Esercizi relativi ai Complementi

4C.1 Determina per quali $\lambda \in \mathbb{C}$ l'equazione $(\lambda^8 - 1)z^3 + (\lambda^4 - 1)z + \lambda^2 - i = 0$ non ammette soluzioni.

4C.2 Trova un polinomio a coefficienti reali di quinto grado con solo una radice reale, e un polinomio a coefficienti reali di quarto grado senza nessuna radice reale.

4C.3 Dato un polinomio $p(t) = a_n t^n + \cdots + a_0 \in \mathbb{Z}[t]$ a coefficienti interi, prendiamo una radice razionale $r = a/b \in \mathbb{Q}$ di p, dove $a, b \in \mathbb{Z}$ sono relativamente primi. Dimostra che a divide a_0, e che b divide a_n.

4C.4 Trova le radici dei polinomi a coefficienti interi $p_1(t) = t^4 - 2t^3 - 5t^2 + 4t + 6$ e $p_2(t) = t^5 - 5t^3 - 2t^2 + 6t + 4$. (*Suggerimento*: usa l'esercizio precedente.)

Applicazioni lineari

5

Sommario

In matematica, ogni volta che si introduce una nuova struttura si studiano immediatamente anche le funzioni che la rispettano. Nel caso degli spazi vettoriali, queste funzioni si chiamano applicazioni lineari, e sono l'argomento di questo capitolo. Ne studieremo le proprietà principali, giungendo a definire il rango di un'applicazione lineare, un numero naturale che ne riassume le principali caratteristiche. Dimostreremo inoltre il fondamentale Teorema della dimensione, che lega il rango di un'applicazione lineare alla dimensione del dominio.

5.1 Esempi e definizioni

Nel capitolo precedente abbiamo visto che le soluzioni di un sistema lineare omogeneo formano un sottospazio vettoriale. La struttura dell'insieme delle soluzioni di un sistema lineare qualunque è descritta nel seguente *Teorema di struttura per i sistemi lineari.*

Proposizione 5.1 (Teorema di struttura)

Sia $v^o \in \mathbb{R}^n$ una soluzione del sistema lineare $Ax = b$ di ordine n. Allora ogni altra soluzione è della forma $v = v^o + w$, dove $w \in \mathbb{R}^n$ è una soluzione del sistema omogeneo $Ax = O$. In altre parole, se $L \subseteq \mathbb{R}^n$ è l'insieme delle soluzioni del sistema $Ax = b$ e $W \subseteq \mathbb{R}^n$ è l'insieme delle soluzioni del sistema omogeneo associato $Ax = O$, si ha

$$L = v^o + W = \{v^o + w \mid w \in W\}.$$

In particolare, v^o è l'unica soluzione del sistema se e solo se le colonne di A sono linearmente indipendenti.

Dimostrazione

Posto $v^o = (v^o_1, \ldots, v^o_n)$, sia $w = (w_1, \ldots, w_n)$ una soluzione del sistema omogeneo. Allora

$$v^o_1 A^1 + \cdots + v^o_n A^n = b \quad \text{e} \quad w_1 A^1 + \cdots + w_n A^n = O,$$

per cui sommando otteniamo

$$(v^o_1 + w_1) A^1 + \cdots + (v^o_n + w_n) A^n = b,$$

cioè $v^o + w$ è un'altra soluzione del sistema $Ax = b$.

Sia ora v un'altra soluzione di $Ax = b$; dobbiamo dimostrare che $w = v - v^o$ è una soluzione del sistema $Ax = O$. Infatti abbiamo

$$v_1 A^1 + \cdots + v_n A^n = b \quad \text{e} \quad v^o_1 A^1 + \cdots + v^o_n A^n = b;$$

sottraendo otteniamo

$$(v_1 - v^o_1) A^1 + \cdots + (v_n - v^o_n) A^n = O,$$

come voluto. L'ultima affermazione segue dalla Proposizione 4.4. □

Osservazione 5.1

Di nuovo, mentre l'esistenza di una soluzione del sistema $Ax = b$ dipende dalla specifica relazione fra A e b (vedi la Proposizione 4.3), l'unicità dipende soltanto dalla matrice dei coefficienti.

Se ci facessimo trasportare dalle notazioni, potremmo credere di avere in mano una dimostrazione molto più veloce della Proposizione 5.1. Infatti, da $Av^o = b$ e $Aw = O$ potremmo essere portati a dedurre $A(v^o + w) = Av^o + Aw = b + O = b$, e prendere questa come dimostrazione del fatto che $v^o + w$ risolve $Ax = b$. Il problema è che, fino a ora, $Ax = b$ era soltanto una notazione abbreviata per indicare l'intero sistema lineare, per cui la prima uguaglianza in questa "dimostrazione" non sembra avere molto senso. L'obiettivo di questo capitolo è introdurre quanto ci serve per renderla corretta, e poi procedere verso lidi anche più interessanti.

Definizione 5.1

Sia $A \in M_{m,n}(\mathbb{R})$ una matrice con m righe e n colonne, cioè

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

con $A^1, \ldots, A^n \in \mathbb{R}^m$. Allora definiamo un'applicazione $L_A : \mathbb{R}^n \to \mathbb{R}^m$ ponendo

$$L_A(x) = \begin{vmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{vmatrix} = x_1 A^1 + \cdots + x_n A^n \in \mathbb{R}^m,$$

per ogni $x \in \mathbb{R}^n$. Nota che x è soluzione del sistema $Ax = b$ se e solo se $L_A(x) = b$ (perché?). A volte, invece di scrivere $L_A(x)$ scriveremo Ax o anche $A \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}$.

Osservazione 5.2

Perché $L_A(x)$ o Ax abbiano senso, occorre che *il numero di colonne di A coincida col numero di coordinate – di righe – del vettore x.* Notiamo poi che utilizzando una matrice $A \in M_{m,n}(\mathbb{K})$ a coefficienti in $\mathbb{K}$ possiamo definire del tutto in maniera analoga un'applicazione $L_A : \mathbb{K}^n \to \mathbb{K}^m$.

Le applicazioni L_A hanno due proprietà fondamentali:

$$L_A(x+y) = L_A(x) + L_A(y) \quad \text{e} \quad L_A(\lambda x) = \lambda L_A(x), \tag{5.1}$$

quali che siano $x, y \in \mathbb{K}^n$ e $\lambda \in \mathbb{K}$ (esercizio). Grazie a queste proprietà possiamo in particolare dare senso (vedi l'Esercizio 5.1) alla precedente "dimostrazione". Questo suggerisce una definizione.

Definizione 5.2

Un'*applicazione* (o *trasformazione*) *lineare* fra due spazi vettoriali V e W è una funzione $T : V \to W$ tale che:

(1) $T(v_1 + v_2) = T(v_1) + T(v_2)$ per tutti i $v_1, v_2 \in V$ (diremo che T è *additiva*);

(2) $T(\lambda v) = \lambda T(v)$ per tutti i $\lambda \in \mathbb{R}$ e $v \in V$ (diremo che T è *omogenea*).

In altre parole, un'applicazione lineare trasforma (rispetta, conserva) le operazioni dallo spazio di partenza a quello d'arrivo. Se $V = W$, si parla di *endomorfismo* (od *operatore lineare*). Nota, inoltre, che se V e W sono definiti su un campo $\mathbb{K}$ possiamo dare la definizione di applicazione lineare nello stesso modo, richiedendo che la proprietà (2) valga per ogni $\lambda \in \mathbb{K}$.

Osservazione 5.3

Affinché $T : V \to W$ sia lineare, deve soddisfare *sia* (1) *sia* (2). Infatti esistono funzioni non lineari che soddisfano (2) ma non (1) – vedi l'Esempio 5.5 – e funzioni non lineari che soddisfano (1) ma non (2) – basta considerare il coniugio come applicazione da $\mathbb{C}$ in sé (Esercizio 5.30).

Osservazione 5.4

Se $T: V \to W$ è lineare, allora $T(O) = O$: infatti, preso $v \in V$ qualunque, si ha

$$T(O) = T(0v) = 0T(v) = O.$$

Inoltre abbiamo anche

$$T(-v) = T\big((-1)v\big) = (-1)T(v) = -T(v)$$

quale che sia $v \in V$.

Vediamo ora qualche esempio.

Esempio 5.1

L'*applicazione identica* (o *identità*) di uno spazio vettoriale V è per definizione l'applicazione $\mathrm{id}_V: V \to V$ data da $\mathrm{id}_V(v) = v$ per tutti i $v \in V$. L'applicazione identica è ovviamente lineare.

Esempio 5.2

Siano V e W due spazi vettoriali; l'*applicazione nulla* $O: V \to W$ è definita da $O(v) = O$ (dove il primo O è l'applicazione nulla, e il secondo è il vettore nullo di W) per tutti i $v \in V$. Di nuovo, O è evidentemente lineare.

Esempio 5.3

Fissato il campo $\mathbb{K}$, l'applicazione $L_A: \mathbb{K}^n \to \mathbb{K}^m$ è lineare per qualsiasi matrice $A \in M_{m,n}(\mathbb{K})$.

Esempio 5.4

Sia $T: \mathbb{R}^3 \to \mathbb{R}^3$ data da

$$T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} 3x - 2y \\ z \\ x + z \end{vmatrix}.$$

Allora abbiamo

$$\begin{aligned} T\left(\begin{vmatrix} x_1 \\ y_1 \\ z_1 \end{vmatrix} + \begin{vmatrix} x_2 \\ y_2 \\ z_2 \end{vmatrix}\right) &= T\begin{vmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{vmatrix} = \begin{vmatrix} 3(x_1 + x_2) - 2(y_1 + y_2) \\ z_1 + z_2 \\ (x_1 + x_2) + (z_1 + z_2) \end{vmatrix} \\ &= \begin{vmatrix} 3x_1 - 2y_1 \\ z_1 \\ x_1 + z_1 \end{vmatrix} + \begin{vmatrix} 3x_2 - 2y_2 \\ z_2 \\ x_2 + z_2 \end{vmatrix} = T\begin{vmatrix} x_1 \\ y_1 \\ z_1 \end{vmatrix} + T\begin{vmatrix} x_2 \\ y_2 \\ z_2 \end{vmatrix}, \end{aligned}$$

e

$$T\left(\lambda\begin{vmatrix} x \\ y \\ z \end{vmatrix}\right) = T\begin{vmatrix} \lambda x \\ \lambda y \\ \lambda z \end{vmatrix} = \begin{vmatrix} 3\lambda x - 2\lambda y \\ \lambda z \\ \lambda x + \lambda z \end{vmatrix} = \lambda\begin{vmatrix} 3x - 2y \\ z \\ x + z \end{vmatrix} = \lambda T\begin{vmatrix} x \\ y \\ z \end{vmatrix},$$

per cui T è lineare. Nota che $T = L_A$, dove

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

Esempio 5.5

Sia $T\colon \mathbb{R}^2 \to \mathbb{R}$ data da $T\begin{vmatrix} x \\ y \end{vmatrix} = (x^3 + y^3)^{1/3}$. Allora $T(\lambda v) = \lambda T(v)$ per ogni $\lambda \in \mathbb{R}$ e $v \in \mathbb{R}^2$, ma T non è additiva: infatti

$$T\left(\begin{vmatrix} 1 \\ 0 \end{vmatrix} + \begin{vmatrix} 0 \\ 1 \end{vmatrix}\right) = T\begin{vmatrix} 1 \\ 1 \end{vmatrix} = 2^{1/3} \neq 2 = 1 + 1 = T\begin{vmatrix} 1 \\ 0 \end{vmatrix} + T\begin{vmatrix} 0 \\ 1 \end{vmatrix}.$$

Osservazione 5.5

Il punto è che le coordinate di un'applicazione lineare devono essere polinomi di primo grado senza termine noto delle coordinate del vettore di partenza (vedi l'Esercizio 5.9 e l'Esercizio 7.11).

Esempio 5.6

Sia $V = \mathbb{R}_2[t]$ lo spazio vettoriale dei polinomi di grado minore o uguale a 2, e sia $T\colon \mathbb{R}_2[t] \to \mathbb{R}_2[t]$ data da $[T(p)](t) = p(t+1)$ per ogni $p \in \mathbb{R}_2[t]$. Per esempio, se $p(t) = t^2$, allora $T(p)$ è il polinomio $(t+1)^2 = t^2 + 2t + 1$. L'applicazione T è effettivamente lineare: infatti

$$[T(p_1+p_2)](t) = (p_1+p_2)(t+1) = p_1(t+1) + p_2(t+1) = [T(p_1)](t) + [T(p_2)](t),$$
$$[T(\lambda p)](t) = (\lambda p)(t+1) = \lambda\big(p(t+1)\big) = \big[\lambda\big(T(p)\big)\big](t)$$

per ogni $t \in \mathbb{R}$, per cui $T(p_1 + p_2) \equiv T(p_1) + T(p_2)$ e $T(\lambda p) \equiv \lambda T(p)$.

