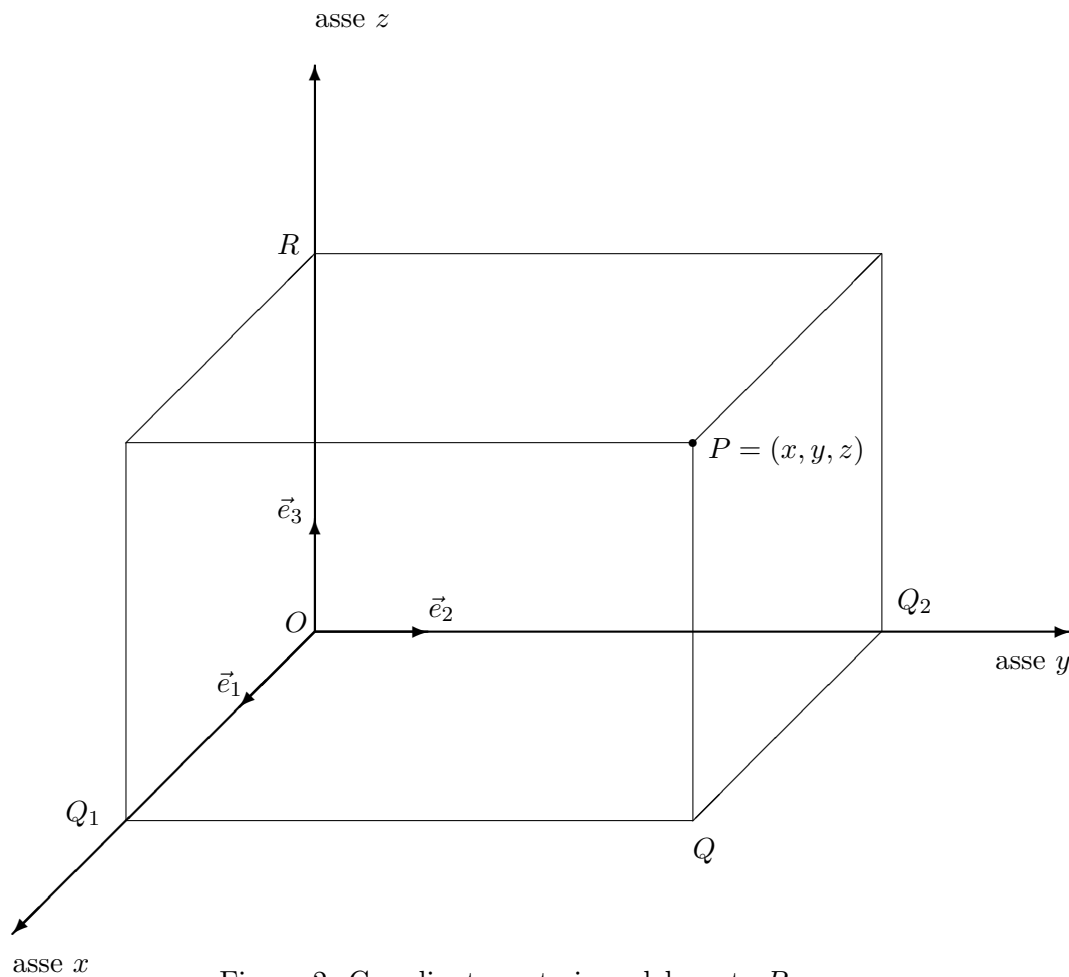


Figura 2: Coordinate cartesiane del punto P

Quindi ogni punto dello spazio si rappresenta con una terna di numeri. L'origine ha coordinate $(0, 0, 0)$. Dalla figura abbiamo che

$$\begin{aligned} x &= \text{ascissa di } P = d(Q_1, O) \\ y &= \text{ordinata di } P = d(Q_2, O) \\ z &= \text{quota di } P = d(R, O) \end{aligned}$$

con l'avvertenza che le distanze sono prese con il segno $+$ o $-$, a seconda che il punto Q_1, Q_2, R segua (rispettivamente, preceda) l'origine rispetto al verso dell'asse corrispondente. (Il punto P nella figura ha tutte le coordinate positive).

Abbiamo tre *piani coordinati* :

- il piano xy , descritto dall'equazione $z = 0$,
- il piano xz , descritto dall'equazione $y = 0$,
- il piano yz , descritto dall'equazione $x = 0$.

Ovviamente gli assi coordinati sono:

- l'asse x , descritto dalle equazioni $y = z = 0$,
- l'asse y , descritto dalle equazioni $x = z = 0$,
- l'asse z , descritto dalle equazioni $x = y = 0$.

Ad esempio, il punto $(2, 0, -1)$ appartiene al piano xz , mentre $(0, 3, 0)$ appartiene all'asse y . Vedremo poi che ogni piano dello spazio si rappresenta con un'equazione del tipo $ax + by + cz + d = 0$.

3.1 Coordinate di un vettore applicato in un punto qualunque

D'ora in poi supporremo di aver fissato nello spazio un sistema di riferimento cartesiano con origine O e base ortonormale $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Ogni vettore applicato nell'origine è quindi individuato dalla terna delle sue coordinate.

Come nel caso del piano, vogliamo ora attribuire coordinate ad un vettore applicato in un punto qualunque dello spazio.

- Dato il vettore $\vec{v} = \overrightarrow{AB}$ applicato nel punto A , le coordinate di \vec{v} sono poste per definizione uguali alle coordinate del vettore \vec{v}_0 , traslato di \vec{v} nell'origine.

Poiché $\vec{v}_0 = \overrightarrow{OB} - \overrightarrow{OA}$ le coordinate di \vec{v} sono date dalla differenza $B - A$. In altre parole

- Se $A = (x_1, y_1, z_1)$ e $B = (x_2, y_2, z_2)$ allora le coordinate del vettore \overrightarrow{AB} sono

$$(x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Dalla definizione è chiaro che

- due vettori sono equipollenti se e solo se hanno coordinate uguali,
- due vettori sono paralleli se e solo se hanno coordinate proporzionali.

Un vettore è identificato dal suo punto di applicazione e dalle sue coordinate. La scrittura

$$\overrightarrow{AB} = (l, m, n)$$

indica l'unico vettore di coordinate (l, m, n) applicato in A : esso unisce il punto di applicazione $A = (x_0, y_0, z_0)$ con il punto $B = (x_0 + l, y_0 + m, z_0 + n)$.

4 Equazioni parametriche di una retta

Vogliamo descrivere una retta con delle equazioni. Una retta dello spazio è determinata da

- un suo punto
- una direzione.

La direzione è specificata da un qualunque vettore parallelo alla retta, che chiameremo *vettore direttore* di r . Le coordinate di un vettore direttore sono dette *parametri direttori* di r .

Procedendo come nel caso del piano, otteniamo equazioni parametriche di una retta.

Proposizione *Una retta del piano si rappresenta con equazioni, dette parametriche, del tipo:*

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}$$

dove t è il parametro, (x_0, y_0, z_0) sono le coordinate di un punto della retta, e (l, m, n) sono i parametri direttori della retta.

Esempio La retta di equazioni parametriche $r : \begin{cases} x = 3t \\ y = 1 - t \\ z = 2 + 2t \end{cases}$ passa per il punto $P_0 = (0, 1, 2)$ e ha parametri direttori $(3, -1, 2)$, dunque è parallela al vettore $\vec{v} = 3\vec{e}_1 - \vec{e}_2 + 2\vec{e}_3$.
□

Rette parallele hanno vettori direttori paralleli; d'altra parte, vettori paralleli hanno coordinate proporzionali. Otteniamo immediatamente:

Proposizione *Due rette sono parallele se e solo se hanno parametri direttori proporzionali.*

Esempio Scrivere equazioni parametriche della retta r' passante per $(1, 2, -1)$ e parallela alla retta $r : \begin{cases} x = 3t \\ y = 1 - t \\ z = 2 + 2t \end{cases}$.

Soluzione. Basta prendere i parametri direttori di r' uguali a quelli di r , e imporre che per $t = 0$ la retta passi per $(1, 2, -1)$. Otteniamo le equazioni parametriche

$$r' : \begin{cases} x = 1 + 3t \\ y = 2 - t \\ z = -1 + 2t \end{cases}.$$

□

4.1 Retta per due punti

Siano $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ due punti distinti. Vogliamo scrivere equazioni parametriche della retta per P_1, P_2 . Ora il vettore $\overrightarrow{P_1 P_2}$ è parallelo alla retta, dunque i parametri direttori della retta cercata saranno proporzionali alle coordinate del vettore, cioè alla terna $P_2 - P_1$. Esplicitamente:

Proposizione *I parametri direttori della retta per $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ sono proporzionali alla terna:*

$$\begin{cases} l = x_2 - x_1 \\ m = y_2 - y_1 \\ n = z_2 - z_1 \end{cases}.$$

Esempio Scriviamo equazioni parametriche della retta passante per i punti $P_1 = (1, 2, 4)$ e $P_2 = (2, 1, 0)$. Possiamo prendere come parametri direttori $l = 1, m = -1, n = -4$; poiché r passa per $(1, 2, 4)$ otteniamo le equazioni

$$\begin{cases} x = 1 + t \\ y = 2 - t \\ z = 4 - 4t \end{cases}.$$

4.2 Condizione di allineamento di tre punti

Proposizione *I punti $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2), P_3 = (x_3, y_3, z_3)$ sono allineati se e solo se*

$$\text{rk} \begin{pmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{pmatrix} \leq 1.$$

Dimostrazione. Come nel caso del piano, basta osservare che i punti sono allineati se e solo se i vettori $\overrightarrow{P_1 P_2}, \overrightarrow{P_1 P_3}$, applicati in P_1 , sono allineati, dunque linearmente dipendenti. Prendendo le rispettive coordinate, si ha l'asserto. \square

Esempio Stabilire se i punti $P_1 = (0, 1, 1), P_2 = (2, 0, 2), P_3 = (4, -1, 3)$ sono allineati.

Soluzione. Si ha

$$\text{rk} \begin{pmatrix} 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix} = 1$$

dunque i tre punti sono allineati. Trovare le equazioni parametriche della retta che li contiene.

4.3 Intersezione di due rette

Illustriamo il problema con due esempi.

Esempio Stabilire se le rette $r : \begin{cases} x = 1 + 2t \\ y = 1 \\ z = 2 + 3t \end{cases}$ e $r' = \begin{cases} x = t \\ y = t \\ z = 3 - t \end{cases}$ si intersecano, e determinare le coordinate dell'eventuale punto d'intersezione.

Soluzione. Notiamo innanzitutto che i parametri che descrivono le due rette sono fra loro indipendenti, dunque per determinare l'intersezione dobbiamo adottare parametri diversi, diciamo t e s :

$$r : \begin{cases} x = 1 + 2t \\ y = 1 \\ z = 2 + 3t \end{cases}, \quad r' = \begin{cases} x = s \\ y = s \\ z = 3 - s \end{cases}.$$

A questo punto uguagliamo le due espressioni per ottenere:

$$\begin{cases} 1 + 2t = s \\ 1 = s \\ 2 + 3t = 3 - s \end{cases}$$

che ammette l'unica soluzione $s = 1, t = 0$. Dunque le rette si incontrano nel punto $(1, 1, 2)$ ottenuto per $t = 0$ dalle equazioni di r e per $s = 1$ da quelle di s . \square

Esempio Stabilire se le rette $r : \begin{cases} x = t \\ y = 1 \\ z = 0 \end{cases}$ e $r' = \begin{cases} x = 0 \\ y = t \\ z = 3 \end{cases}$ si intersecano, e determinare le coordinate dell'eventuale punto d'intersezione.

Soluzione. Cambiamo il nome dei parametri: $r : \begin{cases} x = t \\ y = 1 \\ z = 0 \end{cases}, r' : \begin{cases} x = 0 \\ y = s \\ z = 3 \end{cases}$. Uguagliando le coordinate otteniamo però un sistema incompatibile ($z = 0, z = 3$) dunque r e r' non si intersecano. \square

Osservazione Nel piano due rette distinte o sono parallele oppure si incontrano in un punto. Nello spazio questo non è più vero, come è dimostrato da quest'ultimo esempio: infatti, le rette r e r' sono ovviamente distinte, ma non sono né incidenti né parallele (i parametri direttori sono proporzionali, rispettivamente, alle terne $(1, 0, 0)$ e $(0, 1, 0)$).