Esempio 5.7

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di uno spazio vettoriale V. Allora ogni $v \in V$ si scrive come $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, dove le coordinate $\alpha_1, \dots, \alpha_n$ sono univocamente determinate. Definiamo $F_{\mathcal{B}}\colon V \to \mathbb{K}^n$ ponendo

$$F_{\mathcal{B}}(v) = \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{vmatrix};$$

in altri termini, $F_{\mathcal{B}}$ associa a ogni $v \in V$ le sue coordinate rispetto alla base $\mathcal{B}$ (e generalizza la $F_{\mathcal{B}}$ vista nel Paragrafo 2.2). Anche $F_{\mathcal{B}}$ è lineare: se $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e $w = \beta_1 v_1 + \cdots + \beta_n v_n$, allora $v + w = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n$ e quindi

$$F_{\mathcal{B}}(v+w) = \begin{vmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{vmatrix} = \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{vmatrix} + \begin{vmatrix} \beta_1 \\ \vdots \\ \beta_n \end{vmatrix} = F_{\mathcal{B}}(v) + F_{\mathcal{B}}(w).$$

Analogamente si dimostra che $F_{\mathcal{B}}(\lambda v) = \lambda F_{\mathcal{B}}$. Nota che $F_{\mathcal{B}}$ è bigettiva (perché?).

Esempio 5.8

Prendiamo due sottospazi U e W di uno spazio vettoriale V tali che $V = U \oplus W$. Questo vuol dire che ogni elemento $v \in V$ si scrive in modo unico come $v = u + w$, con $u \in U$ e $w \in W$. Ponendo $P(v) = u$ definiamo un'applicazione $P\colon V \to U \subset V$, detta *proiezione su U lungo W*. È facile dimostrare che P è lineare; per esempio, se $\lambda \in \mathbb{K}$ e $v = u + w \in V$, allora $\lambda v = \lambda u + \lambda w$ e quindi $P(\lambda v) = \lambda P(v)$. In modo analogo si dimostra che P è additiva.

Esempio 5.9

Fissiamo $a \in \mathbb{K}^n$, e definiamo $\varphi_a\colon \mathbb{K}^n \to \mathbb{K}$ ponendo

$$\forall x \in \mathbb{K}^n \quad \varphi_a(x) = a_1x_1 + \cdots + a_nx_n \in \mathbb{K}.$$

Per esempio, se $\mathbb{K} = \mathbb{R}$, $n = 2$ e $a = \begin{vmatrix} 1 \\ 2 \end{vmatrix}$, allora $\varphi_a(x) = x_1 + 2x_2$. Si verifica facilmente (esercizio) che φ_a è un'applicazione lineare da $\mathbb{K}^n$ in $\mathbb{K}$. ⊓⊔

Esempio 5.10

La *trasposizione* ${}^T\colon M_{m,n}(\mathbb{K}) \to M_{n,m}(\mathbb{K})$, che associa a una matrice A la *trasposta* A^T, è l'applicazione data da

$$A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} \longmapsto A^T = \begin{vmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{vmatrix}.$$

Per esempio,

$$A = \begin{vmatrix} 1 & -1 \\ 0 & 27 \\ -\sqrt{33} & \pi^2 \end{vmatrix} \longmapsto A^T = \begin{vmatrix} 1 & 0 & -\sqrt{33} \\ -1 & 27 & \pi^2 \end{vmatrix}.$$

In termini di righe, colonne ed elementi abbiamo

$$(A^T)_{ij} = a_{ji}, \qquad (A^T)_i = (A^i)^T \quad \text{e} \quad (A^T)^j = (A_j)^T, \tag{5.2}$$

dove $A = (a_{ij})$ e $(A^T)_{ij}$ è l'elemento di posto (i,j) della matrice A^T. Capito questo, è evidente che la trasposizione è un'applicazione lineare, e che $(A^T)^T = A$. ⊓⊔

Cominciamo ora a studiare le applicazioni lineari in generale. La prima osservazione importante è che *un'applicazione lineare è completamente determinata dai valori che assume su una base.*

Proposizione 5.2

Siano V e W due spazi vettoriali, $\{v_1, \ldots, v_n\}$ una base di V, e $w_1, \ldots, w_n \in W$. Allora esiste un'unica applicazione lineare $T\colon V \to W$ tale che $T(v_j) = w_j$ per $j = 1, \ldots, n$. L'applicazione T è tale che

$$T(\alpha_1v_1 + \cdots + \alpha_nv_n) = \alpha_1w_1 + \cdots + \alpha_nw_n \tag{5.3}$$

per ogni $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$.

Dimostrazione

Prima di tutto, verifichiamo che la T definita in (5.3) è lineare. Presi v, $v' \in V$, scriviamoli come combinazione lineare dei vettori della base di V

$$v = \alpha_1v_1 + \cdots + \alpha_nv_n, \quad \text{e} \quad v' = \alpha'_1v_1 + \cdots + \alpha'_nv_n.$$

Allora $v + v' = (\alpha_1 + \alpha'_1)v_1 + \cdots + (\alpha_n + \alpha'_n)v_n$ e

$$\begin{aligned} T(v+v') &= (\alpha_1+\alpha'_1)w_1 + \cdots + (\alpha_n+\alpha'_n)w_n \\ &= (\alpha_1w_1 + \cdots + \alpha_nw_n) + (\alpha'_1w_1 + \cdots + \alpha'_nw_n) = T(v) + T(v'). \end{aligned}$$

Nello stesso modo (esercizio) si dimostra che $T(\lambda v) = \lambda T(v)$.

Quindi l'esistenza è sistemata. Per l'unicità, sia $S\colon V \longrightarrow W$ un'altra applicazione lineare tale che $S(v_j) = w_j$ per $j = 1, \ldots, n$; dobbiamo dimostrare che $S(v) = T(v)$ qualunque sia $v \in V$. Prendiamo allora $v \in V$, e scriviamolo di nuovo come combinazione lineare degli elementi della base, $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Allora

$$S(v) = S(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 S(v_1) + \cdots + \alpha_n S(v_n) = \alpha_1 w_1 + \cdots + \alpha_n w_n = T(v). \qquad \square$$

Corollario 5.3

Siano V e W due spazi vettoriali, $\{v_1, \ldots, v_n\}$ una base di V, e $S, T\colon V \longrightarrow W$ due applicazioni lineari. Supponiamo che si abbia $S(v_j) = T(v_j)$ per $j = 1, \ldots, n$. Allora $S \equiv T$, cioè $S(v) = T(v)$ per tutti i vettori $v \in V$.

Dimostrazione

Segue dall'unicità dell'applicazione costruita nella Proposizione 5.2, con $w_j = T(v_j)$ per $j = 1, \ldots, n$. □

Due matrici diverse danno sempre origine ad applicazioni lineari diverse.

Proposizione 5.4

Siano $A, B \in M_{m,n}(\mathbb{K})$ due matrici. Allora $L_A \equiv L_B$ se e solo se $A = B$.

Dimostrazione

Chiaramente se $A = B$ allora $L_A \equiv L_B$. Viceversa, supponiamo che $L_A \equiv L_B$; siccome si ha

$$\forall j = 1, \ldots, n \qquad L_A(e_j) = A^j \tag{5.4}$$

(dove $e_1, \ldots, e_n$ è la base canonica di $\mathbb{K}^n$), le due matrici A e B devono avere le stesse colonne, per cui sono uguali. □

5.2 Nucleo e immagine

A ogni applicazione lineare $T\colon V \longrightarrow W$ possiamo associare due sottoinsiemi:

a) il *nucleo* $\operatorname{Ker} T = \{v \in V \mid T(v) = O\} \subseteq V$;

b) l'*immagine* $\operatorname{Im} T = T(V) = \{T(v) \mid v \in V\} \subseteq W$.

Il fatto interessante è che nucleo e immagine, oltre a essere sottospazi vettoriali, caratterizzano l'iniettività e la surgettività delle applicazioni lineari.

Proposizione 5.5

Sia $T\colon V \longrightarrow W$ un'applicazione lineare. Allora:

(1) $\operatorname{Ker} T$ *è un sottospazio di* V*;*

(2) $\operatorname{Im} T$ *è un sottospazio di* W*;*

(3) T *è surgettiva se e solo se* $\operatorname{Im} T = W$*;*

(4) T *è iniettiva se e solo se* $\operatorname{Ker} T = \{O\}$.

Dimostrazione

(1) Facciamo vedere che il nucleo è chiuso rispetto alla somma; la chiusura rispetto al prodotto per scalari è analoga, e lasciata per esercizio. Prendiamo $v_1, v_2 \in \operatorname{Ker} T$, cioè tali che si abbia $T(v_1) = T(v_2) = O$; dobbiamo dimostrare che $v_1 + v_2 \in \operatorname{Ker} T$. Infatti si ha $T(v_1 + v_2) = T(v_1) + T(v_2) = O + O = O$.

(2) Stavolta dimostriamo che l'immagine è chiusa rispetto al prodotto per scalari, lasciando la chiusura rispetto alla somma per esercizio. Siano $\lambda \in \mathbb{K}$ e $T(v) \in \operatorname{Im} T$; dobbiamo far vedere che $\lambda T(v) \in \operatorname{Im} T$. Infatti si ha $\lambda T(v) = T(\lambda v) \in \operatorname{Im} T$.

(3) È la definizione di surgettività.

(4) Questo è il punto più interessante, che distingue nettamente le applicazioni lineari dalle funzioni qualsiasi. Se T è iniettiva, $T(v) = O = T(O)$ implica $v = O$ e quindi $\operatorname{Ker} T = \{O\}$. Viceversa, supponiamo $\operatorname{Ker} T = \{O\}$. Prendiamo $v_1, v_2 \in V$ tali che $T(v_1) = T(v_2)$; dobbiamo dimostrare che $v_1 = v_2$. Ma $T(v_1) = T(v_2)$ equivale a $T(v_1) - T(v_2) = O$, da cui segue $T(v_1 - v_2) = O$ e dunque $v_1 - v_2 = O$, dove l'ultima implicazione vale in quanto $\operatorname{Ker} T = \{O\}$. □

Come vedremo, calcolare il nucleo di un'applicazione lineare è più o meno equivalente a risolvere un sistema lineare omogeneo. Per il calcolo dell'immagine di un'applicazione lineare è invece utile il lemma seguente.

Lemma 5.6

Sia $T: V \to W$ un'applicazione lineare, e $\{v_1, \ldots, v_n\}$ una base di V. Allora $\operatorname{Im} T = \operatorname{Span}(T(v_1), \ldots, T(v_n))$.

Dimostrazione

Infatti $\operatorname{Im} T$ è l'insieme degli elementi della forma $T(v)$ al variare di $v \in V$. Ma ogni elemento di V è combinazione lineare di $v_1, \ldots, v_n$; quindi

$$\begin{aligned}\operatorname{Im} T &= \{T(\alpha_1 v_1 + \cdots + \alpha_n v_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} \\ &= \{\alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} = \operatorname{Span}(T(v_1), \ldots, T(v_n)).\end{aligned}$$ □

Osservazione 5.6

Non è detto che $\{T(v_1), \ldots, T(v_n)\}$ sia una base di $\operatorname{Im} T$; in generale, è solo un sistema di generatori (vedi l'Esercizio 5.23).

Esempio 5.11

Sia $T: \mathbb{R}^3 \to \mathbb{R}^3$ data da

$$T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x - y \\ 2y - 2x \\ z \end{vmatrix}.$$

È chiaramente lineare; per essere precisi, $T = L_A$ con

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

Il nucleo contiene i vettori mandati in O, cioè le soluzioni del sistema lineare

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

un breve conto con le tecniche del Capitolo 3 mostra che $\operatorname{Ker} T = \mathbb{R}\begin{vmatrix} 1 \\ 1 \\ 0 \end{vmatrix}$. D'altra parte,

$$\operatorname{Im} T = \operatorname{Span}\left(\begin{vmatrix} 1 \\ -2 \\ 0 \end{vmatrix}, \begin{vmatrix} -1 \\ 2 \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix} \right) = \operatorname{Span}\left(\begin{vmatrix} 1 \\ -2 \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix} \right),$$

come segue subito dal Lemma 5.6 applicato usando la base canonica di $\mathbb{R}^3$.

Osservazione 5.7

Se A è una matrice, spesso e volentieri scriveremo $\operatorname{Ker} A$ invece di $\operatorname{Ker} L_A$, e $\operatorname{Im} A$ invece di $\operatorname{Im} L_A$.

Una delle caratteristiche chiave di un'applicazione lineare (che, come vedremo, è in grado da sola di dire se un sistema lineare ammette o no soluzione) è il rango.

Definizione 5.3

Il *rango* $\operatorname{rg} T$ di un'applicazione lineare $T \colon V \longrightarrow W$ è la dimensione dell'immagine: $\operatorname{rg} T = \dim \operatorname{Im} T$.

Esempio 5.12

Nel caso delle matrici, (5.4) e il Lemma 5.6 ci dicono che

$$\forall A \in M_{m,n}(\mathbb{K}) \qquad \operatorname{Im} L_A = \operatorname{Span}(A^1, \dots, A^n) \subseteq \mathbb{K}^m.$$

Quindi il rango di L_A (o, come spesso diremo, il *rango* di A) è la dimensione dello spazio generato dalle colonne. In particolare, $\operatorname{rg} A \leq \min\{m, n\}$.

Siccome l'immagine di T è un sottospazio di W, chiaramente $\operatorname{rg} T \leq \dim W$. D'altra parte, il Lemma 5.6 ci dice che $\operatorname{rg} T \leq \dim V$. È possibile essere più precisi? La risposta è positiva, ed è contenuta nel fondamentale *Teorema della dimensione*.

Teorema 5.7 (della dimensione)

Sia $T \colon V \longrightarrow W$ un'applicazione lineare. Allora

$$\dim V = \dim \operatorname{Ker} T + \operatorname{rg} T.$$

Dimostrazione

Sia $\{u_1, \dots, u_r\}$ una base di $\operatorname{Ker} T$; col Teorema 4.10 la possiamo completare a una base $\{u_1, \dots, u_r, v_{r+1}, \dots, v_n\}$ di V (se $\operatorname{Ker} T = \{O\}$, prendiamo direttamente una base $\{v_1, \dots, v_n\}$ di V, e consideriamo $r = 0$ ed $s = n$ nel seguito). Adesso poniamo $w_j = T(v_{r+j}) \in W$ per $j = 1, \dots, s = n - r$; se dimostriamo che $\mathcal{B} = \{w_1, \dots, w_s\}$ è una base di $\operatorname{Im} T$ abbiamo finito. Grazie al Lemma 5.6 (e al fatto che $T(u_i) = O$ per $i = 1, \dots, r$) sappiamo già che $\mathcal{B}$ è un sistema di generatori di $\operatorname{Im} T$; dobbiamo solo far vedere che $w_1, \dots, w_s$ sono linearmente indipendenti.

Supponiamo che $\alpha_1, \dots, \alpha_s \in \mathbb{K}$ siano tali che $\alpha_1 w_1 + \cdots + \alpha_s w_s = O$. Allora si ha

$$O = \alpha_1 T(v_{r+1}) + \cdots + \alpha_s T(v_{r+s}) = T(\alpha_1 v_{r+1} + \cdots \alpha_s v_{r+s}),$$

per cui $\alpha_1 v_{r+1} + \cdots \alpha_s v_{r+s} \in \operatorname{Ker} T$. Questo vuol dire che esistono $\beta_1, \dots, \beta_r \in \mathbb{K}$ tali che $\alpha_1 v_{r+1} + \cdots + \alpha_s v_{r+s} = \beta_1 u_1 + \cdots + \beta_r u_r$; quindi

$$\beta_1 u_1 + \cdots + \beta_r u_r - \alpha_1 v_{r+1} - \cdots - \alpha_s v_{r+s} = O,$$

e l'indipendenza lineare di $\{u_1, \dots, u_r, v_{r+1}, \dots, v_{r+s}\}$ implica $\alpha_1 = \cdots = \alpha_s = 0$. □

Osservazione 5.8

Nel caso di una matrice $A \in M_{m,n}(\mathbb{K})$ il Teorema 5.7 diventa $n = \dim \operatorname{Ker} A + \operatorname{rg} A$.

Una prima conseguenza del Teorema della dimensione è che per vedere se un'applicazione lineare è iniettiva o surgettiva basta controllare il rango.

Corollario 5.8

Sia $T\colon V \longrightarrow W$ un'applicazione lineare. Allora:

(1) *T è iniettiva se e solo se* $\operatorname{rg} T = \dim V$;

(2) *T è surgettiva se e solo se* $\operatorname{rg} T = \dim W$;

(3) *se* $\dim V = \dim W$ *(in particolare, se $V = W$), T è iniettiva se e solo se è surgettiva.*

Dimostrazione

(1) L'applicazione T è iniettiva se e solo se $\operatorname{Ker} T = \{O\}$, ovvero se e solo se si ha $\dim \operatorname{Ker} T = 0$; basta allora applicare il Teorema 5.7.

(2) T è surgettiva se e solo se $\operatorname{Im} T = W$, che può succedere se e solo se $\operatorname{rg} T = \dim W$ (Proposizione 4.14).

(3) Basta confrontare (1) e (2). □

Un'altra conseguenza importante è il *Teorema di Rouché-Capelli*, un criterio per decidere quando un sistema lineare ammette soluzioni.

Corollario 5.9 (Rouché-Capelli)

Sia $Ax = b$ un sistema lineare di m equazioni in n incognite, e sia A' la matrice completa del sistema. Allora il sistema ammette soluzioni se e solo se $\operatorname{rg} A' = \operatorname{rg} A$. *Inoltre la soluzione, se esiste, è unica se e solo se* $\operatorname{rg} A = n$.

Dimostrazione

La Proposizione 4.3 dice che la soluzione esiste se e solo se $b \in \operatorname{Span}(A^1, \dots, A^n)$, che è equivalente a richiedere

$$\operatorname{Span}(A^1,\dots,A^n) = \operatorname{Span}(A^1,\dots,A^n,b). \tag{5.5}$$

Ma il primo span è l'immagine di L_A, mentre il secondo è l'immagine di $L_{A'}$; poiché si ha sempre $\operatorname{Span}(A^1, \dots, A^n) \subseteq \operatorname{Span}(A^1, \dots, A^n, b)$, l'uguaglianza (5.5) è verificata se e solo se $\operatorname{rg} A = \operatorname{rg} A'$.

L'unicità vale se e solo se $\operatorname{Ker} L_A = \{O\}$ (Proposizione 5.1), il che accade se e solo se $\operatorname{rg} A = \dim \mathbb{R}^n = n$ (Teorema 5.7). □

È dunque necessario trovare un metodo operativo efficace per il calcolo del rango di una matrice (che non sia semplicemente recuperare un insieme massimale di colonne linearmente indipendenti). Questo metodo ci verrà fornito, nel prossimo capitolo, dalla riduzione a scala; per il momento vediamo come il rango si possa calcolare utilizzando le righe della matrice al posto delle colonne. A tale scopo ci serve una relazione molto interessante fra la trasposizione di matrici e le applicazioni φ_a dell'Esempio 5.9.

Lemma 5.10

Per ogni matrice $A \in M_{m,n}(\mathbb{K})$ e vettori $x \in \mathbb{K}^n$ e $a \in \mathbb{K}^m$ si ha

$$\varphi_{A^T a}(x) = \varphi_a(Ax).$$

Dimostrazione

Notiamo prima di tutto che quanto abbiamo scritto ha senso. L_A manda vettori di $\mathbb{K}^n$ in vettori di $\mathbb{K}^m$; quindi $Ax \in \mathbb{K}^m$ e vi possiamo applicare φ_a, che manda vettori di $\mathbb{K}^m$ in $\mathbb{K}$; quindi $\varphi_a(Ax)$ è un ben definito elemento di $\mathbb{K}$. D'altra parte, L_{A^T} manda vettori di $\mathbb{K}^m$ in vettori di $\mathbb{K}^n$; quindi φ_{A^Ta} manda vettori di $\mathbb{K}^n$ in $\mathbb{K}$, e di nuovo $\varphi_{A^Ta}(x)$ è un ben definito scalare in $\mathbb{K}$. Per dimostrare che i due sono uguali, basta calcolarli. Posto

$$A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix}, \qquad x = \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}, \qquad a = \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{vmatrix},$$

otteniamo

$$A^Ta = \begin{vmatrix} a_{11}\alpha_1 + \cdots + a_{m1}\alpha_m \\ \vdots \\ a_{1n}\alpha_1 + \cdots + a_{mn}\alpha_m \end{vmatrix}, \qquad Ax = \begin{vmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{vmatrix},$$

e

$$\begin{aligned}\varphi_{A^Ta}(x) &= (a_{11}\alpha_1 + \cdots + a_{m1}\alpha_m)x_1 + \cdots + (a_{1n}\alpha_1 + \cdots + a_{mn}\alpha_m)x_n \\ &= \alpha_1(a_{11}x_1 + \cdots + a_{1n}x_n) + \cdots + \alpha_m(a_{m1}x_1 + \cdots + a_{mn}x_n) = \varphi_a(Ax).\end{aligned}$$ □

E ora possiamo dimostrare che il rango calcolato usando le righe è uguale al rango calcolato usando le colonne.

Proposizione 5.11

Sia $A \in M_{m,n}(\mathbb{R})$. Allora:

(1) $\mathbb{R}^m = \operatorname{Im} L_A \oplus \operatorname{Ker} L_{A^T}$ *e* $\mathbb{R}^n = \operatorname{Im} L_{A^T} \oplus \operatorname{Ker} L_A$;

(2) $\operatorname{rg} A = \operatorname{rg} A^T$, *ovvero il massimo numero di colonne linearmente indipendenti coincide col massimo numero di righe linearmente indipendenti.*

Dimostrazione

Cominciamo dimostrando che $\operatorname{Im} L_A \cap \operatorname{Ker} L_{A^T} = \{O\}$. Sia $a \in \operatorname{Im} L_A \cap \operatorname{Ker} L_{A^T} \subseteq \mathbb{R}^m$; in particolare, $A^Ta = O$, ed esiste $b \in \mathbb{R}^n$ tale che $Ab = a$. Allora

$$0 = \varphi_O(b) = \varphi_{A^Ta}(b) = \varphi_a(Ab) = \varphi_a(a) = a_1^2 + \cdots + a_m^2,$$

dove abbiamo usato il Lemma 5.10; questo può accadere se e solo se $a_1 = \cdots = a_m = 0$ (in quanto $a_1, \ldots, a_m \in \mathbb{R}$), e quindi $a = O$.

Dunque $\operatorname{Im} L_A \cap \operatorname{Ker} L_{A^T} = \{O\}$. Essendo $\operatorname{Im} L_A + \operatorname{Ker} L_{A^T} \subseteq \mathbb{R}^m$, la Proposizione 4.14 e i Teoremi 4.16 e 5.7 ci dicono che

$$m \geq \dim(\operatorname{Im} L_A + \operatorname{Ker} L_{A^T}) = \dim \operatorname{Im} L_A + \dim \operatorname{Ker} L_{A^T} = \operatorname{rg} A + m - \operatorname{rg} A^T, \tag{5.6}$$

per cui $\operatorname{rg} A^T \geq \operatorname{rg} A$. Rifacendo lo stesso ragionamento a partire da A^T, e ricordandosi che $(A^T)^T = A$, otteniamo $\operatorname{rg} A^T \leq \operatorname{rg} A$; mettendo tutto insieme deduciamo (2). In particolare, (5.6) ci dice che $\dim(\operatorname{Im} L_A + \operatorname{Ker} L_{A^T}) = m$, per cui $\mathbb{R}^m = \operatorname{Im} L_A \oplus \operatorname{Ker} L_{A^T}$. L'ultima affermazione di (1) si ottiene scambiando un'altra volta A e A^T. □

Questo è l'unico risultato in tutta questa parte del testo che dipende da proprietà specifiche di $\mathbb{R}$, e non vale per un campo qualunque. La dimostrazione che abbiamo fatto funziona in $\mathbb{R}$, $\mathbb{Q}$ ed eventuali altri campi contenuti in $\mathbb{R}$; per vedere che cosa diventa questo risultato su $\mathbb{C}$ ci serve un'altra definizione.

Definizione 5.4

Sia $A \in M_{m,n}(\mathbb{C})$ una matrice a coefficienti complessi. La *matrice coniugata* di A è la matrice $\overline{A} \in M_{m,n}(\mathbb{C})$ i cui elementi sono i coniugati degli elementi di A. La *matrice trasposta coniugata* (o *aggiunta*) $A^H \in M_{n,m}(\mathbb{C})$ di A è la trasposta della coniugata di A; in simboli $A^H = \overline{A}^T = \overline{A^T}$.