In effetti, le due rette non possono essere contenute in uno stesso piano, sono cioè *sghembe*. Diremo che due rette dello spazio sono:

- *complanari*, se sono contenute in uno stesso piano,
- *sghembe*, se non sono complanari.

Esercizio Dimostrare che due rette incidenti sono contenute in un unico piano (dunque sono complanari).

Soluzione. Siano r, r' le due rette. Se le rette coincidono, l'asserzione è ovvia. Se non coincidono, le rette si incontrano in un unico punto P . Prendiamo ora un punto $A \neq P$ sulla retta r e un punto $B \neq P$ sulla retta r' . I punti A, B, P non sono allineati, dunque individuano un unico piano π . Ora π contiene due punti distinti di r (cioè P e A), dunque contiene tutta la retta r . Per un motivo analogo π contiene anche r' e si ha dunque la tesi. \square

D'altra parte, osserviamo che due rette dello spazio sono parallele se e solo se coincidono, oppure *sono complanari* e non hanno punti comuni.

In conclusione abbiamo la seguente

Proposizione *Due rette sono complanari se e solo se sono incidenti oppure sono parallele.*

Per contrapposizione:

Proposizione *Due rette sono sghembe se e solo se non sono nè incidenti nè parallele.*

5 Equazione cartesiana di un piano

5.1 Condizione di complanarità di quattro punti

Sappiamo che quattro punti del piano possono essere complanari oppure no. Dati $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2), P_3 = (x_3, y_3, z_3), P_4 = (x_4, y_4, z_4)$ essi sono complanari se e solo se i tre vettori (applicati nel punto P_1):

$$\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \overrightarrow{P_1P_4}$$

sono complanari, cioè linearmente dipendenti. Questo avverrà se e solo se le coordinate dei tre vettori, cioè le terne $P_2 - P_1, P_3 - P_1, P_4 - P_1$, sono vettori linearmente dipendenti di \mathbf{R}^3 . Dunque abbiamo:

Proposizione I punti $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2), P_3 = (x_3, y_3, z_3), P_4 = (x_4, y_4, z_4)$ sono complanari se e solo se

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \end{vmatrix} = 0$$

5.2 Equazione cartesiana di un piano

Proposizione a) Un piano π dello spazio si rappresenta con un'equazione del tipo:

$$ax + by + cz + d = 0, \quad \text{con} \quad (a, b, c) \neq (0, 0, 0),$$

detta equazione cartesiana di π .

b) L'equazione cartesiana del piano per i tre punti non allineati $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2), P_3 = (x_3, y_3, z_3)$ è data da:

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

Dimostrazione. Dimostriamo prima la parte b). Sia $P = (x, y, z)$ il punto generico dello spazio. Allora $P \in \pi$ se e solo se i quattro punti P_1, P_2, P_3, P sono complanari; dalla condizione di complanarità otteniamo (riordinando le righe) l'annullarsi del determinante in b). Ora per ipotesi si ha:

$$\text{rk} \begin{pmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{pmatrix} = 2,$$

poiché P_1, P_2, P_3 non sono allineati. Dunque almeno uno dei minori di ordine due della matrice è non nullo. Sviluppando il determinante lungo la prima riga, l'equazione diventa:

$$ax + by + cz + d = 0$$

con almeno uno fra a, b, c non nullo. Questo dimostra la parte a). \square

- Si può dimostrare anche il viceversa: le soluzioni di un'equazione del tipo $ax + by + cz + d = 0$, con a, b, c non tutti nulli, individuano un unico piano dello spazio.

Esempio Sono dati i punti $P_1 = (1, 2, 1), P_2 = (0, 1, 3), P_3 = (1, -1, 2)$. Verificare che i tre punti non sono allineati, e trovare l'equazione cartesiana dell'unico piano che li contiene.

Soluzione. Le coordinate di $\overrightarrow{P_1P_2}$ sono $(-1, -1, 2)$ mentre quelle di $\overrightarrow{P_1P_3}$ sono $(0, -3, 1)$.
Ora

$$\text{rk} \begin{pmatrix} -1 & -1 & 2 \\ 0 & -3 & 1 \end{pmatrix} = 2,$$

dunque i punti non sono allineati. L'equazione del piano è dunque:

$$\begin{vmatrix} x-1 & y-2 & z-1 \\ -1 & -1 & 2 \\ 0 & -3 & 1 \end{vmatrix} = 0,$$

che diventa $5x + y + 3z - 10 = 0$. \square

Esempio Abbiamo visto che le rette $r : \begin{cases} x = 1 + 2t \\ y = 1 \\ z = 2 + 3t \end{cases}$ e $r' = \begin{cases} x = t \\ y = t \\ z = 3 - t \end{cases}$ si intersecano nel punto $P_0 = (1, 1, 2)$: quindi sono complanari, contenute in un unico piano π .

Vogliamo determinare l'equazione del piano π .

Per fare ciò, è sufficiente trovare un punto $P \neq P_0$ sulla retta r , e un punto $Q \neq P_0$ sulla retta r' : il piano π sarà quello passante per P_0, P e Q . Il punto P si può ottenere ponendo $t = 1$ nelle equazioni parametriche di r :

$$P = (3, 1, 5).$$

Il punto Q si può ottenere ponendo ad esempio $t = 0$ nelle equazioni parametriche di r :

$$Q = (0, 0, 3).$$

L'equazione del piano π sarà dunque

$$\begin{vmatrix} x-1 & y-1 & z-2 \\ 2 & 0 & 3 \\ -1 & -1 & 1 \end{vmatrix} = 0,$$

ovvero

$$\pi : 3x - 5y - 2z + 6 = 0.$$

In effetti, si verifica che π contiene il punto generico di r , che ha coordinate $(1+2t, 1, 2+3t)$ con $t \in \mathbf{R}$, e contiene anche il punto generico della retta r' , che ha coordinate $(t, t, 3-t)$ con $t \in \mathbf{R}$. \square

5.3 Forme particolari

Abbiamo già osservato che i tre *piani coordinati* sono definiti dalle equazioni: $x = 0$ (piano yz), $y = 0$ (piano xz), $z = 0$ (piano xy).

Abbiamo immediatamente che

- se $d = 0$ il piano passa per l'origine.

Esempio Il piano $\pi : x - y + 2z = 0$ passa per l'origine.

Esempio L'equazione $2y - z = 0$ non contiene la variabile x , ed è soddisfatta da tutte le terne del tipo $(x, 0, 0)$: dunque il piano $\pi : 2y - z = 0$ contiene tutti i punti dell'asse x . Più in generale:

- se $a = d = 0$ il piano contiene l'asse x . Discutere i casi analoghi ($b = d = 0$ etc.)

6 Intersezione e parallelismo di due piani

I piani π e π' si dicono *paralleli* se coincidono oppure non hanno punti in comune.

Teorema Dati i piani $\pi : ax + by + cz + d = 0$ e $\pi' : a'x + b'y + c'z + d' = 0$, consideriamo la matrice: $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$. Allora

- I piani π, π' sono paralleli se e solo se $\text{rk}A = 1$.*
- I piani π, π' si incontrano in una retta se e solo se $\text{rk}A = 2$.*

Dimostrazione. I punti comuni a π, π' si ottengono risolvendo il sistema:

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

Sia $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$ la matrice dei coefficienti e A' la matrice completa. Se $\text{rk}A = 2$ allora anche $\text{rk}A' = 2$: il sistema è compatibile e ammette ∞^1 soluzioni. È allora evidente che in tal caso l'intersezione è una retta.

Supponiamo ora $\text{rk}A = 1$. Se $\text{rk}A' = 1$ allora il sistema è compatibile e ammette ∞^2 soluzioni: i piani sono coincidenti. Se invece $\text{rk}A' = 2$ allora il sistema è incompatibile, e i piani sono paralleli e distinti. \square

In conclusione, i due piani sono paralleli se e solo se i rispettivi coefficienti sono proporzionali (o uguali):

$$(a', b', c') = k(a, b, c)$$

per qualche $k \neq 0$.

Esempio I piani $\pi : x - y + 2z + 2 = 0$ e $\pi' : 2x - 2y + 4z + 1 = 0$ sono paralleli.

Notiamo che possiamo riscrivere $\pi' : x - y + 2z + \frac{1}{2} = 0$ e dunque π e π' differiscono solo per il termine noto. Questo è sempre vero:

- Le equazioni cartesiane di due piani paralleli possono ridursi a differire solo per il termine noto.

Esempio Il piano generico parallelo a $\pi : x - y + 2z + 2 = 0$ ha equazione $x - y + 2z + k = 0$, dove $k \in \mathbf{R}$, detta *equazione del fascio di piani paralleli a π* .

In generale, fissato un piano $\pi : ax + by + cz + d = 0$, il fascio di piani paralleli a π ha equazione:

$$ax + by + cz + k = 0,$$

dove $k \in \mathbf{R}$. Otteniamo così ∞^1 piani, tutti paralleli fra loro.

Esempio Determinare l'equazione cartesiana del piano passante per $(1, -1, 2)$ e parallelo al piano $\pi : x + 3y - z + 5 = 0$.

Soluzione. Scriviamo l'equazione del fascio di piani paralleli a π :

$$x + 3y - z + k = 0.$$

Imponiamo ora il passaggio per il punto $(1, -1, 2)$ e otteniamo $-4 + k = 0$ cioè $k = 4$. Dunque il piano cercato ha equazione $x + 3y - z + 4 = 0$. \square

7 Equazioni cartesiane di una retta

Abbiamo visto che due piani non paralleli si incontrano in una retta. Viceversa, una retta è sempre intersezione di due piani non paralleli (in infiniti modi). Abbiamo quindi la seguente

Proposizione Una retta si può rappresentare come intersezione di due piani non paralleli:

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

dove $\text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2$. Le equazioni di tale rappresentazione sono dette le equazioni cartesiane della retta r .

Dunque abbiamo due modi per rappresentare una retta:

- con equazioni parametriche,
- con equazioni cartesiane.

Per passare dalle equazioni parametriche alle equazioni cartesiane si elimina il parametro; mentre per passare dalle equazioni cartesiane alle equazioni parametriche si risolve il sistema.

Esempio È data la retta $r : \begin{cases} x - y - z + 2 = 0 \\ x + y + 3z = 0 \end{cases}$.

- Scrivere le equazioni parametriche di r e calcolare i suoi parametri direttori.
- Trovare equazioni parametriche ed equazioni cartesiane della retta r' parallela a r e passante per l'origine.

Soluzione. a) Si verifica che i piani che definiscono r non sono paralleli. Risolvendo il sistema otteniamo ∞^1 soluzioni:

$$\begin{cases} x = -1 - t \\ y = 1 - 2t \\ z = t \end{cases}$$

con parametro $t \in \mathbf{R}$, che danno le equazioni parametriche cercate. I parametri direttori di r sono proporzionali a $(l, m, n) = (-1, -2, 1)$ o anche a $(1, 2, -1)$.