Chiaramente, $(A^H)^H = A$; inoltre, si ha $\operatorname{rg} A = \operatorname{rg} \overline{A}$ (Esercizio 5.24). Ora, per ogni $a \in \mathbb{C}^m$ definiamo l'applicazione lineare $\psi_a\colon \mathbb{C}^m \to \mathbb{C}$ ponendo

$$\forall x \in \mathbb{C}^m \quad \psi_a(x) = \overline{a_1}x_1 + \cdots + \overline{a_m}x_m. \tag{5.7}$$

Ovviamente questa è una semplice variazione dell'applicazione lineare definita nell'Esempio 5.9; è infatti facile vedere che

$$\psi_a(x) = \varphi_{\overline{a}}(x)$$

per ogni a, $x \in \mathbb{C}^m$. Di conseguenza, grazie al Lemma 5.10 e a questa identificazione tra ψ_a e $\varphi_{\overline{a}}$, otteniamo che se $A \in M_{m,n}(\mathbb{C})$, $a \in \mathbb{C}^m$ e $x \in \mathbb{C}^n$ allora

$$\psi_{A^H a}(x) = \psi_a(Ax). \tag{5.8}$$

La versione complessa della Proposizione 5.11 è dunque la seguente.

Proposizione 5.12

Sia $A \in M_{m,n}(\mathbb{C})$. Allora:

(1) $\mathbb{C}^m = \operatorname{Im} L_A \oplus \operatorname{Ker} L_{A^H}$ *e* $\mathbb{C}^n = \operatorname{Im} L_{A^H} \oplus \operatorname{Ker} L_A$;

(2) $\operatorname{rg} A = \operatorname{rg} A^T = \operatorname{rg} A^H$.

Dimostrazione

Prima di tutto notiamo che (Esercizio 5.24) $\operatorname{rg} A^T = \operatorname{rg} A^H$; pertanto ci basta dimostrare che $\operatorname{Im} L_A \cap \operatorname{Ker} L_{A^H} = \{O\}$, e poi possiamo procedere esattamente come nella dimostrazione della Proposizione 5.11.

Prendiamo $a \in \operatorname{Im} L_A \cap \operatorname{Ker} L_{A^H}$, cioè tale che $A^H a = O$ ed esista $b \in \mathbb{C}^n$ con $Ab = a$. Allora (5.8) ci dice che

$$0 = \psi_O(b) = \psi_{A^H a}(b) = \psi_a(Ab) = \psi_a(a) = |a_1|^2 + \cdots + |a_m|^2.$$

Dunque abbiamo di nuovo una somma di numeri reali positivi uguale a zero; quindi necessariamente $|a_1| = \cdots = |a_m| = 0$, per cui $a = O$, e abbiamo finito. □

Osservazione 5.9

Per matrici A a coefficienti complessi, *non* è detto che l'intersezione fra l'immagine di L_A e il nucleo della trasposta sia il solo vettore nullo. Per esempio, se

$$A = \begin{vmatrix} i & 1 \\ 1 & -i \end{vmatrix} = A^T,$$

allora

$$\operatorname{Im} L_A = \operatorname{Ker} L_{A^T} = \mathbb{C}\begin{vmatrix} i \\ 1 \end{vmatrix}.$$

Esercizi 5

5.1 Dimostra di nuovo la Proposizione 5.1 utilizzando L_A.

5.2 Sia V lo spazio vettoriale dell'Esempio 4.12. Definite T, S_1, S_2: $V \to V$ con

$$T\begin{vmatrix} x \\ y \\ 0 \end{vmatrix} = \begin{vmatrix} 2x-y \\ x \\ 0 \end{vmatrix}, \quad S_1\begin{vmatrix} x \\ y \\ 0 \end{vmatrix} = \begin{vmatrix} x^2 \\ 2y \\ 0 \end{vmatrix}, \quad S_2\begin{vmatrix} x \\ y \\ 0 \end{vmatrix} = \begin{vmatrix} x+1 \\ y \\ 0 \end{vmatrix},$$

dimostra che T è lineare mentre S_1 ed S_2 non lo sono.

5.3 Al variare di $k \in \mathbb{R}$, considera l'applicazione $T_k : \mathbb{R}^3 \to \mathbb{R}^3$ data da

$$T_k(x) = \begin{vmatrix} kx_1 + x_2 \\ (k+2)e^{x_2} \\ x_2 - x_3 \end{vmatrix}.$$

Determina per quali k l'applicazione assegnata sia lineare.

5.4 Al variare di $k \in \mathbb{R}$, considera l'applicazione $T_k : \mathbb{R}^2 \to \mathbb{R}^3$ data da

$$T_k(x) = \begin{vmatrix} x_1 + x_2 \\ (k-1)\sin(x_1) + (k^2-1)x_2 \\ 2x_1 + 2x_2 \end{vmatrix}.$$

Determina per quali k l'applicazione assegnata sia lineare e per tali valori trovane il nucleo.

5.5 Considera l'applicazione $T : \mathbb{R}^3 \to \mathbb{R}^2$ data da $T(x) = \begin{vmatrix} x_1 - x_3 \\ x_2 - 2x_3 \end{vmatrix}$. Verifica che T è lineare e trovane nucleo e immagine.

5.6 Sia T: $\mathbb{R}^n \to \mathbb{R}^m$ (con $n > m$) data da

$$T\begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix} = \begin{vmatrix} x_1 \\ \vdots \\ x_m \end{vmatrix}.$$

Dimostra che T è lineare e trova una matrice $A \in M_{m,n}(\mathbb{R})$ tale che $T = L_A$.

5.7 Sia $T : \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione data da $T(x) = \begin{vmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_1 - 2x_2 \end{vmatrix}$. Trova $A \in M_{3,2}(\mathbb{R})$ tale che $T = L_A$ e verifica, usando il nucleo, che T è iniettiva.

5.8 Considera l'applicazione $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ data da $T(x) = \begin{vmatrix} x_1 + x_2 + x_3 \\ 5x_2 + x_3 \\ 3x_3 \end{vmatrix}$. Trova $A \in M_{3,3}(\mathbb{R})$ tale che $T = L_A$ e verifica che T è iniettiva.

5.9 Sia $T\colon V \to W$ un'applicazione fra due spazi vettoriali, e fissiamo una base $\mathcal{B}$ di V e una base $\mathcal{C}$ di W. Dimostra che se le coordinate (rispetto a $\mathcal{C}$) di $T(v)$ si ottengono dalle coordinate di v (rispetto a $\mathcal{B}$) tramite polinomi di primo grado senza termini noti allora l'applicazione T è lineare.

5.10 La *traccia* $\operatorname{tr} A$ di una matrice quadrata $A \in M_{n,n}(\mathbb{R})$ è per definizione la somma degli elementi sulla diagonale principale: $\operatorname{tr} A = a_{11} + \cdots + a_{nn}$. Dimostra che $\operatorname{tr}\colon M_{n,n}(\mathbb{R}) \to \mathbb{R}$ è lineare.

5.11 Considera l'applicazione $T\colon M_{2,2}(\mathbb{R}) \to \mathbb{R}$ data da $T\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a - 2b + c - d$. Dimostra che T è lineare, calcola la dimensione del nucleo e verifica che $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$, $\begin{vmatrix} -1 & 0 \\ 1 & 0 \end{vmatrix}$ e $\begin{vmatrix} 2 & 1 \\ 0 & 0 \end{vmatrix}$ formano una base di $\operatorname{Ker} T$.

5.12 Considera l'applicazione $T\colon \mathbb{R}_2[t] \to \mathbb{R}^2$ data da $T(p(t)) = \begin{vmatrix} p(1) \\ p(0) \end{vmatrix}$. Verifica che T è lineare e trovane il nucleo.

5.13 Sia $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'unica applicazione lineare tale che

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

e per ogni $a \in \mathbb{R}$ sia $S_a\colon \mathbb{R}^2 \to \mathbb{R}^3$ l'unica applicazione lineare tale che

$$S_a\begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} 6 \\ 4 \\ 2 \end{vmatrix} \quad \text{e} \quad S_a\begin{vmatrix} 2 \\ -1 \end{vmatrix} = \begin{vmatrix} a \\ 0 \\ 4 \end{vmatrix}.$$

Trova per quali $a \in \mathbb{R}$ si ha $\operatorname{Im} T = \operatorname{Im} S_a$ e calcola la dimensione di $\operatorname{Im} T \cap \operatorname{Im} S_a$ al variare di $a \in \mathbb{R}$.

5.14 Sia $T\colon \mathbb{R}^n \to \mathbb{R}$ un'applicazione lineare. Dimostra che esiste $a \in \mathbb{R}^n$ tale che $T \equiv \varphi_a$, dove $\varphi_a\colon \mathbb{R}^n \to \mathbb{R}$ è l'applicazione dell'Esempio 5.9. (*Suggerimento:* prendi $a_j = T(e_j)$ per $j = 1, \ldots, n$, dove $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{R}^n$.)

5.15 Una matrice quadrata $A \in M_{n,n}(\mathbb{R})$ è detta *simmetrica* (rispettivamente, *antisimmetrica*) se $A^T = A$ (rispettivamente, $A^T = -A$). Indichiamo con $\mathcal{S}_n(\mathbb{R})$ (rispettivamente, $\mathcal{A}_n(\mathbb{R})$) l'insieme delle matrici simmetriche (rispettivamente, antisimmetriche) di ordine n. Dimostra che i sottoinsiemi $\mathcal{S}_n(\mathbb{R})$ e $\mathcal{A}_n(\mathbb{R})$ sono sottospazi di $M_{n,n}(\mathbb{R})$ e calcolane la dimensione. (*Suggerimento:* può esserti utile l'Esercizio 3.28.)

5.16 Dimostra che $M_{n,n}(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$. (*Suggerimento:* prova a usare $(A + A^T)/2$ e $(A - A^T)/2$.)

5.17 Sia $T\colon V \to W$ un'applicazione lineare, $V_1 \subseteq V$ un sottospazio di V e $W_1 \subseteq W$ un sottospazio di W. Dimostra che:

1) $T^{-1}(W_1) = \{v \in V \mid T(v) \in W_1\} \subseteq V$ è un sottospazio di V;
2) $T(V_1) = \{T(v) \mid v \in V_1\} \subseteq W$ è un sottospazio di W.

5.18 Dimostra che il sistema lineare $Ax = b$ è compatibile se e solo se $b \in \operatorname{Im} L_A$, e che la soluzione, se esiste, è unica se e solo se $\operatorname{Ker} L_A = \{O\}$.

5.19 In ciascuno dei tre casi seguenti, scopri se è possibile costruire applicazioni lineari che soddisfano le condizioni indicate, e in caso ne esistano più di una, trovane almeno due distinte:

1) $T\colon \mathbb{R}^4 \to \mathbb{R}^3$ surgettiva e tale che $\operatorname{Ker} T = \mathbb{R}\begin{vmatrix} 1 \\ 0 \\ 1 \\ 0 \end{vmatrix}$;

2) $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che $\operatorname{Im} T = \mathbb{R}\begin{vmatrix} 1 \\ 1 \end{vmatrix}$;

3) $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ iniettiva e tale che $\operatorname{Im} T = \operatorname{Span}\left(\begin{vmatrix} 1 \\ 1 \\ 1 \end{vmatrix}, \begin{vmatrix} -1 \\ 2 \\ 0 \end{vmatrix}\right)$.

5.20 Sia $T\colon V \to W$ un'applicazione lineare. Dimostra che:

1) se $\dim V > \dim W$ l'applicazione T non può essere iniettiva;
2) se $\dim V < \dim W$ l'applicazione T non può essere surgettiva.

5.21 Sia $T\colon V \to W$ un'applicazione lineare. Dimostra che se $v_1, \dots, v_k \in V$ sono linearmente dipendenti, anche $T(v_1), \dots, T(v_k) \in W$ lo sono. Dimostra pure che se $T(v_1), \dots, T(v_k) \in W$ sono linearmente indipendenti, anche $v_1, \dots, v_k \in V$ lo sono.

5.22 Sia $T\colon V \to W$ un'applicazione lineare *iniettiva*. Dimostra che se i vettori $v_1, \dots, v_k \in V$ sono linearmente indipendenti anche $T(v_1), \dots, T(v_k) \in W$ lo sono.

5.23 Trova un'applicazione lineare $T\colon V \to W$ e dei vettori $v_1, v_2 \in V$ tali che v_1 e v_2 siano linearmente indipendenti e $T(v_1)$ e $T(v_2)$ non lo siano.

5.24 Sia $A \in M_{m,n}(\mathbb{C})$. Dimostra che $\operatorname{rg} \overline{A} = \operatorname{rg} A$.

5.25 Trova rango, nucleo e immagine delle seguenti matrici a coefficienti complessi.

$$\begin{vmatrix} 1 & i & -1 \\ 1 & -i & -1 \\ 1 & 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & i & 0 \\ i & -1 & 1 \\ 2 & 2i & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & i & 1 \\ i & -1 & i \\ 1 & i & 1 \end{vmatrix}.$$

5.26 Considera l'applicazione $T : \mathbb{C}^3 \to \mathbb{C}^2$ data da $T(z) = \begin{vmatrix} z_1 - iz_3 \\ z_2 + 3iz_3 \end{vmatrix}$. Verifica che T è lineare e trovane nucleo e immagine.

5.27 Considera l'applicazione $T : \mathbb{C}^3 \to \mathbb{C}^3$ data da $T(z) = \begin{vmatrix} 4z_1 + iz_2 - 2z_3 \\ iz_1 + z_3 \\ iz_1 - iz_3 \end{vmatrix}$. Trova $A \in M_{3,3}(\mathbb{C})$ tale che $T = L_A$ e verifica che T è iniettiva.

5.28 Considera l'applicazione $T : M_{2,2}(\mathbb{C}) \to \mathbb{C}^2$ data da $T\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a - ic \\ b + 2id \end{vmatrix}$. Prova che T è lineare, calcola la dimensione del nucleo e verifica che $\begin{vmatrix} 0 & -2i \\ 0 & 1 \end{vmatrix}$ e $\begin{vmatrix} i & 0 \\ 1 & 0 \end{vmatrix}$ formano una base di $\operatorname{Ker} T$.

5.29 Dimostra che l'insieme $M_{n,n}(\mathbb{C})$ delle matrici quadrate a coefficienti complessi ha dimensione n^2 come spazio vettoriale su $\mathbb{C}$, ma dimensione $2n^2$ se considerato come spazio vettoriale su $\mathbb{R}$. Dimostra che l'insieme $\{A \in M_{n,n}(\mathbb{C}) \mid A^H = A\}$ è uno spazio vettoriale su $\mathbb{R}$ di dimensione n^2.

5.30 Dimostra che il coniugio è lineare se consideriamo $\mathbb{C}$ come spazio vettoriale su $\mathbb{R}$, mentre non lo è se consideriamo $\mathbb{C}$ come spazio vettoriale su se stesso.

5.31 Sia $T: \mathbb{C}^n \to \mathbb{C}$ un'applicazione lineare. Dimostra che esiste $a \in \mathbb{C}^n$ tale che $T \equiv \psi_a$, dove $\psi_a: \mathbb{C}^n \to \mathbb{C}$ è l'applicazione definita in (5.7). (*Suggerimento:* prendi $a_j = \overline{T(e_j)}$ per $j = 1, \dots, n$, dove $\{e_1, \dots, e_n\}$ è la base canonica di $\mathbb{C}^n$.)

5.32 Sia $T: V \to W$ un'applicazione fra spazi vettoriali. Dimostra che T è lineare se e solo se il grafico $\Gamma = \{(v, T(v)) \in V \times W \mid v \in V\}$ è un sottospazio vettoriale di $V \times W$ (vedi l'Esercizio 4.26).

Sistemi lineari

6

Sommario

È giunto il momento di affrontare un certo numero di problemi riguardanti tecniche di calcolo: come si trova il rango di una matrice? Come si risolve un sistema lineare non quadrato? Come si trovano dimensione e base di un sottospazio generato da certi vettori? Come si trova dimensione e base del nucleo di un'applicazione lineare? Come si estrae una base da un sistema di generatori? Come si trovano dimensione e base dell'unione e dell'intersezione di due sottospazi? Si tratta di problemi tutti collegati fra loro, e che possono essere risolti con l'ausilio di un unico strumento: la riduzione a scala, una versione dell'eliminazione di Gauss per matrici non quadrate. In questo capitolo risolveremo questi problemi in $\mathbb{R}^n$; in $\mathbb{C}^n$ si opera in maniera assolutamente identica, e vedremo nel Capitolo 8 come procedere in spazi vettoriali qualunque (ma non sarà molto diverso). Infine, parleremo di equazioni parametriche e cartesiane, e definiremo ufficialmente i sottospazi affini.

6.1 Sistemi a scala

Cominciamo col vedere che cosa rimpiazza le matrici triangolari superiori nel caso non quadrato.

Definizione 6.1

Una *matrice a scala* è una matrice $m \times n$ siffatta:

$$\begin{vmatrix} 0 & \cdots & 0 & p_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & \vdots & 0 & 0 & \cdots & 0 & p_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & 0 & \cdots & 0 & p_3 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \ddots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & p_r & * \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & 0 & 0 \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \cdots\cdots & 0 & 0 & 0 \cdots & 0 \end{vmatrix}$$

dove un $*$ indica che quell'elemento può essere qualsiasi. I numeri $p_1, \ldots, p_r \in \mathbb{R}^*$, tutti non nulli, sono i *pivot* della matrice a scala. Un *sistema a scala* è un sistema lineare la cui matrice dei coefficienti è una matrice a scala.

Osservazione 6.1

I pivot di una matrice quadrata triangolare superiore descritti nel Paragrafo 3.3 coincidono con quelli definiti qui solo per matrici non singolari. In questo capitolo gli unici pivot che useremo sono quelli che abbiamo appena definito, e nel seguito sarà sempre chiaro di quali pivot si sta parlando.

Le proprietà principali di una matrice a scala sono contenute nel seguente lemma.

Lemma 6.1

Sia $S \in M_{m,n}(\mathbb{R})$ una matrice a scala con r pivot, e poniamo

$$V_r = \left\{ \begin{vmatrix} b_1 \\ \vdots \\ b_r \\ 0 \\ \vdots \\ 0 \end{vmatrix} \in \mathbb{R}^m \;\middle|\; b_1, \ldots, b_r \in \mathbb{R} \right\} = \mathrm{Span}(e_1, \ldots, e_r) \subset \mathbb{R}^m.$$

Per $k = 1, \ldots, r$, indichiamo con S^{j_k} la colonna della matrice S in cui compare il k-esimo pivot p_k. Allora $\mathrm{Im}\, S = V_r$, $\mathrm{rg}\, S = r$ e $\{S^{j_1}, \ldots, S^{j_r}\}$ è una base di $\mathrm{Im}\, S$.

Dimostrazione

Siccome tutte le colonne di S appartengono a V_r e generano $\mathrm{Im}\, S$, è chiaro che $\mathrm{Im}\, S \subseteq V_r$. È pure evidente che $\dim V_r = r$; quindi per concludere basta far vedere (perché?) che $S^{j_1}, \ldots, S^{j_r}$ sono linearmente indipendenti.

Consideriamo il sistema omogeneo $\alpha_1 S^{j_1} + \cdots + \alpha_r S^{j_r} = O$, di incognite $\alpha_1, \ldots, \alpha_r$. Questo sistema ha come matrice dei coefficienti una matrice della forma

$$\begin{vmatrix} p_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & p_r \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{vmatrix}.$$

Quindi siamo ricondotti a un sistema omogeneo quadrato triangolare superiore con elementi non nulli sulla diagonale principale; per la Proposizione 3.1, l'unica soluzione possibile è $\alpha_1 = \ldots = \alpha_r = 0$. □

Corollario 6.2

Sia $S \in M_{m,n}(\mathbb{R})$ una matrice a scala di rango r. Allora il sistema $Sx = c$ ha soluzione se e solo se le ultime $m - r$ coordinate di c sono zero, e lo spazio delle soluzioni del sistema omogeneo $Sx = O$ ha dimensione $n - r$.

Dimostrazione

Il sistema $Sx = c$ ha soluzione se e solo se $c \in \operatorname{Im} S = V_r$. Lo spazio delle soluzioni di $Sx = O$ è $\operatorname{Ker} S$ che ha dimensione $n - \operatorname{rg} S$ (vedi l'Osservazione 5.8). □

Esempio 6.1

Consideriamo il sistema

$$\begin{cases} x_2 - x_3 + 3x_4 - x_5 + \frac{1}{2}x_6 = 1, \\ 2x_4 - x_6 = 0, \\ x_5 + 4x_6 = 1. \end{cases}$$

Si tratta di un sistema a scala $Sx = c$, dove

$$S = \begin{vmatrix} 0 & 1 & -1 & 3 & -1 & 1/2 \\ 0 & 0 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{vmatrix} \quad \text{e} \quad c = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}.$$

La matrice $S \in M_{3,6}(\mathbb{R})$ ha tre pivot: $p_1 = 1$, $p_2 = 2$ e $p_3 = 1$ nelle colonne 2, 4 e 5. Dalla terza equazione ricaviamo x_5 in funzione di x_6: si ha $x_5 = 1 - 4x_6$. La seconda equazione ci fornisce x_4 in funzione di x_6: abbiamo $x_4 = \frac{1}{2}x_6$. Sostituendo nella prima equazione ricaviamo anche x_2 in funzione di x_3 e x_6: $x_2 = x_3 - 6x_6 + 2$. Infine, x_1 è libera, in quanto il sistema non impone alcuna condizione su di essa. Pertanto le soluzioni del nostro sistema dipendono da 3 parametri liberi, e sono

$$\begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{vmatrix} = \begin{vmatrix} x_1 \\ x_3 - 6x_6 + 2 \\ x_3 \\ x_6/2 \\ 1 - 4x_6 \\ x_6 \end{vmatrix} = \begin{vmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{vmatrix} + x_1 \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} + x_3 \begin{vmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{vmatrix} + x_6 \begin{vmatrix} 0 \\ -6 \\ 0 \\ 1/2 \\ -4 \\ 1 \end{vmatrix}. \tag{6.1}$$

Il metodo che si utilizza (e che abbiamo visto in azione nell'esempio precedente) per risolvere i sistemi a scala si chiama *risoluzione all'indietro*: ne diamo ora una descrizione generale. Sia $Sx = c$ un sistema a scala, dove S è una matrice con r pivot sulle colonne $j_1, \ldots, j_r$ (e sulle righe $1, \ldots, r$, ovviamente), e $c \in V_r$, in modo che il sistema sia compatibile. Essendo $p_r \neq 0$ per definizione, dalla r-esima equazione possiamo ricavare x_{j_r} in funzione delle variabili successive. Sostituendo nell'equazione precedente ricaviamo $x_{j_{r-1}}$ in funzione delle variabili successive (x_{j_r} esclusa). Procediamo così fino a giungere alla prima equazione, da cui ricaviamo x_{j_1} in funzione delle variabili successive ($x_{j_2}, \ldots, x_{j_r}$ escluse).

La soluzione è allora la somma di un vettore costante e del sottospazio (di dimensione $n - r$) delle soluzioni del sistema omogeneo associato $Sx = O$, dipendente da $n - r$ parametri, le variabili che non corrispondono ai pivot, le cosiddette *variabili libere*. Le altre variabili, quelle corrispondenti ai pivot e che si esprimono in funzione delle variabili libere, si chiamano *variabili dipendenti*, e sono $r = \operatorname{rg} S$.

Osservazione 6.2

La risoluzione all'indietro fornisce sempre le soluzioni del sistema nella forma $v = v^o + x_{i_1} w_1 + \cdots + x_{i_{n-r}} w_{n-r}$, dove $x_{i_1}, \ldots, x_{i_{n-r}}$ sono le variabili libere, $v^o \in \mathbb{R}^n$ è una soluzione particolare, e $w_1, \ldots, w_{n-r} \in \mathbb{R}^n$ sono un sistema di generatori di $\operatorname{Ker} S$. In realtà, i w_j sono proprio una base di $\operatorname{Ker} S$ (perché?).

6.2 La riduzione a scala

La *riduzione a scala* è un algoritmo (simile all'eliminazione di Gauss) che trasforma una matrice $A \in M_{m,n}(\mathbb{R})$ in una matrice a scala $S \in M_{m,n}(\mathbb{R})$ in modo da trasformare qualunque sistema $Ax = b$ in un sistema a scala $Sx = c$ a esso equivalente, chiamato *riduzione a scala* di $Ax = b$. La riduzione a scala procede per passi successivi. Il passo i-esimo opera in maniera tale da: trasformare la riga i-esima della matrice dei coefficienti nella forma adatta a una matrice a scala; produrre un numero reale non nullo $p_i \in \mathbb{R}^*$, l'i-esimo *pivot* della matrice, situato nella colonna j_i; annullare tutti gli elementi della matrice sotto la i-esima riga nelle colonne dalla prima alla j_i-esima inclusa – e tutto senza cambiare le soluzioni del sistema.