- Le equazioni parametriche di r' sono date da $r' : \begin{cases} x = t \\ y = 2t \\ z = -t \end{cases}$. Eliminiamo il parametro

t per ottenere le equazioni cartesiane:

$$r' : \begin{cases} x + z = 0 \\ 2x - y = 0 \end{cases}.$$

7.1 Parametri direttori di una retta data con equazioni cartesiane

Sia r una retta descritta con equazioni cartesiane:

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

Proposizione *I parametri direttori di r sono proporzionali alla terna dei minori di ordine*

due (presi a segni alterni) della matrice dei coefficienti $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$, precisamente:

$$l = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, \quad m = -\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \quad n = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}. \quad (1)$$

Dimostrazione. Osserviamo che la retta r_0 , parallela a r e passante per l'origine, ha equazioni cartesiane:

$$r_0 : \begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0. \end{cases}$$

Se Q è un punto di r_0 diverso dall'origine, allora un vettore direttore di r sarà \overrightarrow{OQ} , e possiamo prendere come parametri direttori proprio le coordinate di Q . A questo punto basta osservare che in effetti la terna $Q = (l, m, n)$ definita in (1) è una soluzione non nulla del sistema che definisce r_0 . \square

Esempio Trovare le equazioni cartesiane della retta r' passante per $P_0 = (1, -1, 2)$ e parallela alla retta

$$r : \begin{cases} x - y + z = 0 \\ 3x + y + 5 = 0 \end{cases}.$$

Soluzione. I parametri direttori di r si ottengono dai minori della matrice

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

e sono proporzionali a $(-1, 3, 4)$. La retta cercata ha equazioni parametriche

$$r' : \begin{cases} x = 1 - t \\ y = -1 + 3t \\ z = 2 + 4t \end{cases}.$$

8 Parallelismo di una retta e un piano

Data una retta r e un piano π abbiamo tre possibilità:

- r e π si incontrano in un punto: diremo allora che sono *incidenti*.
- r e π non hanno intersezione.
- r è interamente contenuta in π .

Negli ultimi due casi, diremo che la retta r è *parallela* al piano π .

È chiaro che, se r è parallela a π e se π contiene un punto di r allora π contiene l'intera retta r .

Proposizione *Il piano $\pi : ax + by + cz + d = 0$ e la retta r di parametri direttori (l, m, n) sono paralleli se e solo se:*

$$al + bm + cn = 0.$$

• Nell'equazione di un piano $\pi : ax + by + cz + d = 0$ la terna (a, b, c) è detta la terna dei *parametri di giacitura* del piano. Dunque la proposizione può essere riformulata come segue.

Proposizione *Un piano di parametri di giacitura (a, b, c) e una retta di parametri direttori (l, m, n) sono paralleli se e solo se*

$$al + bm + cn = 0.$$

Dimostrazione. Sia r_0 la retta parallela a r passante per l'origine, e sia π_0 il piano parallelo a π passante per l'origine. Allora r è parallela a π se e solo se r_0 è contenuta in π_0 . Dalla definizione di parametri direttori, sappiamo che il punto (l, m, n) appartiene a r_0 ; d'altra parte, l'equazione del piano π_0 è data da $ax + by + cz = 0$. Dunque $r_0 \subseteq \pi_0$ se e solo se $(l, m, n) \in \pi$, cioè se e solo se $al + bm + cn = 0$. \square

Esempio Stabilire se la retta $r : \begin{cases} x = t \\ y = 2 + t \\ z = 1 \end{cases}$ e il piano $\pi : x - 3y + z = 0$ sono paralleli o incidenti.

Soluzione. I parametri direttori di r sono $(1, 1, 0)$ mentre i parametri di giacitura di π sono $(1, -3, 1)$. La condizione di parallelismo $al + bm + cn = 0$ non è verificata dunque retta e piano si incontrano in un punto. Per trovare il punto, basta sostituire le equazioni parametriche della retta nell'equazione del piano e si osserva che la retta incontra il piano per il valore $t = -\frac{5}{2}$. Dunque il punto d'intersezione ha coordinate $(-\frac{5}{2}, -\frac{1}{2}, 1)$. \square

8.1 Fascio di piani di asse una retta

Data una retta in equazioni cartesiane

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

il piano generico contenente r ha equazione:

$$\pi : h(ax + by + cz + d) + k(a'x + b'y + c'z + d') = 0,$$

con h, k parametri reali, non entrambi nulli. L'espressione è anche detta il *fascio di piani di asse r* .

Osservazione Data una retta r e un punto P non appartenente a r , esiste uno ed un solo piano contenente r e P .

Infatti, siano A e B due punti distinti di r . Siccome A, B e P non sono allineati, esiste uno ed un solo piano passante per A, B, P . Tale piano contiene sia r che P .

Esempio Determinare l'equazione cartesiana dell'unico piano passante per il punto $P = (1, 1, 1)$ e contenente la retta $r : \begin{cases} x + y - 1 = 0 \\ 3x + y - z = 0 \end{cases}$.

Soluzione. L'equazione del fascio di piani di asse r è: $h(x + y - 1) + k(3x + y - z) = 0$. Imponiamo il passaggio per $P = (1, 1, 1)$ e otteniamo la condizione:

$$h + 3k = 0.$$

Possiamo dunque prendere $k = 1$ e di conseguenza $h = -3$. Il piano cercato è dunque:

$$2y + z - 3 = 0.$$

Sembrerebbe che il problema ammetta più di una soluzione. In realtà non è così, poiché prendendo un'altra soluzione $h = -3k$ con $k \neq 0$ avremmo ottenuto il piano $2ky + kz - 3k = 0$ che coincide con il piano trovato in precedenza (basta dividere per k ambo i membri).

In effetti, potevamo scrivere il fascio di piani di asse r nella *forma ridotta* :

$$x + y - 1 + k(3x + y - z) = 0,$$

che ha il vantaggio di dipendere dal solo parametro k . L'unico problema è che nel fascio ridotto manca un piano, precisamente $3x + y - z = 0$.

Quindi si poteva procedere anche così: si cerca la soluzione fra i piani del fascio ridotto; se non la troviamo, significa che il piano cercato è quello che manca.

Infine, per risolvere il problema si poteva procedere anche nel modo seguente. Scegliamo due punti su r , ad esempio $A = (1, 0, 3), B = (0, 1, 1)$. Il piano cercato è l'unico contenente A, B, P , e ha equazione $2y + z - 3 = 0$.

Esempio Determinare l'equazione cartesiana del piano contenente la retta $r_1 : \begin{cases} x + y - 1 = 0 \\ 3x + y - z = 0 \end{cases}$ e parallelo alla retta $r_2 : \begin{cases} x - y + 3 = 0 \\ z + 4 = 0 \end{cases}$

Soluzione. Primo metodo. Scriviamo l'equazione del fascio ridotto di piani di asse r_1 , cioè $x + y - 1 + k(3x + y - z) = 0$. L'equazione si scrive anche così:

$$\pi : (1 + 3k)x + (1 + k)y - kz - 1 = 0.$$

Si trova facilmente che i parametri direttori di r_2 sono proporzionali a $(-1, -1, 0)$ ovvero a $(l, m, n) = (1, 1, 0)$. Dobbiamo ora imporre che il piano del fascio π sia parallelo a r_2 :

$$1 + 3k + 1 + k = 0,$$

da cui $k = -\frac{1}{2}$. Sostituendo, troviamo che il piano cercato è

$$x - y - z + 2 = 0.$$

Secondo metodo. Partiamo dall'equazione generica di un piano $ax + by + cz + d = 0$, e determiniamo i coefficienti a, b, c e il termine noto d .

1. Prendiamo due punti di r_1 , ad esempio $A = (1, 0, 3)$ e $B = (0, 1, 1)$.
2. Imponiamo che A appartenga al piano: $a + 3c + d = 0$
2. Imponiamo che B appartenga al piano: $b + c + d = 0$
3. Imponiamo che il piano sia parallelo a r_2 (di parametri direttori $(1, 1, 0)$): $a + b = 0$.

Dunque a, b, c, d sono soluzione del sistema lineare omogeneo:

$$\begin{cases} a + 3c + d = 0 \\ b + c + d = 0 \\ a + b = 0 \end{cases}$$

Il sistema ammette ∞^1 soluzioni, tutte proporzionali alla soluzione

$$\begin{cases} a = 1 \\ b = -1 \\ c = -1 \\ d = 2 \end{cases}$$

e il piano cercato è $x - y - z + 2 = 0$. \square

8.2 Stella di piani di centro un punto

L'insieme di tutti i piani passanti per un punto P_0 è detto *la stella di piani di centro P_0* . Si vede subito che un piano di tale insieme ha equazione del tipo

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

con $a, b, c \in \mathbf{R}$. In particolare, ci sono ∞^2 piani passanti per un punto dato (la terna (a, b, c) può essere alterata per un fattore di proporzionalità non nullo).

8.3 Piano parallelo a due direzioni

Supponiamo ora di fissare due rette dello spazio r, r' , e un punto P_0 .

Osservazione Se le rette r, r' non sono parallele, allora esiste un unico piano parallelo a entrambe le rette e passante per P_0 .

Infatti, siano r_0 e r'_0 le rette per l'origine parallele, rispettivamente, a r e r' . Allora r_0, r'_0 sono incidenti nell'origine, e definiscono un piano π_0 che le contiene entrambe. Ora il piano π parallelo a π_0 e passante per P_0 soddisfa chiaramente i requisiti.

Proposizione Siano date le rette r, r' non parallele, di parametri direttori $(l, m, n), (l', m', n')$, rispettivamente. Allora il piano per $P_0 = (x_0, y_0, z_0)$ parallelo a r e r' ha equazione

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ l & m & n \\ l' & m' & n' \end{vmatrix} = 0.$$

Dimostrazione. I vettori \vec{v}, \vec{w} di coordinate (l, m, n) e (l', m', n') , applicati in P_0 , sono

entrambi contenuti in π . Se $P = (x, y, z)$ è un punto di π anche il vettore $\overrightarrow{P_0P}$ è contenuto in π . I tre vettori $\vec{v}, \vec{w}, \overrightarrow{P_0P}$ sono dunque complanari, e di conseguenza linearmente dipendenti. Le coordinate di tali vettori dovranno essere linearmente dipendenti, e quindi

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ l & m & n \\ l' & m' & n' \end{vmatrix} = 0.$$

□

Esempio Trovare l'equazione del piano passante per $P_0 = (1, 0, 0)$ e parallelo a entrambe

$$\text{le rette } r : \begin{cases} x = 1 + t \\ y = -2 + t \\ z = 2t \end{cases} \text{ e } r' = \begin{cases} x - 2z = 0 \\ y = 0 \end{cases}.$$

Soluzione. I parametri direttori di r si ottengono immediatamente: $(l, m, n) = (1, 1, 2)$.
 Quelli di r' sono $(2, 0, 1)$ dunque l'equazione del piano cercato è:

$$\begin{vmatrix} x-1 & y & z \\ 1 & 1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 0,$$

cioè $x + 3y - 2z - 1 = 0$.

Metodo alternativo. Partiamo dal piano generico $ax + by + cz + d = 0$.

1. Imponiamo il passaggio per $P_0 : a + d = 0$.
2. Imponiamo il parallelismo alla retta $r : a + b + 2c = 0$.
3. Imponiamo il parallelismo alla retta $r' : 2a + c = 0$.

Il piano si ottiene risolvendo il sistema $\begin{cases} a + d = 0 \\ a + b + 2c = 0 \\ 2a + c = 0 \end{cases}$ che ha ∞^1 soluzioni $a = -t, b =$

$-3t, c = 2t, d = t$, con $t \in \mathbf{R}$, tutte proporzionali alla soluzione $a = 1, b = 3, c = -2, d = -1$ dunque il piano cercato è

$$x + 3y - 2z - 1 = 0.$$

□

Parte 11. Geometria dello spazio II

A. Savo – Appunti del Corso di Geometria 2010-11

INDICE DELLE SEZIONI

- 1 Il prodotto scalare, 1
- 2 Distanze, angoli, aree, 4
- 3 Il prodotto vettoriale, 6
- 4 Condizioni di perpendicolarità, 8
- 5 Distanza di un punto da un piano, 13
- 6 Distanza di un punto da una retta, 14
- 7 Distanza di due rette, 16
- 8 Circonferenze, 18
- 9 Sfere, 24

1 Il prodotto scalare

La definizione di prodotto scalare di due vettori dello spazio è identica a quella già vista nel piano.

Definizione Siano \vec{v}, \vec{w} vettori non nulli dello spazio applicati in uno stesso punto. Definiamo prodotto scalare di \vec{v} e \vec{w} il numero

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

dove θ è l'angolo convesso (cioè tale che $0 \leq \theta \leq \pi$) definito da \vec{v} e \vec{w} .