Descriviamo ora in dettaglio questo algoritmo. Partiamo da un sistema lineare $Ax = b$ con $A \in M_{m,n}(\mathbb{R})$ e $b \in \mathbb{R}^m$. Se A è la matrice nulla, abbiamo finito. Altrimenti, sia A^{j_1} la prima colonna non nulla di A; scambiando se necessario la prima riga con una sottostante possiamo supporre che sia il primo elemento a_{1j_1} della colonna A^{j_1} a essere diverso da zero; poniamo $p_1 = a_{1j_1}$. Sommiamo alla riga h-esima (con $h = 2, \ldots, m$) un adeguato multiplo della prima in modo da annullare tutti gli elementi della j_1-esima colonna tranne il primo. La matrice diventa

$$\begin{vmatrix} 0 & \cdots & 0 & \multicolumn{1}{|l}{\underline{p_1}} & * & \cdots & * \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{vmatrix},$$

con $p_1 \neq 0$. I passi successivi sono molto simili. Supponiamo di aver già trattato le prime $i - 1$ righe; se le righe dalla i-esima in poi sono tutte nulle (oppure se la matrice ha solo $i - 1$ righe) abbiamo finito. Altrimenti, sia j_i l'indice della prima colonna che contiene un elemento non nullo in una riga sotto la $(i-1)$-esima. Scambiando se necessario la riga i-esima con una sottostante, possiamo supporre che l'elemento diverso da zero della colonna A^{j_i} sia sulla riga i-esima; poniamo p_i uguale a questo elemento. Sommiamo alla riga h-esima (per $h = i+1, \ldots, m$) un adeguato multiplo della i-esima riga in modo da annullare tutti gli elementi della colonna A^{j_i} sotto la riga i-esima. In questo modo arriviamo a una matrice della forma

$$
\left|\begin{array}{ccccccccccccccc}
0 & \cdots & 0 & p_1 & * & \cdots & * & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & \vdots & 0 & \cdots & \cdots & 0 & p_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & \vdots & \vdots & & & \vdots & 0 & 0 & & & & & & \vdots \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & \vdots & & \vdots & & & & \vdots \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & \vdots & & 0 & p_i & * & \cdots & * \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & \vdots & & 0 & 0 & * & \cdots & * \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots & *
\end{array}\right|
$$

con $p_1, \ldots, p_i \neq 0$. Procedendo in questo modo arriviamo prima o poi a una matrice in cui le ultime righe sono tutte nulle, oppure in cui l'ultimo pivot appartiene all'ultima riga; in entrambi i casi abbiamo ridotto a scala la matrice di partenza.

Ma vediamo come funziona la riduzione a scala su un esempio.

Esempio 6.2

Consideriamo il sistema

$$
\begin{cases}
2x_1 - x_2 + 4x_3 + x_4 = -2, \\
-2x_1 + x_2 - 7x_3 + x_4 = -1, \\
4x_1 - 2x_2 + 5x_3 + 4x_4 = -7,
\end{cases}
$$

con matrice A e vettore dei termini noti b dati da

$$
A = \begin{vmatrix} 2 & -1 & 4 & 1 \\ -2 & 1 & -7 & 1 \\ 4 & -2 & 5 & 4 \end{vmatrix}, \qquad b = \begin{vmatrix} -2 \\ -1 \\ -7 \end{vmatrix}.
$$

La matrice A non è nulla, e il primo elemento della prima colonna è diverso da zero, per cui $j_1 = 1$ e $p_1 = 2$. Per concludere il primo passo dobbiamo sommare la prima equazione alla seconda, e sottrarre il doppio della prima alla terza, ottenendo

$$
\left|\begin{array}{cccc|c} 2 & -1 & 4 & 1 & -2 \\ -2 & 1 & -7 & 1 & -1 \\ 4 & -2 & 5 & 4 & -7 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 2 & -1 & 4 & 1 & -2 \\ 0 & 0 & -3 & 2 & -3 \\ 0 & 0 & -3 & 2 & -3 \end{array}\right|.
$$

Passiamo al passo 2; siccome la seconda colonna ha tutti zero dalla seconda riga in giù, mentre la terza colonna ha un termine non nullo nella seconda riga, troviamo $j_2 = 3$ e $p_2 = -3$. Sottraendo la seconda riga alla terza otteniamo

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

Troviamo una matrice a scala S con due pivot, e il vettore c dei termini noti sta in V_2 (cioè ha nulli gli elementi dal terzo in poi); il Corollario 6.2 implica che $Sx = c$ (e quindi il sistema iniziale) è compatibile. Risolvendo all'indietro, troviamo

$$\begin{cases} x_1 = \frac{1}{2}x_2 - \frac{11}{6}x_4 - 3, \\ x_3 = \frac{2}{3}x_4 + 1, \end{cases}$$

cioè

$$\left|\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array}\right| = \left|\begin{array}{c} -3 \\ 0 \\ 1 \\ 0 \end{array}\right| + x_2 \left|\begin{array}{c} 1/2 \\ 1 \\ 0 \\ 0 \end{array}\right| + x_4 \left|\begin{array}{c} -11/6 \\ 0 \\ 2/3 \\ 1 \end{array}\right|,$$

con due variabili dipendenti (x_1 e x_3, corrispondenti alle colonne dei pivot) e due variabili libere (x_2 e x_4), coerentemente con quanto ci aspettavamo.

Esempio 6.3

Consideriamo invece il sistema

$$\begin{cases} 2x_1 - x_2 + 4x_3 + x_4 = -2, \\ -2x_1 + x_2 - 7x_3 + x_4 = -1, \\ 4x_1 - 2x_2 + 5x_3 + 4x_4 = 7, \end{cases}$$

con matrice A e vettore dei termini noti b dati da

$$A = \left|\begin{array}{cccc} 2 & -1 & 4 & 1 \\ -2 & 1 & -7 & 1 \\ 4 & -2 & 5 & 4 \end{array}\right|, \qquad b = \left|\begin{array}{c} -2 \\ -1 \\ 7 \end{array}\right|.$$

In questo caso la riduzione a scala ci dà

$$\left|\begin{array}{cccc|c} 2 & -1 & 4 & 1 & -2 \\ -2 & 1 & -7 & 1 & -1 \\ 4 & -2 & 5 & 4 & 7 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 2 & -1 & 4 & 1 & -2 \\ 0 & 0 & -3 & 2 & -3 \\ 0 & 0 & -3 & 2 & 11 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 2 & -1 & 4 & 1 & -2 \\ 0 & 0 & -3 & 2 & -3 \\ 0 & 0 & 0 & 0 & 14 \end{array}\right|$$

Il vettore c *non* appartiene a V_2, per cui il sistema originale non è compatibile.

Riassumiamo i legami tra un sistema e una sua riduzione a scala.

Teorema 6.3

Sia $Ax = b$ un sistema lineare, e $Sx = c$ una sua riduzione a scala. Allora:

(1) *lo spazio delle soluzioni di $Ax = b$ coincide con quello delle soluzioni di $Sx = c$;*

(2) $\operatorname{Ker} A = \operatorname{Ker} S$;

(3) $\operatorname{rg} A = \operatorname{rg} S$ *(ma in generale $\operatorname{Im} A \neq \operatorname{Im} S$);*

(4) *siano $S^{j_1}, \dots, S^{j_r}$, dove $r = \operatorname{rg} S$, le colonne corrispondenti ai pivot di S; allora $\{A^{j_1}, \dots, A^{j_r}\}$ è una base di $\operatorname{Im} A$.*

Dimostrazione

(1) Infatti i due sistemi sono equivalenti.

(2) Il sistema $Ax = O$ è equivalente al sistema $Sx = O$.

(3) Sia n il numero di colonne di A (e S). Allora il Teorema 5.7 ci dice che

$$\operatorname{rg} A = n - \dim \operatorname{Ker} A = n - \dim \operatorname{Ker} S = \operatorname{rg} S.$$

(4) Poiché si ha $\operatorname{rg} A = \operatorname{rg} S$, è sufficiente dimostrare (perché?) che le colonne $\{A^{j_1}, \ldots, A^{j_r}\}$ sono indipendenti, cioè che il sistema $x_1 A^{j_1} + \cdots + x_r A^{j_r} = O$ ha solo la soluzione $x_1 = \ldots = x_r = 0$. Applicando la stessa riduzione a scala che ha trasformato A in S questo nuovo sistema si riduce al sistema a scala equivalente $x_1 S^{j_1} + \cdots + x_r S^{j_r} = O$, che nella dimostrazione del Lemma 6.1 abbiamo visto avere solo la soluzione nulla. □

Osservazione 6.3

I pivot di una matrice non sono univocamente determinati, dato che possiamo scegliere quali righe scambiare. Invece, il numero dei pivot ottenuti *non* dipende dalle scelte fatte. Infatti, il numero dei pivot di una matrice a scala è uguale al suo rango; siccome, grazie al Teorema 6.3.(3), il rango di una riduzione a scala S di una matrice A è uguale al rango di A, il numero dei pivot di una qualunque riduzione a scala di A è esattamente uguale a $\operatorname{rg} A$.

Osservazione 6.4

Anche se il rango di A è uguale a quello di una sua riduzione a scala S, l'immagine di L_A è in generale diversa dall'immagine di L_S (controllalo sull'Esempio 6.2). Questo è dovuto al fatto che la riduzione a scala opera sulle righe della matrice, mentre l'immagine dipende dalle colonne.

6.3 Tecniche di calcolo

Adesso abbiamo tutti gli strumenti necessari per la risoluzione dei nostri problemi.

a) *Per risolvere il sistema $Ax = b$*: basta applicare la riduzione a scala al sistema per trasformarlo in un sistema a scala equivalente, la cui compatibilità è subito vista col Corollario 6.2, e che viene risolto con la risoluzione all'indietro.

b) *Per trovare il rango e una base dell'immagine di una matrice A*: basta applicare la riduzione a scala per trasformarla in una matrice a scala dello stesso rango, e utilizzare il Teorema 6.3 e il Lemma 6.1.

c) *Per trovare dimensione e base del nucleo di un'applicazione lineare L_A*: basta applicare la riduzione a scala ad A per trasformarla in una matrice a scala S e poi trovare $\operatorname{Ker} S$; questo ci fornisce dimensione e base di $\operatorname{Ker} A$ grazie al Teorema 6.3 e all'Osservazione 6.2.

d) *Per trovare dimensione e base del sottospazio generato da $v_1, \ldots, v_k \in \mathbb{R}^n$*: consideriamo la matrice $A = \left| v_1 \ \cdots \ v_k \right| \in M_{n,k}(\mathbb{R})$, in cui abbiamo messo i vettori v_j per colonna. Siccome $\operatorname{Span}(v_1, \ldots, v_k) = \operatorname{Im} A$, la dimensione dello span è il rango di A, e una base la si trova applicando il Teorema 6.3.(4).

e) *Per completare un insieme $\{v_1, \ldots, v_k\}$ di vettori linearmente indipendenti di $\mathbb{R}^n$ a una base*: basta estrarre una base dall'insieme $\{v_1, \ldots, v_k, e_1, \ldots, e_n\}$ col metodo del punto *d*), dove $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{R}^n$. Infatti la riduzione a scala è fatta in modo tale da assicurare (perché?) la presenza dei vettori $v_1, \ldots, v_k$ nella base estratta.

f) *Per trovare dimensione e base di $U + W$, dove U e W sono sottospazi di $\mathbb{R}^n$*: sia $\mathcal{B} \subset \mathbb{R}^n$ una base di U, e $\mathcal{C} \subset \mathbb{R}^n$ una base di W; allora $U + W$ è generato da $\mathcal{B} \cup \mathcal{C}$, per cui basta applicare il punto *d*) a $\mathcal{B} \cup \mathcal{C}$.

g) *Per trovare dimensione e base di $U \cap W$, dove U e W sono sottospazi di $\mathbb{R}^n$*: sia $\{u_1, \ldots, u_r\} \subset \mathbb{R}^n$ una base di U, $\{w_1, \ldots, w_s\} \subset \mathbb{R}^n$ una base di W, e consideriamo la matrice[1] $A = |u_1 \cdots u_r\ w_1 \cdots w_s| \in M_{n,r+s}(\mathbb{R})$. Sappiamo (perché?) che si ha $U + W = \operatorname{Im} A$; dunque i Teoremi di Grassmann e della dimensione implicano

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) = r + s - \operatorname{rg} A = \dim \operatorname{Ker} A.$$

Allora per calcolare la dimensione dell'intersezione basta applicare il punto *c*) alla matrice A; per trovare una base di $U \cap W$ bisogna lavorare ancora. Un vettore $v \in \mathbb{R}^n$ appartiene all'intersezione $U \cap W$ se e solo se esistono (unici) $x_1, \ldots, x_r, y_1, \ldots, y_s \in \mathbb{R}$ tali che

$$x_1 u_1 + \cdots + x_r u_r = v = y_1 w_1 + \cdots + y_s w_s, \tag{6.2}$$

e quindi se e solo se

$$\begin{vmatrix} -x_1 \\ \vdots \\ -x_r \\ y_1 \\ \vdots \\ y_s \end{vmatrix} \in \operatorname{Ker} A. \tag{6.3}$$

In altre parole, indicando con $\psi(v)$ il vettore (6.3) definiamo un'applicazione $\psi\colon U \cap W \to \operatorname{Ker} A$ lineare, iniettiva (esercizio) e quindi surgettiva, per il Corollario 5.8.(3). L'applicazione inversa $\psi^{-1}\colon \operatorname{Ker} A \to U \cap W$ è data da

$$\psi^{-1}\begin{vmatrix} -x_1 \\ \vdots \\ -x_r \\ y_1 \\ \vdots \\ y_s \end{vmatrix} = y_1 w_1 + \cdots + y_s w_s, \tag{6.4}$$

grazie a (6.2) e (6.3). Dunque se troviamo[2] col punto *c*) una base $\{z_1, \ldots, z_p\}$ di $\operatorname{Ker} A$, allora $\{\psi^{-1}(z_1), \ldots, \psi^{-1}(z_p)\}$ sarà una base di $U \cap W$ (in quanto è un sistema di generatori per il Lemma 5.6, e quindi una base per l'Esercizio 5.22).