Se uno dei due vettori è nullo poniamo semplicemente $\langle \vec{v}, \vec{w} \rangle = 0$.

Abbiamo immediatamente:

Proposizione a) Due vettori sono ortogonali se e solo se il loro prodotto scalare è nullo.

b) Si ha $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$.

Per definizione, il vettore nullo è ortogonale a tutti i vettori.

Dunque il prodotto scalare permette di misurare il modulo (lunghezza) di un vettore, e l'angolo fra due vettori.

Le proprietà del prodotto scalare nel piano si estendono inalterate al prodotto scalare nello spazio. In particolare, abbiamo le proprietà di bilinearità:

$$\langle a\vec{v}_1 + b\vec{v}_2, \vec{w} \rangle = a\langle \vec{v}_1, \vec{w} \rangle + b\langle \vec{v}_2, \vec{w} \rangle,$$

valide per ogni scelta dei vettori $\vec{v}_1, \vec{v}_2, \vec{w}$ e degli scalari $a, b \in \mathbf{R}$.

1.1 Calcolo in un riferimento cartesiano

Ricordiamo che assegnare un riferimento cartesiano nello spazio equivale a fissare un'origine O e una base ortonormale di V_O^3 , diciamo $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Una base ortonormale è formata da vettori a due a due ortogonali e di modulo 1; dunque dalla definizione di prodotto scalare abbiamo che

$$\langle \vec{e}_i, \vec{e}_j \rangle = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}. \quad (1)$$

Ogni vettore si esprime, in modo unico:

$$\vec{v} = \overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3,$$

e la coppia (x, y, z) dà luogo alle coordinate di \vec{v} nel riferimento scelto. Spesso scriveremo semplicemente

$$\vec{v} = (x, y, z)$$

identificando un vettore applicato nell'origine con le sue coordinate.

Conoscendo le coordinate dei vettori, possiamo calcolare il prodotto scalare con facilità.

Proposizione Siano $\vec{v} = (x, y, z)$ e $\vec{w} = (x', y', z')$ vettori applicati nell'origine. Allora

$$\langle \vec{v}, \vec{w} \rangle = xx' + yy' + zz'.$$

In particolare, $\|\vec{v}\| = \sqrt{x^2 + y^2 + z^2}$.

Dimostrazione. Basta scrivere $\vec{v} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3$, $\vec{w} = x'\vec{e}_1 + y'\vec{e}_2 + z'\vec{e}_3$, usare la bilinearità del prodotto scalare e le relazioni (1). \square

Esempio Calcolare il prodotto scalare dei vettori $\vec{v} = 3\vec{e}_1 - \vec{e}_2 + 4\vec{e}_3$ e $\vec{w} = -\vec{e}_1 + 5\vec{e}_2$.

Soluzione. Si ha $\vec{v} = (3, -1, 4)$, $\vec{w} = (-1, 5, 0)$ quindi $\langle \vec{v}, \vec{w} \rangle = -8$. \square

Vogliamo ora derivare una formula analoga per calcolare il prodotto scalare di due vettori applicati in un punto qualunque dello spazio.

Ricordiamo che, dato un vettore $\vec{v} = \overrightarrow{AB}$ applicato nel punto A , le coordinate di \vec{v} sono per definizione le coordinate del vettore \vec{v}_0 , traslato di \vec{v} nell'origine, e uguagliano la differenza $B - A$.

Poichè la traslazione di due vettori mantiene inalterati sia i moduli che l'angolo compreso, il prodotto scalare fra \vec{v} e \vec{w} è uguale al prodotto scalare dei rispettivi vettori traslati \vec{v}_0 e \vec{w}_0 :

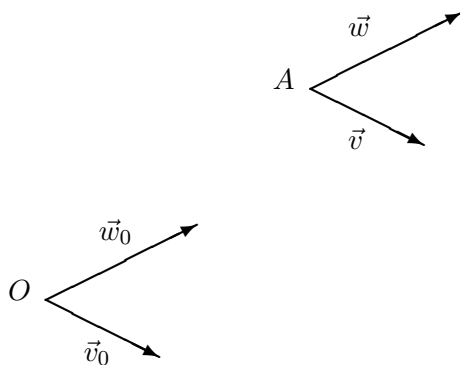


Figura 1: $\langle \vec{v}, \vec{w} \rangle = \langle \vec{v}_0, \vec{w}_0 \rangle$

Di conseguenza abbiamo la seguente

Proposizione Siano \vec{v} e \vec{w} vettori dello spazio applicati in un qualunque punto A dello spazio, di coordinate rispettive (x, y, z) e (x', y', z') . Allora:

$$\langle \vec{v}, \vec{w} \rangle = xx' + yy' + zz'.$$

In particolare, $\|\vec{v}\| = \sqrt{x^2 + y^2 + z^2}$.

Esempio Dati i punti $A = (1, 2, -1)$, $B = (3, 0, 2)$, $C = (2, 0, -3)$ calcolare il prodotto scalare dei vettori \overrightarrow{AB} , \overrightarrow{AC} , entrambi applicati in A .

Soluzione. Le coordinate di \overrightarrow{AB} sono $(2, -2, 3)$ mentre quelle di \overrightarrow{AC} sono $(1, -2, -2)$. Dunque:

$$\langle \overrightarrow{AB}, \overrightarrow{AC} \rangle = 2 + 4 - 6 = 0.$$

I vettori sono perpendicolari, e il triangolo ABC è rettangolo in A . \square

2 Distanze, angoli, aree

2.1 Distanza di due punti

Se $A = (x_1, y_1, z_1)$ e $B = (x_2, y_2, z_2)$ sono due punti dello spazio, allora la distanza di A da B uguaglia il modulo del vettore \overrightarrow{AB} , che ha coordinate

$$B - A = (x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Otteniamo dunque la *formula della distanza* :

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

2.2 Angoli

Dalla formula che definisce il prodotto scalare, vediamo subito che il coseno dell'angolo compreso fra i vettori non nulli \vec{v} e \vec{w} è dato da:

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

2.3 Area di un parallelogramma

Proposizione *Dati due vettori non nulli \vec{v}, \vec{w} , l'area \mathcal{A} del parallelogramma definito dai due vettori è data dalla formula:*

$$\mathcal{A} = \sqrt{\begin{vmatrix} \langle \vec{v}, \vec{v} \rangle & \langle \vec{v}, \vec{w} \rangle \\ \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{vmatrix}},$$

ovvero dalla formula equivalente

$$\mathcal{A}^2 = \|\vec{v}\|^2 \|\vec{w}\|^2 - \langle \vec{v}, \vec{w} \rangle^2.$$

Dimostrazione. Sia θ l'angolo formato da \vec{v} e \vec{w} . Dalla geometria elementare, il parallelogramma definito dai due vettori ha area

$$\mathcal{A} = \|\vec{v}\| \|\vec{w}\| \sin \theta.$$

Elevando ambi i membri al quadrato e osservando che $\sin^2 \theta = 1 - \cos^2 \theta$, otteniamo:

$$\begin{aligned} \mathcal{A}^2 &= \|\vec{v}\|^2 \|\vec{w}\|^2 - \|\vec{v}\|^2 \|\vec{w}\|^2 \cos^2 \theta \\ &= \|\vec{v}\|^2 \|\vec{w}\|^2 - \langle \vec{v}, \vec{w} \rangle^2. \end{aligned}$$

□

Esempio È dato il triangolo T di vertici $A = (1, 0, 0)$, $B = (0, 2, 0)$, $C = (0, 0, 3)$. Calcolare:

- a) Il perimetro di T .
- b) Il coseno di ciascuno degli angoli interni di T .
- c) L'area di T .

Soluzione. a) Abbiamo $\overrightarrow{AB} = (-1, 2, 0)$, $\overrightarrow{AC} = (-1, 0, 3)$, $\overrightarrow{BC} = (0, -2, 3)$. Dunque il perimetro vale

$$\|\overrightarrow{AB}\| + \|\overrightarrow{AC}\| + \|\overrightarrow{BC}\| = \sqrt{5} + \sqrt{10} + \sqrt{13}.$$

b) Sia θ_A l'angolo in A . Si ha

$$\cos \theta_A = \frac{\langle \overrightarrow{AB}, \overrightarrow{AC} \rangle}{\|\overrightarrow{AB}\| \|\overrightarrow{AC}\|} = \frac{1}{\sqrt{50}}.$$

Analogamente otteniamo

$$\cos \theta_B = \frac{4}{\sqrt{65}}, \quad \cos \theta_C = \frac{9}{\sqrt{130}}.$$

c) Consideriamo i vettori $\vec{v} = \overrightarrow{AB} = (-1, 2, 0)$ e $\vec{w} = \overrightarrow{AC} = (-1, 0, 3)$. L'area del parallelogramma su \vec{v} e \vec{w} vale:

$$\mathcal{A} = \sqrt{\begin{vmatrix} 5 & 1 \\ 1 & 10 \end{vmatrix}} = 7,$$

dunque l'area del triangolo T (che è la metà dell'area del parallelogramma) vale $\frac{7}{2}$.

Esempio Sono dati i vettori $\vec{v}_1 = (1, 1, -1)$ e $\vec{v}_2 = (2, 0, 1)$, applicati nell'origine.

- a) Descrivere l'insieme dei vettori \vec{w} ortogonali sia a \vec{v}_1 che a \vec{v}_2 .
- b) Trovare tutti i vettori \vec{u} ortogonali sia a \vec{v}_1 che a \vec{v}_2 e aventi modulo 1.

Soluzione. a) Poniamo $\vec{w} = (x, y, z)$. Allora \vec{w} è ortogonale a \vec{v}_1 se e solo se $\langle \vec{w}, \vec{v}_1 \rangle = 0$ dunque se e solo se $x + y - z = 0$. Analogamente, \vec{w} è ortogonale a \vec{v}_2 se e solo se $\langle \vec{w}, \vec{v}_2 \rangle = 0$, cioè se e solo se $2x + z = 0$. In conclusione, le coordinate di \vec{w} sono le soluzioni del sistema lineare omogeneo

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

La matrice dei coefficienti $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix}$ ha rango 2; dunque il sistema ammette ∞^1 soluzioni e un calcolo mostra che

$$\text{Sol}(S) = \{t(1, -3, -2) : t \in \mathbf{R}\}.$$

Dunque i vettori \vec{w} ortogonali sia a \vec{v}_1 che a \vec{v}_2 sono tutti e soli i vettori di coordinate:

$$\vec{w} = (t, -3t, -2t),$$

dove $t \in \mathbf{R}$. Notiamo che il vettore \vec{w} sta sulla retta per l'origine ortogonale al piano generato dai vettori \vec{v}_1, \vec{v}_2 .

b) Imponiamo ora che il vettore \vec{w} abbia modulo 1. Si ha:

$$\|\vec{w}\|^2 = 14t^2,$$

e $14t^2 = 1$ se $t = \frac{1}{\sqrt{14}}, t = -\frac{1}{\sqrt{14}}$. Otteniamo due vettori:

$$\vec{u}_1 = \frac{1}{\sqrt{14}}(1, -3, -2), \quad \vec{u}_2 = -\frac{1}{\sqrt{14}}(1, -3, -2),$$

l'uno opposto dell'altro. \square

3 Il prodotto vettoriale

Il prodotto vettoriale è un'operazione che associa a due vettori \vec{v}, \vec{w} applicati in un punto A un terzo vettore applicato in A e denotato $\vec{v} \wedge \vec{w}$. È sufficiente definire le coordinate di $\vec{v} \wedge \vec{w}$.

Definizione Il prodotto vettoriale di $\vec{v} = (a, b, c)$ e $\vec{w} = (a', b', c')$ è il vettore

$$\vec{v} \wedge \vec{w} = (\alpha, \beta, \gamma),$$

dove

$$\alpha = \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, \quad \beta = -\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \quad \gamma = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$$

sono i minori di ordine due della matrice $\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$, presi a segni alterni: $+-+$.

Esempio Dati i vettori $\vec{v} = (1, 1, -1), \vec{w} = (2, 0, 1)$ si ha

$$\vec{v} \wedge \vec{w} = (1, -3, -2).$$

Osserviamo che $\vec{v} \wedge \vec{w}$ è ortogonale sia a \vec{v} che a \vec{w} . \square

Il prodotto vettoriale ha proprietà notevoli.

Proposizione *Dati due vettori dello spazio \vec{v}, \vec{w} , si ha:*

- a) \vec{v}, \vec{w} sono allineati (paralleli) se e solo se $\vec{v} \wedge \vec{w} = \vec{0}$.
- b) $\vec{v} \wedge \vec{w}$ è un vettore ortogonale sia a \vec{v} che a \vec{w} .
- c) Il modulo del prodotto vettoriale $\vec{v} \wedge \vec{w}$ uguaglia l'area del parallelogramma definito da \vec{v} e \vec{w} .

Dimostrazione. a) Supponiamo che $\vec{v} = (a, b, c), \vec{w} = (a', b', c')$. I vettori \vec{v}, \vec{w} sono allineati se e solo se $\text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 1$ e questo è vero se e solo se i tre minori di ordine due hanno determinante nullo, cioè $(\alpha, \beta, \gamma) = (0, 0, 0)$.

b) Si ha

$$\begin{aligned} \langle \vec{v} \wedge \vec{w}, \vec{v} \rangle &= a\alpha + b\beta + c\gamma \\ &= a \begin{vmatrix} b & c \\ \beta & \gamma \end{vmatrix} - b \begin{vmatrix} a & c \\ \alpha & \gamma \end{vmatrix} + c \begin{vmatrix} a & b \\ \alpha & \beta \end{vmatrix} \\ &= \begin{vmatrix} a & b & c \\ a & b & c \\ \alpha & \beta & \gamma \end{vmatrix} \\ &= 0 \end{aligned}$$

perchè il determinante ha due righe uguali. Quindi $\vec{v} \wedge \vec{w}$ è ortogonale a \vec{v} . Con un procedimento analogo si dimostra che $\langle \vec{v} \wedge \vec{w}, \vec{w} \rangle = 0$, e $\vec{v} \wedge \vec{w}$ è ortogonale anche a \vec{w} .

c) Si ha $\|\vec{v} \wedge \vec{w}\|^2 = \alpha^2 + \beta^2 + \gamma^2$. Un calcolo esplicito mostra che

$$\alpha^2 + \beta^2 + \gamma^2 = \begin{vmatrix} \langle \vec{v}, \vec{v} \rangle & \langle \vec{v}, \vec{w} \rangle \\ \langle \vec{v}, \vec{w} \rangle & \langle \vec{w}, \vec{w} \rangle \end{vmatrix}.$$

L'espressione a secondo membro è il quadrato dell'area \mathcal{A} del parallelogramma definito da \vec{v} e \vec{w} , dunque $\|\vec{v} \wedge \vec{w}\| = \mathcal{A}$. \square

Osserviamo che il prodotto vettoriale è *anticommutativo* : $\vec{w} \wedge \vec{v} = -\vec{v} \wedge \vec{w}$.

Esempio Calcolare l'area del triangolo di vertici $A = (1, 0, 0), B = (0, 2, 0), C = (0, 0, 3)$.

Soluzione. Consideriamo $\vec{v} = \overrightarrow{AB} = (-1, 2, 0)$ e $\vec{w} = \overrightarrow{AC} = (-1, 0, 3)$. L'area del triangolo sarà la metà dell'area del parallelogramma \mathcal{P} definito da \vec{v} e \vec{w} . Ora

$$\vec{v} \wedge \vec{w} = (6, 3, 2);$$

l'area di \mathcal{P} è $\|\vec{v} \wedge \vec{w}\| = 7$ e quindi l'area del triangolo è $\frac{7}{2}$. \square

4 Condizioni di perpendicolarità

4.1 Perpendicolarità di due rette

Due rette si dicono ortogonali (o perpendicolari) se hanno direzioni ortogonali. Poiché la direzione di una retta è definita da un suo vettore direttore, che ha coordinate date dai parametri direttori (l, m, n) , abbiamo che:

Proposizione Due rette di parametri direttori $(l, m, n), (l', m', n')$ sono ortogonali se e solo se

$$ll' + mm' + nn' = 0.$$

Osservazione Occorre notare che due rette possono essere ortogonali fra loro senza necessariamente incontrarsi in un punto; in particolare, esistono rette sghembe ortogonali tra loro.

Esempio L'asse x , di equazioni $\begin{cases} y = 0 \\ z = 0 \end{cases}$, e la retta $\begin{cases} x = 0 \\ z = 1 \end{cases}$ sono sghembe e ortogonali.

- Diremo che le rette r, r' sono *ortogonali e incidenti*, se sono ortogonali e si incontrano in un punto.

Esempio Determinare equazioni parametriche della retta s passante per l'origine e ortogonale a entrambe le rette r :

$$r : \begin{cases} x = 1 + 2t \\ y = t \\ z = -3t \end{cases} \quad \text{e } r' : \begin{cases} x - y = 0 \\ y - z = 0 \end{cases}.$$

Soluzione. Un vettore direttore di r è $\vec{v} = (2, 1, -3)$ mentre un vettore direttore di r' è $\vec{w} = (1, 1, 1)$. La retta s avrà vettore direttore ortogonale sia a \vec{v} che a \vec{w} . Dunque possiamo prendere come vettore direttore della retta s il prodotto vettoriale $\vec{v} \wedge \vec{w} = (4, -5, 1)$.

Otteniamo le equazioni parametriche: $s : \begin{cases} x = 4t \\ y = -5t \\ z = t \end{cases} . \square$

In generale, abbiamo la seguente

Proposizione Dato un punto P_0 e due rette non parallele r, r' , esiste un'unica retta s passante per P_0 e ortogonale sia a r che a r' .

In particolare, due rette non parallele definiscono un'unica direzione ortogonale a entrambe, data dal prodotto vettoriale dei rispettivi parametri direttori.

4.2 Vettore normale a un piano

Fissato un piano π e un qualunque punto $P_0 \in \pi$, diremo che il vettore \vec{n} , applicato in P_0 , è *ortogonale* (o *normale*) a π se è ortogonale a tutti i vettori applicati in P_0 e contenuti in π :

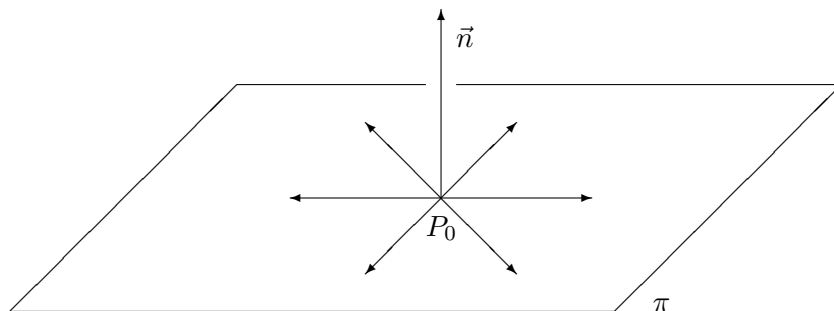


Figura 2: Vettore normale al piano π

È evidente che un piano ammette infiniti vettori normali, tutti allineati tra loro, e appartenenti alla retta per P_0 perpendicolare al piano.

La proposizione che segue fornisce un'interpretazione geometrica dei parametri di giacitura di un piano.

Proposizione Dato il piano $\pi : ax + by + cz + d = 0$, il vettore \vec{n} di coordinate (a, b, c) , applicato in un punto qualunque di π , è ortogonale a π . In altre parole, i parametri di giacitura di un piano rappresentano le coordinate di un vettore normale al piano.

Dimostrazione. Fissiamo un punto $P_0 = (x_0, y_0, z_0)$ del piano. Allora

$$ax_0 + by_0 + cz_0 + d = 0.$$

Sottraendo membro a membro dall'equazione di π osserviamo che $P = (x, y, z)$ appartiene a π se e solo se:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

Consideriamo il vettore $\vec{n} = (a, b, c)$, applicato in P_0 , e un vettore arbitrario contenuto nel piano, diciamo $\overrightarrow{P_0P}$. Se $P = (x, y, z)$ allora $\overrightarrow{P_0P}$ ha coordinate $(x - x_0, y - y_0, z - z_0)$. Il prodotto scalare:

$$\langle \vec{n}, \overrightarrow{P_0P} \rangle = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

dunque \vec{n} è ortogonale a $\overrightarrow{P_0P}$. Poiché $P \in \pi$ è arbitrario, \vec{n} è ortogonale a tutti i vettori di π . \square

4.3 Perpendicolarità di una retta e un piano

Con l'interpretazione geometrica dei parametri di giacitura data in precedenza abbiamo la seguente

Proposizione *La retta di parametri direttori (l, m, n) è ortogonale al piano di parametri di giacitura (a, b, c) se e solo se*

$$\text{rk} \begin{pmatrix} a & b & c \\ l & m & n \end{pmatrix} = 1,$$

vale a dire, se e solo se i parametri direttori della retta sono proporzionali ai parametri di giacitura del piano.

Dimostrazione. Il vettore $\vec{n} = (a, b, c)$ è ortogonale al piano, dunque è parallelo alla retta. Ne segue che la terna (a, b, c) è proporzionale alla terna dei parametri direttori (l, m, n) .
 \square

Esempio Determinare equazioni parametriche della retta passante per $P_0 = (1, 2, 3)$ e ortogonale al piano $x - y + 2z = 0$.

Soluzione. Per imporre l'ortogonalità, è sufficiente prendere i parametri direttori della retta uguali ai parametri di giacitura del piano: $(l, m, n) = (1, -1, 2)$. Dunque la retta cercata ha equazioni parametriche

$$\begin{cases} x = 1 + t \\ y = 2 - t \\ z = 3 + 2t \end{cases}.$$

\square

Esempio Trovare l'equazione cartesiana del piano π passante per $P_0 = (1, -2, 0)$ e perpendicolare alla retta di equazioni cartesiane

$$r : \begin{cases} x + y - 2z = 0 \\ 2x + 3z = 0 \end{cases}$$

Soluzione. Un calcolo mostra che i parametri direttori di r sono proporzionali alla terna $(3, -7, -2)$. Allora il piano generico perpendicolare a r avrà equazione

$$3x - 7y - 2z + k = 0,$$

con $k \in \mathbf{R}$. Imponendo il passaggio per il punto P_0 otteniamo $k = -17$. Dunque il piano cercato è unico, di equazione

$$3x - 7y - 2z - 17 = 0.$$

□

4.4 Perpendicolarità di due piani

Siano π, π' due piani non paralleli.

Diremo che π è *ortogonale* a π' se, dato un punto qualunque $P_0 \in \pi \cap \pi'$, la retta r passante per P_0 e ortogonale a π è interamente contenuta in π' :

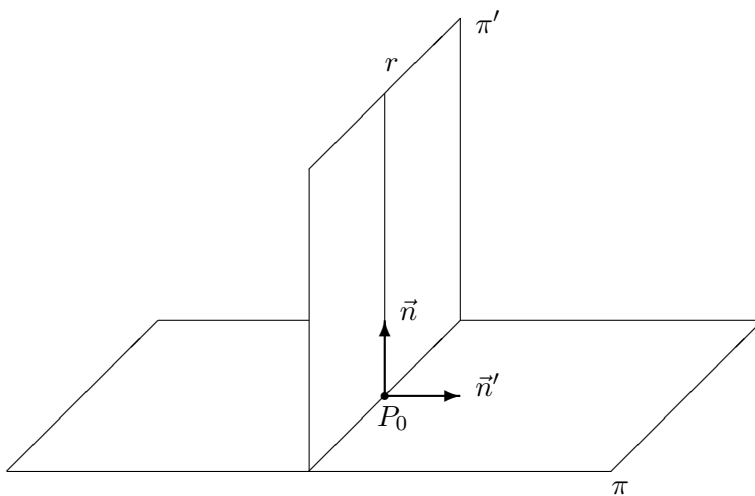


Figura 3: Piani ortogonali

Se \vec{n}, \vec{n}' sono vettori normali a π, π' , rispettivamente, allora risulta che $\pi \perp \pi'$ se e solo se $\vec{n} \perp \vec{n}'$. Ricordando che i parametri di giacitura di un piano sono le coordinate di un vettore normale al piano, abbiamo la seguente condizione di perpendicolarità di due piani.