Vediamo ora questi metodi in azione su alcuni esempi.

Esempio 6.4

Vogliamo dimensione e base di $\operatorname{Ker} L_A$ e $\operatorname{Im} L_A$, dove

$$A = \begin{vmatrix} 1 & 0 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \\ -1 & -1 & 2 \end{vmatrix} \in M_{4,3}(\mathbb{R}).$$

---

[1] Se $r < s$, conviene invece mettere prima i w_j e poi gli u_i per ridurre i conti che dovremo fare in seguito.

[2] Non è necessario calcolare completamente i vettori della base di $\operatorname{Ker} A$; bastano le loro ultime s coordinate, le uniche che servono per scrivere ψ^{-1}. Ciò riduce i conti necessari, come vedremo nell'Esempio 6.6.

Dobbiamo prima di tutto ridurre a scala la matrice A

$$\left|\begin{matrix} 1 & 0 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 4 \\ -1 & -1 & 2 \end{matrix}\right| \longrightarrow \left|\begin{matrix} 1 & 0 & 3 \\ 0 & 1 & -5 \\ 0 & 1 & -5 \\ 0 & -1 & 5 \end{matrix}\right| \longrightarrow \left|\begin{matrix} 1 & 0 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}\right| = S.$$

Dunque abbiamo trovato due pivot; questo vuol dire che $\operatorname{rg} A = \dim \operatorname{Im} L_A = 2$ e $\dim \operatorname{Ker} L_A = 1$. Inoltre sappiamo anche che una base di $\operatorname{Im} A$ è data da $\{A^1, A^2\}$. Per trovare una base del nucleo, dobbiamo semplicemente risolvere all'indietro il sistema $Sx = O$; troviamo $x_1 = -3x_3$ e $x_2 = 5x_3$, per cui

$$\operatorname{Ker} L_A = \mathbb{R}\left|\begin{matrix} -3 \\ 5 \\ 1 \end{matrix}\right|.$$

$\square$

Esempio 6.5

Vogliamo dimensione e base di $U = \operatorname{Span}(u_1, u_2, u_3) \subset \mathbb{R}^4$, dove

$$u_1 = \left|\begin{matrix} 1 \\ -1 \\ 3 \\ 2 \end{matrix}\right|, \qquad u_2 = \left|\begin{matrix} -2 \\ 1 \\ 0 \\ 1 \end{matrix}\right|, \qquad u_3 = \left|\begin{matrix} 8 \\ -5 \\ 6 \\ 1 \end{matrix}\right|.$$

La matrice A è data da

$$A = \left|\begin{matrix} 1 & -2 & 8 \\ -1 & 1 & -5 \\ 3 & 0 & 6 \\ 2 & 1 & 1 \end{matrix}\right|.$$

Applicando la riduzione a scala otteniamo

$$S = \left|\begin{matrix} 1 & -2 & 8 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}\right|;$$

quindi $\dim U = \operatorname{rg} S = 2$, e una base di U è data da $\{u_1, u_2\}$. $\square$

Esempio 6.6

Vogliamo dimensione e base di $U + W$ e $U \cap W$, dove $U \subset \mathbb{R}^4$ è il sottospazio dell'esempio precedente, mentre $W = \operatorname{Span}(w_1, w_2) \subset \mathbb{R}^4$, con

$$w_1 = \left|\begin{matrix} 1 \\ 0 \\ -3 \\ -3 \end{matrix}\right|, \quad w_2 = \left|\begin{matrix} 0 \\ 1 \\ 0 \\ 1 \end{matrix}\right|.$$

Una veloce riduzione a scala mostra che $\dim W = 2$, cioè che i vettori w_1 e w_2 sono linearmente indipendenti. Stavolta la matrice da considerare è

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

Effettuiamo una riduzione a scala di A, ottenendo

$$S = \begin{vmatrix} 1 & -2 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{vmatrix}.$$

Allora

$$\dim(U + W) = \operatorname{rg} A = \operatorname{rg} S = 3 \quad \text{e} \quad \dim(U \cap W) = \dim \operatorname{Ker} A = \dim \operatorname{Ker} S = 1.$$

Guardando la matrice S – o meglio, guardando dove sono situati i suoi pivot – vediamo subito che una base di $U + W$ è data da $\{u_1, u_2, w_2\}$; per trovare una base di $U \cap W$ dobbiamo studiare $\operatorname{Ker} A = \operatorname{Ker} S$. Il sistema omogeneo di matrice S è

$$\begin{cases} x_1 - 2x_2 + y_1 \phantom{{}+y_2} = 0, \\ -x_2 + y_1 + y_2 = 0, \\ 6y_2 = 0, \end{cases}$$

quindi $y_2 = 0$ e y_1 è libera. Una base di $\operatorname{Ker} S$ è $\{z\}$, dove $z \in \mathbb{R}^4$ è della forma

$$z = \begin{vmatrix} * \\ * \\ 1 \\ 0 \end{vmatrix}$$

e gli asterischi servono a indicare che non ci interessano gli effettivi valori delle prime due coordinate di z, in quanto (vedi la Nota 2) in (6.4) compaiono solo le y_j. Quindi una base di $U \cap W$ è $\{\psi^{-1}(z)\}$, che si ottiene facendo una combinazione lineare della base di W usando come coefficienti gli elementi di z. In altre parole, $\psi^{-1}(z) = 1w_1 + 0w_2$, per cui una base di $U \cap W$ è data da $\{w_1\}$.

6.4 Equazioni parametriche e cartesiane

L'obiettivo di questo paragrafo è vedere come si possono rappresentare i sottospazi vettoriali (e poi i sottospazi affini) di $\mathbb{R}^n$.

Ora, come si descrive un sottospazio o, più in generale, un sottoinsieme X di $\mathbb{R}^n$? Esamineremo due possibilità: tramite una parametrizzazione o tramite equazioni.

a) *Parametrizzazione*: vuol dire che esiste un insieme $U \subseteq \mathbb{R}^k$ (di solito con $k \neq n$) e una applicazione (possibilmente iniettiva) $F\colon U \to \mathbb{R}^n$ tale che $F(U) = X$. In altri termini, X è l'immagine di un'applicazione; un punto $x \in \mathbb{R}^n$ appartiene a X se e solo se $x = F(t)$ per qualche $t \in U$. L'applicazione F è detta *parametrizzazione* di X, e gli elementi di U *parametri*. Abbiamo visto esempi di parametrizzazione di rette e piani (con $U = \mathbb{R}$ o $U = \mathbb{R}^2$, rispettivamente) nel Paragrafo 2.3; più avanti vedremo come parametrizzare anche altri oggetti.

b) *Equazioni*: vuol dire che diamo delle condizioni che un punto di $\mathbb{R}^n$ deve soddisfare per appartenere a X. Formalmente, scegliamo un sottoinsieme $V \subseteq \mathbb{R}^n$ (spesso $V = \mathbb{R}^n$) e un'applicazione $G: V \to \mathbb{R}^k$ (con $k \neq n$, di solito), e diciamo che $x \in X$ se e solo se $G(x) = O$; queste sono *equazioni cartesiane* per X. Per esempio, se prendiamo $A \in M_{m,n}(\mathbb{R})$ e $G = L_A: \mathbb{R}^n \to \mathbb{R}^m$, allora $G(x) = O$ è un sistema di equazioni che descrivono $\operatorname{Ker} A$ – ma ne riparleremo più avanti.

Esempio 6.7

Per descrivere la circonferenza X di centro l'origine e raggio 1 nel piano una possibile parametrizzazione è data da $F: \mathbb{R} \to \mathbb{R}^2$ con $F(t) = (\cos t, \sin t)$; se invece prendiamo la funzione $G: \mathbb{R}^2 \to \mathbb{R}$ data da $G(x) = x_1^2 + x_2^2 - 1$, allora $G(x) = 0$ è un'equazione cartesiana della circonferenza.

Vediamo come si concretizzano queste descrizioni nel caso dei sottospazi vettoriali, cominciando con le parametrizzazioni. Sia V un sottospazio di dimensione d di $\mathbb{R}^n$, e fissiamo una base $\{v_1, \dots, v_d\}$ di V. Allora possiamo definire una parametrizzazione $F: \mathbb{R}^d \to \mathbb{R}^n$ di V ponendo

$$F\begin{vmatrix} t_1 \\ \vdots \\ t_d \end{vmatrix} = t_1 v_1 + \cdots + t_d v_d.$$

Nota che F è data da una matrice; infatti, essendo $F(e_j) = v_j$ per $j = 1, \dots, d$, la Proposizione 5.2 ci dice che $F = L_A$, dove A è la matrice $A = \begin{vmatrix} v_1 & \cdots & v_d \end{vmatrix} \in M_{n,d}(\mathbb{R})$.

Questo ci suggerisce una definizione.

Definizione 6.2

Una *parametrizzazione lineare* di un sottospazio V di $\mathbb{R}^n$ è un'applicazione lineare $L_A: \mathbb{R}^k \to \mathbb{R}^n$ tale che $V = \operatorname{Im} A$, con $A \in M_{n,k}(\mathbb{R})$; in altri termini, V viene descritto come span delle colonne di A. In coordinate, i vettori $v \in V$ saranno descritti dalle equazioni $v = t_1 A^1 + \cdots + t_k A^k$, per cui *equazioni parametriche* del sottospazio $V \subseteq \mathbb{R}^n$ sono

$$\begin{cases} x_1 = a_{11} t_1 + \cdots + a_{1k} t_k, \\ \qquad \vdots \\ x_n = a_{n1} t_1 + \cdots + a_{nk} t_k, \end{cases}$$

dove abbiamo posto $v = \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}$.

Osservazione 6.5

È evidente che le equazioni parametriche non sono per niente uniche. Inoltre, $\dim V = \operatorname{rg} A \leq k$; quindi il numero dei parametri in una parametrizzazione potrebbe essere eccessivo, superiore alla dimensione del sottospazio se le colonne della matrice A non sono linearmente indipendenti (cioè se L_A non è iniettiva; è quanto capitava nell'Esempio 4.6). Di solito, in tal caso si sostituisce ad A una matrice $\tilde{A}$ di ugual immagine e con rango esattamente uguale a $\dim V$, ottenuta per esempio considerando solo le colonne di A che formano una base di $\operatorname{Im} A = V$.

Esempio 6.8

Vogliamo trovare equazioni parametriche per il sottospazio $U \subset \mathbb{R}^4$ dell'Esempio 6.5. Un primo insieme di equazioni parametriche l'abbiamo subito

$$\begin{cases} x_1 = t_1 - 2t_2 + 8t_3, \\ x_2 = -t_1 + t_2 - 5t_3, \\ x_3 = 3t_1 + 6t_3, \\ x_4 = 2t_1 + t_2 + t_3. \end{cases}$$

In realtà, come abbiamo visto nell'Esempio 6.5, ci sono dei parametri superflui, in quanto la dimensione di U è 2, non 3. Utilizzando la base di U che avevamo trovato in quell'esempio otteniamo come equazioni parametriche

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

Vediamo ora le equazioni cartesiane. Cerchiamo di usare soltanto applicazioni lineari, come abbiamo fatto per le equazioni parametriche. Dato un sottospazio V di $\mathbb{R}^n$ vogliamo descriverlo come nucleo di una matrice $B \in M_{p,n}(\mathbb{R})$, cioè dire che $x \in V$ se e solo se $Bx = O$, ovvero, in coordinate, che $x \in V$ se e solo se

$$\begin{cases} b_{11}x_1 + \cdots + b_{1n}x_n = 0, \\ \vdots \\ b_{p1}x_1 + \cdots + b_{pn}x_n = 0. \end{cases} \tag{6.5}$$

Definizione 6.3

Il sistema (6.5) raccoglie *equazioni cartesiane (lineari)* per il sottospazio V, che risulta essere lo spazio delle soluzioni del sistema omogeneo di matrice B.