Proposizione Il piano $\pi : ax + by + cz + d = 0$ è ortogonale al piano $\pi' : a'x + b'y + c'z + d' = 0$ se e solo se:

$$aa' + bb' + cc' = 0.$$

Osserviamo che, dato un piano π e un punto P_0 dello spazio, esistono infiniti piani passanti per P_0 e ortogonali al piano π . Tutti questi piani contengono la retta che passa per P_0 ed è ortogonale a π .

D'altra parte, si ha che:

- dato un piano π e una retta r esiste sempre un piano π' perpendicolare a π e contenente r . Se la retta r non è ortogonale a π , tale piano è unico.

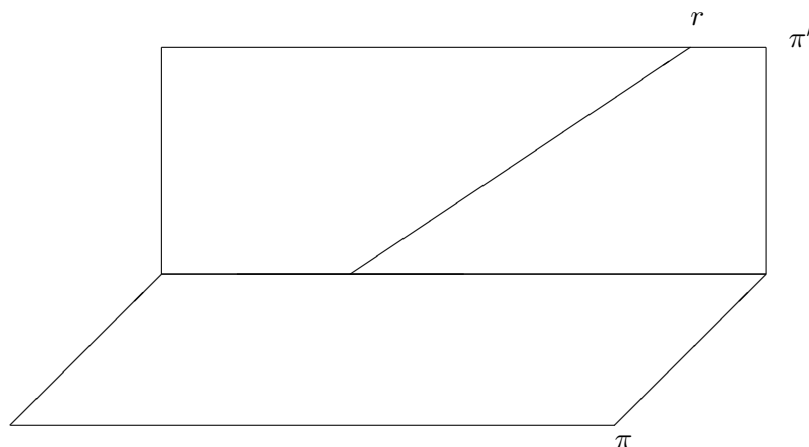


Figura 4: Piano ortogonale a π contenente r

Esempio Sia $\pi : 3x - y - 2z = 0$ e $r : \begin{cases} x - y = 0 \\ x + 2z = 0 \end{cases}$. Determinare il piano π' contenente r e perpendicolare a π .

Soluzione. Primo metodo. Il fascio (ridotto) di piani di asse r ha equazione cartesiana $x - y + k(x + 2z) = 0$ ovvero:

$$(k + 1)x - y + 2kz = 0.$$

Imponiamo la perpendicolarità al piano π e otteniamo:

$$3(k + 1) + 1 - 4k = 0,$$

ovvero $k = 4$. Il piano cercato è unico, di equazione $\pi' : 5x - y + 8z = 0$.

Secondo metodo. Consideriamo due punti sulla retta, ad esempio $A = (0, 0, 0)$, $B = (2, 2, -1)$. Il piano cercato dovrà passare per A e B e dovrà essere ortogonale a π . Partiamo dall'equazione del piano generico $\pi' : ax + by + cz + d = 0$. Imponendo le tre condizioni

otteniamo, nell'ordine, le equazioni:

$$\begin{cases} d = 0 \\ 2a + 2b - c + d = 0 \\ 3a - b - 2c = 0 \end{cases}$$

Risolvendo, otteniamo ∞^1 soluzioni, tutte proporzionali alla soluzione $a = 5, b = -1, c = 8, d = 0$, che dà luogo al piano $\pi' : 5x - y + 8z = 0$. \square

5 Distanza di un punto da un piano

Dato un piano π e un punto P_0 dello spazio, consideriamo il piede della perpendicolare condotta da P_0 al piano. Questo è un punto H di π , detto *proiezione ortogonale* di P_0 su π .

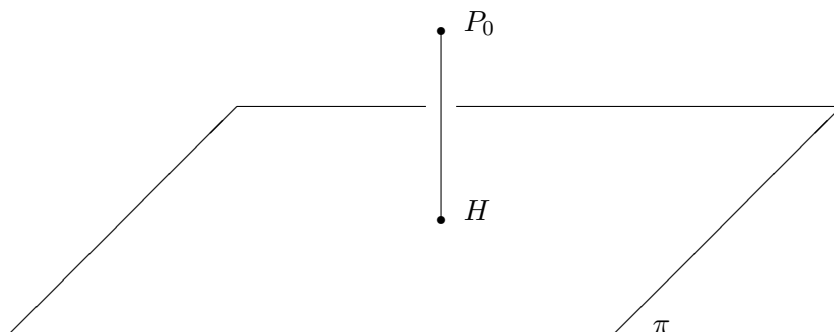


Figura 5: Proiezione ortogonale di un punto su un piano

È facile dimostrare che H è il punto di π più vicino a P_0 : per il teorema di Pitagora, ogni punto Q del piano, diverso da H , ha distanza $d(P_0, Q) > d(P_0, H)$.

La distanza di P_0 dalla sua proiezione ortogonale H è detta *distanza di P_0 da π* e si denota $d(P_0, \pi)$. Dunque

- $d(P_0, \pi)$ è la distanza minima di P_0 da un punto del piano π .

In conclusione, per trovare la distanza di P_0 dal piano π , occorre:

1. trovare la retta r passante per P_0 e ortogonale al piano π ;
2. calcolare il punto d'intersezione $H = r \cap \pi$ (proiezione ortogonale);
3. calcolare la distanza di P_0 da H .

Esempio Dato il punto $P_0 = (1, 1, 3)$ e il piano $x - y + 2z = 0$:

- a) Trovare le coordinate della proiezione ortogonale di P_0 su π .
 b) Calcolare la distanza di P_0 da π .

Soluzione. La retta per P_0 ortogonale al piano ha equazioni parametriche:

$$r : \begin{cases} x = 1 + t \\ y = 1 - t \\ z = 3 + 2t \end{cases}.$$

La proiezione ortogonale H si ottiene come intersezione $\pi \cap r$. Sostituendo le equazioni parametriche nell'equazione del piano, otteniamo $t = -1$ e il punto $H = (0, 2, 1)$. Si ha infine:

$$d(P_0, \pi) = d(P_0, H) = \sqrt{6}.$$

5.1 Formula della distanza di un punto da un piano

C'è una formula esplicita per il calcolo della distanza di un punto da un piano.

Proposizione Dato il piano di equazione $\pi : ax + by + cz + d = 0$ e il punto $P_0 = (x_0, y_0, z_0)$ si ha:

$$d(P_0, \pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Dimostrazione. Omessa. \square

Esempio Siano $P_0 = (1, 1, 3)$ e $\pi : x - y + 2z = 0$. Allora

$$d(P_0, \pi) = \frac{|1 - 1 + 6|}{\sqrt{1 + 1 + 4}} = \frac{6}{\sqrt{6}} = \sqrt{6}.$$

6 Distanza di un punto da una retta

La distanza di un punto P_0 dalla retta r è definita, in modo naturale, come la distanza minima di P_0 da un punto di r .

Il punto di distanza minima si ottiene come intersezione di r con il piano passante per P_0 e ortogonale a r (tale punto H è detto *proiezione ortogonale* di P_0 su r). La distanza minima di P_0 da r sarà quindi la distanza del punto P_0 dalla sua proiezione ortogonale H .

Esempio Calcolare la distanza minima dell'origine dalla retta $r : \begin{cases} x = 1 - t \\ y = 2t \\ z = 5 - t \end{cases}$.

Soluzione. Il piano per l'origine perpendicolare a r ha equazione $\pi : -x + 2y - z = 0$ e quindi $r \cap \pi$ si ottiene sostituendo le equazioni parametriche nell'equazione del piano. Si ottiene l'equazione

$$-(1-t) + 4t - (5-t) = 0,$$

che dà $t = 1$. La proiezione ortogonale è dunque $H = (0, 2, 4)$ e la distanza di P_0 da r vale $d(P_0, H) = \sqrt{20}$.

In modo alternativo, il punto mobile su r ha coordinate $P = (1-t, 2t, 5-t)$. La sua distanza dall'origine, elevata al quadrato, è:

$$f(t) = (1-t)^2 + 4t^2 + (5-t)^2 = 6t^2 - 12t + 26.$$

Cerchiamo il valore di t che rende minima $f(t)$: tale valore renderà minima anche la distanza di P_0 da un punto di r . Annullando la derivata prima di $f(t)$ otteniamo il solo valore $t = 1$, che dà luogo al risultato trovato in precedenza. \square

6.1 Proiezione ortogonale di una retta su un piano

Data una retta r e un piano π , consideriamo l'insieme dei punti che si ottengono proiettando ortogonalmente tutti i punti di r sul piano. Si ottiene così una retta r' (contenuta in π) detta *proiezione ortogonale* di r su π . Osserviamo che:

- la proiezione ortogonale di r su π è la retta r' intersezione di π con il piano π' contenente r e perpendicolare a π :

$$r' = \pi' \cap \pi.$$

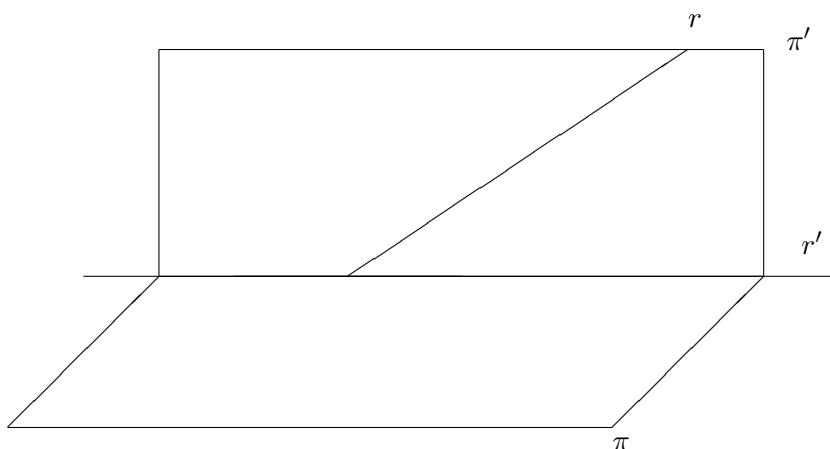


Figura 6: Proiezione ortogonale della retta r sul piano π

Esempio Calcoliamo la retta proiezione ortogonale della retta $r : \begin{cases} x - y = 0 \\ x + 2z = 0 \end{cases}$ sul piano $\pi : 3x - y - 2z = 0$.

Soluzione. Il piano π' contenente r e perpendicolare a π è $\pi' : 5x - y + 8z = 0$. Dunque

$$r' : \begin{cases} 3x - y - 2z = 0 \\ 5x - y + 8z = 0 \end{cases}.$$

7 Distanza di due rette

Vogliamo ora calcolare la distanza tra due rette r, r' dello spazio, definita come distanza minima di un punto di r da un punto di r' , e denotata con $d(r, r')$.

Distinguiamo due casi:

1. le rette sono complanari,
2. le rette sono sghembe.

Nel primo caso, r e r' sono incidenti oppure parallele. Se sono incidenti, la distanza è ovviamente nulla. Se sono parallele distinte, è evidente che la distanza di r da r' uguaglia la distanza di un punto qualunque H di r dalla retta r' :

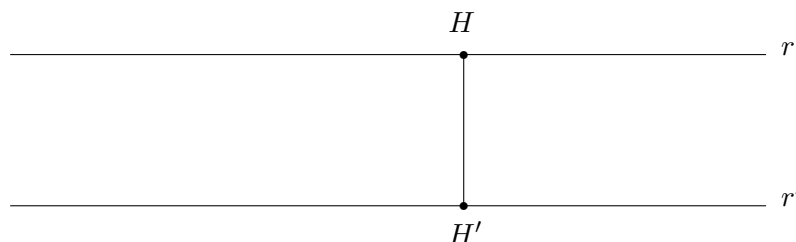


Figura 7: $d(r, r') = d(H, H')$

Esempio Calcoliamo la distanza delle rette parallele $r : \begin{cases} x = t \\ y = -2t \\ z = t \end{cases}$, $r' : \begin{cases} x = 4 - t \\ y = 2t \\ z = 2 - t \end{cases}$.

Soluzione. Le rette sono parallele perché hanno parametri direttori proporzionali dati da, rispettivamente: $(1, -2, 1)$ e $(-1, 2, -1)$. Osserviamo che r passa per l'origine: dunque

$d(r, r') = d(O, r')$. Il piano π , passante per l'origine e ortogonale a r' ha equazione cartesiana $x - 2y + z = 0$ e la proiezione ortogonale di O su r' è $H' = \pi \cap r' = (3, 2, 1)$. Dunque la distanza di r da r' vale $d(O, H') = \sqrt{14}$. \square

Supponiamo ora che r e r' siano sghembe. Si può dimostrare che esistono sempre due punti $H \in r$ e $H' \in r'$ che minimizzano la distanza di r da r' , cioè tali che:

$$d(H, H') \leq d(P, P') \quad \text{per ogni } P \in r, P' \in r',$$

e si ha per definizione $d(r, r') = d(H, H')$. Usando il teorema di Pitagora, si osserva che il segmento HH' dovrà essere ortogonale sia a r che a r' . Dunque abbiamo la seguente

Proposizione *La distanza di due rette r, r' è minimizzata da due punti $H \in r$ e $H' \in r'$ tali che il vettore $\overrightarrow{HH'}$ (quindi la retta per H e H') risulta ortogonale sia a r che a r' .*

Per calcolare la distanza minima di due rette sghembe, dovremo quindi trovare due tali punti H e H' .

Esempio Calcolare la distanza minima delle due rette $r : \begin{cases} x - y = 0 \\ z - 2 = 0 \end{cases}, r' : \begin{cases} x - z = 0 \\ y - 1 = 0 \end{cases}$.

Soluzione. Si verifica che r e r' non hanno intersezione. Le equazioni parametriche di r e r' sono

$$r : \begin{cases} x = t \\ y = t \\ z = 2 \end{cases}, \quad r' : \begin{cases} x = s \\ y = 1 \\ z = s \end{cases}$$

dove t e s sono i rispettivi parametri. I parametri direttori di r sono $(1, 1, 0)$ mentre quelli di r' sono $(1, 0, 1)$ e le rette non sono parallele. Dunque r e r' sono sghembe.

Il punto mobile su r ha coordinate $H = (t, t, 2)$ mentre quello su r' ha coordinate $H' = (s, 1, s)$. Consideriamo il vettore

$$\overrightarrow{HH'} = (s - t, 1 - t, s - 2)$$

e imponiamo che sia ortogonale sia a r che a r' . Utilizzando il prodotto scalare, dobbiamo quindi imporre che $(s - t, 1 - t, s - 2)$ abbia prodotto scalare nullo sia con $(1, 1, 0)$ che con $(1, 0, 1)$. Otteniamo il sistema:

$$\begin{cases} s - 2t + 1 = 0 \\ 2s - t - 2 = 0 \end{cases}$$

che dà l'unica soluzione $t = \frac{4}{3}, s = \frac{5}{3}$. Dunque

$$H = (\frac{4}{3}, \frac{4}{3}, 2), \quad H' = (\frac{5}{3}, 1, \frac{5}{3}).$$

Osserviamo che i punti H e H' sono unici. In conclusione la distanza minima di r da r' vale

$$d(r, r') = d(H, H') = \frac{1}{\sqrt{3}}.$$

□

Nell'esempio precedente, i punti H e H' sono unici, dunque la retta s per H e H' è ortogonale e incidente sia a r che a r' , ed è l'unica con tale proprietà.

Questo è un fatto generale, che sottolinea una proprietà interessante delle rette sghembe.

Teorema *Date due rette sghembe r, r' , esiste un'unica retta s perpendicolare e incidente sia a r che a r' , detta retta di minima distanza.*

Dimostrazione. Se $H \in r$ e $H' \in r'$ sono i punti che minimizzano la distanza di r da r' , allora $H \neq H'$ e la retta s per H e H' è ortogonale sia a r che a r' per quanto osservato in precedenza. Supponiamo che s' sia una seconda retta con tale proprietà, diversa da s . Poiché esiste un'unica direzione ortogonale sia a r che a r' , le rette s e s' risultano parallele e sono quindi contenute in un piano π . Detti $K = s' \cap r$ e $K' = s' \cap r'$, si avrà che π contiene tutti e quattro i punti H, H', K, K' . Ora π contiene $H, K \in r$, dunque contiene tutta la retta r . Analogamente, π contiene $H', K' \in r'$, dunque contiene anche r' : ma questo è impossibile, poiché r e r' , essendo per ipotesi sghembe, non sono complanari. La conclusione è che $s' = s$ e la retta s è unica. □

Se r e r' sono parallele distinte, ci sono infinite rette perpendicolari e incidenti sia a r che a r' , ottenute intersecando il piano di r e r' con un qualunque piano ortogonale a entrambe.

8 Circonferenze

In questa sezione torniamo alla geometria del piano, in cui supponiamo di aver fissato un sistema di riferimento cartesiano. Vedremo quindi come rappresentare una circonferenza del piano tramite un'equazione nelle variabili x, y .

Definizione *Una circonferenza è il luogo dei punti del piano equidistanti da un punto fisso, detto centro. La distanza di un qualunque punto della circonferenza dal centro è detta raggio.*

Per determinare l'equazione di una circonferenza, supponiamo che il centro C abbia coordinate (α, β) e che il raggio sia uguale a R . Allora il punto $P = (x, y)$ appartiene alla circonferenza se e solo se

$$d(P, C) = R. \quad (2)$$

Elevando al quadrato ambo i membri della (2) e utilizzando la formula della distanza otteniamo

$$(x - \alpha)^2 + (y - \beta)^2 = R^2, \quad (3)$$

e sviluppando:

$$x^2 + y^2 - 2\alpha x - 2\beta y + \alpha^2 + \beta^2 - R^2 = 0.$$

Ponendo $a = -2\alpha, b = -2\beta, c = \alpha^2 + \beta^2 - R^2$, tale equazione diventa:

$$x^2 + y^2 + ax + by + c = 0, \quad (4)$$

detta *equazione cartesiana* della circonferenza.

Esempio Scrivere l'equazione cartesiana della circonferenza di centro $C = (2, -1)$ e raggio 2.

Soluzione. Si ha:

$$(x - 2)^2 + (y + 1)^2 = 4,$$

ovvero

$$x^2 + y^2 - 4x + 2y + 1 = 0.$$

□

Notiamo che l'equazione di una circonferenza ha le seguenti caratteristiche:

- è di secondo grado;
- i coefficienti di x^2 e y^2 sono uguali a 1;
- non è presente il termine misto xy .

Ci chiediamo ora se un'equazione del tipo $x^2 + y^2 + ax + by + c = 0$ rappresenta sempre una circonferenza. A tale scopo, cerchiamo, se possibile, di riscriverla nella forma (3).

Sviluppando i quadrati in (3) e uguagliando i coefficienti otteniamo

$$\alpha = -\frac{a}{2}, \quad \beta = -\frac{b}{2}, \quad R^2 = \frac{a^2}{4} + \frac{b^2}{4} - c.$$

Poichè il raggio è un numero positivo, si dovrà avere $\frac{a^2}{4} + \frac{b^2}{4} - c > 0$, e questa è l'unica condizione. In conclusione, abbiamo dimostrato il seguente risultato.

Proposizione a) Una circonferenza del piano si rappresenta con un'equazione di secondo grado in x e y , del tipo

$$x^2 + y^2 + ax + by + c = 0. \quad (5)$$

b) Viceversa, un'equazione del tipo (5) rappresenta una circonferenza se e solo se:

$$a^2 + b^2 - 4c > 0.$$

Se tale condizione è verificata, il centro ha coordinate $C = (-\frac{a}{2}, -\frac{b}{2})$ e il raggio vale:

$$R = \frac{1}{2}\sqrt{a^2 + b^2 - 4c}.$$

Esempio L'equazione $x^2 + y^2 + x - y + 6 = 0$ non rappresenta una circonferenza poiché:

$$a^2 + b^2 - 4c = -22 < 0.$$

Esempio L'equazione $x^2 + y^2 + x - 2y - 1 = 0$ è una circonferenza, poiché:

$$a^2 + b^2 - 4c = 9 > 0.$$

Il centro è $C = (-\frac{1}{2}, 1)$ e il raggio vale $R = \frac{3}{2}$.

Notiamo infine che, se $c \leq 0$, si ha sempre una circonferenza; inoltre $c = 0$ se e solo se tale circonferenza passa per l'origine.

8.1 Circonferenza per tre punti

Dati due punti distinti del piano, diciamo A e B , ci sono infinite circonferenze che passano per A e B : i centri di tali circonferenze saranno sull'asse del segmento AB che, come sappiamo, è il luogo dei punti del piano equidistanti da A e B . Vogliamo ora dimostrare la seguente

Proposizione Per tre punti non allineati del piano passa una e una sola circonferenza.

Dimostrazione. Siano P_1, P_2, P_3 i tre punti e consideriamo l'asse del segmento P_1P_2 , diciamo s , e l'asse del segmento P_1P_3 , diciamo s' . Poiché i tre punti non sono allineati, i due assi non sono paralleli, e si incontrano in un punto $C = s \cap s'$. È evidente che C è equidistante da P_1, P_2 e P_3 . Se R è la distanza di C da uno qualunque dei tre punti:

$$R = d(C, P_1) = d(C, P_2) = d(C, P_3)$$

allora P_1, P_2 e P_3 appartengono tutti alla circonferenza di centro C e raggio R . Inoltre, è anche chiaro che la circonferenza per P_1, P_2 e P_3 è unica. \square

Esempio Determinare il centro e il raggio della circonferenza per i tre punti non allineati $O = (0, 0), A = (4, 2), B = (1, -2)$.

Soluzione. Primo metodo. Troviamo il centro intersecando l'asse del segmento OA con l'asse del segmento OB . Il punto medio del segmento OA è $M = (2, 1)$, e quindi il suo asse (retta per M ortogonale a OA) ha equazione $2x + y - 5 = 0$. In modo analogo, risulta che l'asse del segmento OB ha equazione $2x - 4y - 5 = 0$. Quindi le coordinate del centro si ottengono risolvendo il sistema

$$\begin{cases} 2x + y - 5 = 0 \\ 2x - 4y - 5 = 0 \end{cases}$$

da cui $C = (\frac{5}{2}, 0)$. Il raggio è la distanza di uno qualunque dei tre punti dal centro, e vale $R = \frac{5}{2}$. L'equazione è $x^2 + y^2 - 5x = 0$.

Secondo metodo. Partiamo dall'equazione della circonferenza generica:

$$x^2 + y^2 + ax + by + c = 0,$$

e imponiamo il passaggio per i tre punti. Otteniamo le equazioni:

$$\begin{cases} c = 0 \\ 4a + 2b + c + 20 = 0 \\ a - 2b + c + 5 = 0 \end{cases}$$

Il sistema ammette un'unica soluzione: $a = -5, b = 0, c = 0$ dunque l'equazione della circonferenza è

$$x^2 + y^2 - 5x = 0,$$

da cui il centro: $C = (\frac{5}{2}, 0)$ e il raggio $R = \frac{5}{2}$. \square

Esempio a) Scrivere l'equazione della circonferenza γ di centro $C = (1, 3)$ passante per il punto $A = (-1, -1)$.

b) Stabilire se il punto $B = (4, 0)$ è interno alla circonferenza.

Soluzione. a) Il raggio uguaglia la distanza di C da A , che vale $R = \sqrt{20}$. Dunque l'equazione è

$$(x - 1)^2 + (y - 3)^2 = 20,$$

ovvero $x^2 + y^2 - 2x - 6y - 10 = 0$.

b) I punti interni alla circonferenza hanno distanza da C minore del raggio, che vale $\sqrt{20}$. Ora $d(C, B) = \sqrt{18} < \sqrt{20}$ dunque B è interno a γ . \square

8.2 Retta tangente a una circonferenza

L'intersezione di una circonferenza γ con una retta r risulta di tre tipi:

- L'insieme vuoto (la retta è *esterna* alla circonferenza): in tal caso la distanza della retta r dal centro è maggiore del raggio.
- Un insieme di due punti: cio' avviene se la distanza di r dal centro è minore del raggio.
- Un punto: in tal caso la retta si dice *tangente* alla circonferenza, e la distanza di r dal centro è uguale al raggio.

Osserviamo che, se P è il punto di tangenza, allora il vettore \overrightarrow{PC} è ortogonale alla retta.

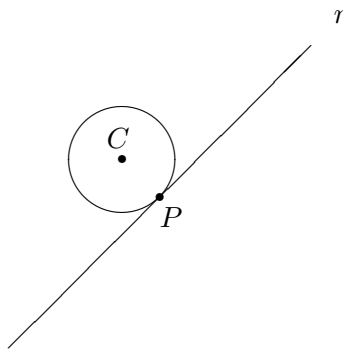


Figura 8: Retta tangente a una circonferenza

Esempio Determinare la circonferenza di centro $C = (2, 0)$ tangente alla retta $x - y = 0$.

Soluzione. Il raggio è uguale a $d(C, r) = \frac{2}{\sqrt{2}} = \sqrt{2}$. Dunque γ ha equazione $(x-2)^2 + y^2 = 2$ ovvero:

$$x^2 + y^2 - 4x + 2 = 0.$$

8.3 Altri esempi

Esempio Data la circonferenza $\gamma : x^2 + y^2 - 2x + 4y = 0$ determinare:

- a) L'equazione della retta tangente a γ nell'origine.
- b) Le equazioni delle rette tangenti a γ e parallele alla retta $x + 2y = 0$.

Soluzione. γ è una circonferenza poiché $c = 0$, inoltre γ passa per l'origine. Il suo centro è $C = (1, -2)$. Dunque il raggio vale $d(C, O) = \sqrt{5}$.

a) La tangente a γ nell'origine è la retta passante per l'origine perpendicolare al raggio OC , dunque al vettore \overrightarrow{OC} . Le coordinate di \overrightarrow{OC} sono $(1, -2)$, e la tangente ha equazione

$$x - 2y = 0.$$

b) La retta generica parallela a $x + 2y = 0$ ha equazione $r : x + 2y + k = 0$. Imponiamo la condizione di tangenza, cioè $d(C, r) = \sqrt{5}$:

$$\frac{|k - 3|}{\sqrt{5}} = \sqrt{5},$$

e otteniamo le due soluzioni $k = 8$ e $k = -2$. Le rette cercate sono due, di equazioni:

$$x + 2y + 8 = 0, \quad x + 2y - 2 = 0.$$

□

Esempio Sono dati i punti $A = (2, 0)$, $B = (4, 4)$. Determinare, se possibile:

- a) una circonferenza di raggio 5 passante per A e B ,
- b) una circonferenza di raggio 1 passante per A e B .

Soluzione. a) Il centro appartiene all'asse del segmento AB : il punto medio è $M = (3, 2)$ dunque l'asse di AB ha equazione cartesiana

$$\alpha : x + 2y - 7 = 0,$$

ed equazioni parametriche

$$\begin{cases} x = -2t + 7 \\ y = t \end{cases}.$$

Quindi il centro ha coordinate $C = (-2t + 7, t)$, con $t \in \mathbf{R}$. Affinche' il raggio sia uguale a 5 imponiamo $d(C, A)^2 = 25$ e otteniamo l'equazione: $(-2t + 5)^2 + t^2 = 25$ ovvero

$$5t^2 - 20t = 0,$$

che ha due soluzioni: $t = 0, t = 4$. Dunque otteniamo due circonferenze, con centri $C_1 = (7, 0)$, $C_2 = (-1, 4)$ ed equazioni, rispettivamente:

$$\gamma_1 : (x - 7)^2 + y^2 = 25, \quad \gamma_2 : (x + 1)^2 + (y - 4)^2 = 25.$$

b) Non esiste. Infatti il raggio minimo di una circonferenza passante per A e B si ottiene quando AB è un diametro, e vale $\sqrt{5}$. Poiché $\sqrt{5}$ è maggiore di 1, tale circonferenza non esiste. □

9 Sfere

Studiamo ora le sfere dello spazio tridimensionale, in cui abbiamo fissato un sistema di riferimento cartesiano.

Definizione Una sfera è il luogo dei punti dello spazio equidistanti da un punto fisso, detto centro. La distanza di un qualunque punto della sfera dal centro è detta raggio.

Per descrivere una sfera con un'equazione, procediamo come nel caso della circonferenza. Se il centro C ha coordinate (α, β, γ) e il raggio è R , allora il punto di coordinate (x, y, z) appartiene alla sfera se e solo se:

$$(x - \alpha)^2 + (y - \beta)^2 + (z - \gamma)^2 = R^2.$$

Tale equazione può essere riscritta nel modo seguente

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0,$$

dando luogo alla *equazione cartesiana* della sfera.

Esempio Scrivere l'equazione cartesiana della sfera di centro $C = (2, -1, 1)$ e raggio 3.

Soluzione. Si ha:

$$(x - 2)^2 + (y + 1)^2 + (z - 1)^2 = 9,$$

ovvero

$$x^2 + y^2 + z^2 - 4x + 2y - 2z - 3 = 0$$

□

Viceversa, sotto opportune condizioni, si ha che l'insieme dei punti dello spazio che soddisfano un'equazione del tipo precedente è una sfera. Precisamente:

Proposizione Un'equazione del tipo $x^2 + y^2 + z^2 + ax + by + cz + d = 0$ rappresenta una sfera se e solo se:

$$a^2 + b^2 + c^2 - 4d > 0.$$

Se tale condizione è verificata, il centro C ha coordinate $(-\frac{a}{2}, -\frac{b}{2}, -\frac{c}{2})$ e il raggio vale:

$$R = \frac{1}{2} \sqrt{a^2 + b^2 + c^2 - 4d}.$$

Esempio L'equazione $x^2 + y^2 + z^2 = 1$ rappresenta la sfera di centro l'origine e raggio 1.

Esempio L'equazione $x^2 + y^2 + z^2 + 1 = 0$ non ha soluzioni, quindi non rappresenta una sfera.

Esempio L'equazione $x^2 + y^2 + z^2 + x - 2y - 1 = 0$ è una sfera, poiché

$$a^2 + b^2 + c^2 - 4d = 9 > 0.$$

Il centro è $C = (-\frac{1}{2}, 1, 0)$ e il raggio vale $R = \frac{3}{2}$.

9.1 Asse di un segmento

L'*asse* di un segmento dello spazio è il luogo dei punti equidistanti dagli estremi del segmento. Si dimostra facilmente che

- L'asse di un segmento è il piano ortogonale al segmento passante per il suo punto medio.

Esempio Determiniamo l'equazione dell'asse del segmento di estremi $A = (1, 2, 3), B = (2, 0, 1)$. Il punto medio è $M = (\frac{3}{2}, 1, 2)$; un piano ortogonale al segmento ha vettore normale $\overrightarrow{AB} = (1, -2, -2)$. Dunque l'asse ha equazione $(x - \frac{3}{2}) - 2(y - 1) - 2(z - 2) = 0$, cioè

$$2x - 4y - 4z + 9 = 0.$$

9.2 Sfera per quattro punti

Osserviamo innanzitutto che per tre punti non allineati dello spazio passano infinite sfere (dove sono i centri?). Abbiamo però il seguente risultato.

Proposizione *Per quattro punti non complanari dello spazio passa una e una sola sfera.*

Dimostrazione. Siano P_1, P_2, P_3, P_4 i quattro punti, e siano (x_i, y_i, z_i) le coordinate di $P_i, i = 1, 2, 3, 4$. Imponiamo all'equazione generica della sfera:

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0$$

il passaggio per $P_i, i = 1, 2, 3, 4$. Otteniamo così un sistema lineare di quattro equazioni nelle quattro incognite a, b, c, d , la cui matrice dei coefficienti è

$$A = \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{pmatrix}.$$

Vogliamo dimostrare che A ha determinante non nullo. In effetti, sottraendo la prima riga dalle altre vediamo che

$$\det A = \det \begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 & 0 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 & 0 \\ x_4 - x_1 & y_4 - y_1 & z_4 - z_1 & 0 \end{pmatrix} = \det \begin{pmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \end{pmatrix},$$

e l'ultimo determinante è non nullo poiché per ipotesi i punti non sono complanari. Per il teorema di Cramer, esiste un'unica soluzione (a, b, c, d) del sistema, quindi una unica sfera passante per i quattro punti. \square

Corollario *Fissati quattro punti non complanari, esiste un unico punto dello spazio equidistante dai quattro punti dati.*

Dimostrazione. Il punto cercato è il centro dell'unica sfera passante per i quattro punti. \square

Esempio Determinare l'equazione cartesiana della sfera passante per l'origine e per i punti $A = (1, 0, 0)$, $B = (0, 2, 0)$, $C = (0, 0, 3)$.

Soluzione. Dall'equazione generica della sfera: $x^2 + y^2 + z^2 + ax + by + cz + d = 0$, imponendo il passaggio per i quattro punti otteniamo $a = -1, b = -4, c = -9, d = 0$ dunque l'equazione è

$$x^2 + y^2 + z^2 - x - 2y - 3z = 0.$$

Il centro è $C = (\frac{1}{2}, 1, \frac{3}{2})$ e il raggio vale $r = \frac{1}{2}\sqrt{14}$. \square

Esempio Determinare il punto del piano $\pi : 2x + y - z = 0$ equidistante dai punti $O = (0, 0, 0)$, $A = (1, 0, 0)$, $B = (0, 2, 0)$.

Soluzione. Partiamo dall'equazione generica della sfera e imponiamo il passaggio per i punti O, A, B : otteniamo l'equazione

$$x^2 + y^2 + z^2 - x - 2y + cz = 0$$

con $c \in \mathbf{R}$. In particolare, esistono infinite sfere per i tre punti. Il centro della sfera generica per O, A, B è $C = (\frac{1}{2}, 1, -\frac{c}{2})$, che per definizione è equidistante dai tre punti. Ora imponiamo che C appartenga a π : questo avviene se e solo se $c = -4$. Dunque $C = (\frac{1}{2}, 1, 2)$ è il punto cercato. \square

9.3 Piano tangente

- Un piano si dice *tangente* a una sfera se l'intersezione del piano con la sfera si riduce a un punto.

Risulta che il piano π è tangente alla sfera σ se e solo se la distanza di π dal centro di σ è uguale al raggio. (Osserviamo che se tale distanza è minore del raggio, allora π incontra σ in una circonferenza).

Inoltre, il piano tangente alla sfera σ in un punto $P \in \sigma$ ha vettore normale \overrightarrow{PC} , dove C è il centro della sfera.

Esempio a) Verificare che l'equazione $x^2 + y^2 + z^2 - 2x + 4z - 6 = 0$ rappresenta una sfera σ .

b) Determinare l'equazione del piano tangente a σ nel suo punto $P_0 = (2, 1, 1)$.

Soluzione. a) Possiamo riscrivere l'equazione nella forma:

$$(x - 1)^2 + y^2 + (z + 2)^2 = 11,$$

quindi σ è la sfera di centro $C = (1, 0, -2)$ e raggio $\sqrt{11}$.

b) Il piano π , tangente alla sfera in P_0 , è ortogonale al vettore $\overrightarrow{P_0C}$. Dunque i parametri di giacitura di π saranno proporzionali alle coordinate del vettore $\overrightarrow{P_0C}$, che sono $(-1, -1, -3)$. L'equazione del piano tangente è dunque $(x - 2) + (y - 1) + 3(z - 1) = 0$, ovvero:

$$x + y + 3z - 6 = 0.$$

Si può verificare che $\pi \cap \sigma$ si riduce al punto P_0 . \square