Osservazione 6.6

In questo caso, la dimensione di V è data da $n - \operatorname{rg} B$ (Osservazione 5.8), con $\operatorname{rg} B \leq p$. Dunque in generale la dimensione di V potrebbe essere maggiore del numero di incognite (n) meno il numero di equazioni (p), in quanto potrebbero esservi delle equazioni inutili (come accade nel sistema omogeneo associato al sistema studiato nell'Esempio 6.2). Di solito, in tal caso si sostituisce a B una matrice $\tilde{B}$ di ugual nucleo e con esattamente $\operatorname{rg} B = n - \dim V$ righe, ottenuta per esempio tramite una riduzione a scala di B. In particolare, anche le equazioni cartesiane non sono uniche.

Esempio 6.9

Il sottospazio $V \subset \mathbb{R}^4$ delle soluzioni del sistema

$$\begin{cases} 2x_1 - x_2 + 4x_3 + x_4 = 0, \\ -2x_1 + x_2 - 7x_3 + x_4 = 0, \\ 4x_1 - 2x_2 + 5x_3 + 4x_4 = 0, \end{cases}$$

ha come equazioni cartesiane il sistema appena scritto, cioè il sistema $Bx = O$ con

$$B = \begin{vmatrix} 2 & -1 & 4 & 1 \\ -2 & 1 & -7 & 1 \\ 4 & -2 & 5 & 4 \end{vmatrix}.$$

Per quanto visto nell'Esempio 6.2, a B possiamo sostituire $\tilde{B} = \begin{vmatrix} 2 & -1 & 4 & 1 \\ 0 & 0 & -3 & 2 \end{vmatrix}$, ottenendo

$$\begin{cases} 2x_1 - x_2 + 4x_3 + x_4 = 0, \\ -3x_3 + 2x_4 = 0, \end{cases} \tag{6.6}$$

come altre equazioni cartesiane per V.

Si pongono ora due problemi: è vero che ogni sottospazio di $\mathbb{R}^n$ è descrivibile da equazioni cartesiane lineari? E come si passa da equazioni parametriche a equazioni cartesiane o viceversa?

Passare da equazioni cartesiane a equazioni parametriche è facile: basta risolvere il sistema omogeneo corrispondente.

Esempio 6.10

Vogliamo equazioni parametriche per il sottospazio $V \subset \mathbb{R}^4$ dell'Esempio 6.9. Risolvendo all'indietro il sistema (6.6) troviamo

$$\begin{cases} x_1 = \frac{1}{2}x_2 - \frac{11}{6}x_4, \\ x_3 = \frac{2}{3}x_4; \end{cases}$$

quindi equazioni parametriche sono

$$\begin{cases} x_1 = \frac{1}{2}t_1 - \frac{11}{6}t_2, \\ x_2 = t_1, \\ x_3 = \frac{2}{3}t_2, \\ x_4 = t_2, \end{cases}$$

ovvero $V = \operatorname{Im} A$ con

$$A = \begin{vmatrix} 1/2 & -11/6 \\ 1 & 0 \\ 0 & 2/3 \\ 0 & 1 \end{vmatrix}.$$

Viceversa, si può sempre passare da equazioni parametriche a equazioni cartesiane (per cui, in particolare, ogni sottospazio può venire descritto da equazioni cartesiane lineari). Sia $V = \operatorname{Im} A \subseteq \mathbb{R}^n$, in modo che i vettori $v \in V$ siano della forma $v = At$. Per passare a equazioni cartesiane, dobbiamo trovare quali condizioni devono soddisfare le coordinate di $v \in \mathbb{R}^n$ perché il sistema $At = v$ ammetta soluzione. Effettuiamo una riduzione a scala; otteniamo il sistema equivalente $St = c$, dove S è a scala con le ultime $n - \operatorname{rg} A$ righe nulle. Dunque il sistema $At = v$ è compatibile se e solo se le ultime $n - \operatorname{rg} A$ componenti di c sono uguali a zero – e queste sono le equazioni cartesiane cercate.

Esempio 6.11

Usiamo questo metodo per il sottospazio $U = \operatorname{Im} A$ dell'Esempio 6.8, dove

$$A = \left|\begin{matrix} 1 & -2 \\ -1 & 1 \\ 3 & 0 \\ 2 & 1 \end{matrix}\right|.$$

Operando una riduzione a scala sul sistema $At = v$ otteniamo

$$\left|\begin{array}{cc|c} 1 & -2 & x_1 \\ -1 & 1 & x_2 \\ 3 & 0 & x_3 \\ 2 & 1 & x_4 \end{array}\right| \longrightarrow \left|\begin{array}{cc|c} 1 & -2 & x_1 \\ 0 & -1 & x_2 + x_1 \\ 0 & 6 & x_3 - 3x_1 \\ 0 & 5 & x_4 - 2x_1 \end{array}\right| \longrightarrow \left|\begin{array}{cc|c} 1 & -2 & x_1 \\ 0 & -1 & x_1 + x_2 \\ 0 & 0 & 3x_1 + 6x_2 + x_3 \\ 0 & 0 & 3x_1 + 5x_2 + x_4 \end{array}\right|,$$

per cui delle possibili equazioni cartesiane sono

$$\begin{cases} 3x_1 + 6x_2 + x_3 = 0, \\ 3x_1 + 5x_2 + x_4 = 0. \end{cases}$$

Osservazione 6.7

È molto facile trovare le equazioni cartesiane di un'intersezione. Infatti se $B_1x = O$ sono equazioni cartesiane del sottospazio $U \subseteq \mathbb{R}^n$, e $B_2x = O$ sono equazioni cartesiane del sottospazio $W \subseteq \mathbb{R}^n$, allora $x \in U \cap W$ se e solo se soddisfa sia $B_1x = O$ sia $B_2x = O$, per cui le equazioni cartesiane di $U \cap W$ sono semplicemente

$$\begin{cases} B_1x = O, \\ B_2x = O. \end{cases}$$

6.5 Sottospazi affini

Abbiamo quindi visto come descrivere i sottospazi vettoriali di $\mathbb{R}^n$, e dunque in particolare l'insieme delle soluzioni di un sistema lineare omogeneo. In questo paragrafo discuteremo come descrivere l'insieme delle soluzioni di un sistema lineare non omogeneo – o, meglio, l'analogo in spazi vettoriali qualunque.

Definizione 6.4

Un *sottospazio affine* L di uno spazio vettoriale V è un sottoinsieme di V della forma $L = v_0 + W = \{v_0 + w \mid w \in W\}$, dove $v_0 \in V$ e W è un sottospazio vettoriale di V; diremo che L è *parallelo* a W, e che W è il *sottospazio di giacitura* di L. La *dimensione* $\dim L$ di L è per definizione uguale alla dimensione di W.

Osservazione 6.8

Di solito, un sottospazio affine di dimensione 1 si chiama *retta*, e uno di dimensione 2 *piano*. A volte, un sottospazio affine di V di dimensione $\dim V - 1$ viene chiamato *iperpiano*.

In parole povere, un sottospazio affine è un *traslato* di un sottospazio vettoriale. Per esempio, una retta che non passa per l'origine (o anche che ci passa, in quanto i sottospazi vettoriali sono particolari sottospazi affini, basta prendere $v_0 = O$) è un sottospazio affine. Inoltre, il Teorema di struttura (Proposizione 5.1) ci dice che l'insieme delle soluzioni di un sistema lineare $Ax = b$ è un sottospazio affine, parallelo allo spazio delle soluzioni del sistema omogeneo associato $Ax = O$.

Definizione 6.5

Siano $L_1, L_2 \subseteq V$ sottospazi affini, con sottospazi di giacitura rispettivamente W_1 e W_2. Diremo che L_1 e L_2 sono *paralleli* se $W_1 \subseteq W_2$ oppure $W_2 \subseteq W_1$.

Un sottospazio affine di $\mathbb{R}^n$ può essere descritto con equazioni parametriche o cartesiane, a seconda di come viene descritto il sottospazio di giacitura W.

Definizione 6.6

Un sottospazio affine $L \subseteq \mathbb{R}^n$ si dice descritto da *equazioni parametriche* se $L = \{v = At + v_0 \in \mathbb{R}^n \mid t \in \mathbb{R}^k\} = v_0 + \operatorname{Im} A$, per qualche $v_0 \in \mathbb{R}^n$ e $A \in M_{n,k}(\mathbb{R})$. Il sottospazio L è l'immagine dell'applicazione $F \colon \mathbb{R}^k \to \mathbb{R}^n$ data da $F(t) = At + v_0$, e il sottospazio di giacitura ha equazioni parametriche $w = At$.

Definizione 6.7

Un sottospazio affine $L \subseteq \mathbb{R}^n$ si dice descritto da *equazioni cartesiane* se $L = \{x \in \mathbb{R}^n \mid Bx = b\}$ per qualche $B \in M_{p,n}(\mathbb{R})$ e $b \in \mathbb{R}^p$; il sottospazio affine L è l'insieme dei vettori di $\mathbb{R}^n$ dove l'applicazione $G \colon \mathbb{R}^n \to \mathbb{R}^p$ data da $G(x) = Bx - b$ si annulla. In questo caso, il sottospazio di giacitura ha equazioni cartesiane $Bx = O$.

Di nuovo si pone il problema di passare dalle equazioni parametriche a quelle cartesiane e viceversa. Se $Bx = b$ sono equazioni cartesiane, risolvendo il sistema troviamo equazioni parametriche; in altre parole, "risolvere" un sistema lineare vuol dire passare da equazioni cartesiane a equazioni parametriche. Viceversa:

Proposizione 6.4

Sia $L \subseteq \mathbb{R}^n$ un sottospazio affine di $\mathbb{R}^n$ di equazioni parametriche $v = At + v_0$, e siano $Bx = O$ equazioni cartesiane del sottospazio di giacitura W. Allora equazioni cartesiane di L sono date da

$$Bx = Bv_0. \tag{6.7}$$

Dimostrazione

Infatti, un vettore $x \in \mathbb{R}^n$ appartiene a L se e solo se $x - v_0 \in W$, che accade se e solo se $B(x - v_0) = O$.

□

Esempio 6.12

Consideriamo il sistema lineare

$$\begin{cases} 2x_1 - x_2 + x_3 = 1, \\ \qquad\; x_2 + x_3 = 3, \\ x_1 \qquad\; + x_3 = 2. \end{cases}$$

L'insieme L delle soluzioni di questo sistema è un sottospazio affine di $\mathbb{R}^3$; vogliamo descriverlo tramite equazioni parametriche. Effettuando una riduzione a scala troviamo che la terza equazione è inutile; risolvendo all'indietro il sistema a scala formato dalle prime due, otteniamo le equazioni parametriche

$$\begin{cases} x_1 = 2 - t, \\ x_2 = 3 - t, \\ x_3 = t, \end{cases} \qquad \text{cioè} \qquad x = \begin{vmatrix} 2 \\ 3 \\ 0 \end{vmatrix} + t \begin{vmatrix} -1 \\ -1 \\ 1 \end{vmatrix}.$$

Quindi L è una retta passante per $\begin{vmatrix} 2 \\ 3 \\ 0 \end{vmatrix}$ e parallela a $\begin{vmatrix} -1 \\ -1 \\ 1 \end{vmatrix}$.

Esempio 6.13

Consideriamo il sottospazio affine $L \subseteq \mathbb{R}^4$ dato da

$$\begin{cases} x_1 = 3 + s - t, \\ x_2 = -1 + 2s, \\ x_3 = 1 + 2t, \\ x_4 = s + t, \end{cases}$$

cioè $v = At + v_0$, con

$$A = \begin{vmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 2 \\ 1 & 1 \end{vmatrix} \qquad \text{e} \qquad v_0 = \begin{vmatrix} 3 \\ -1 \\ 1 \\ 0 \end{vmatrix}.$$

Effettuando una riduzione a scala del sistema $At = v$, troviamo

$$\left|\begin{array}{cc|c} 1 & -1 & x_1 \\ 2 & 0 & x_2 \\ 0 & 2 & x_3 \\ 1 & 1 & x_4 \end{array}\right| \longrightarrow \left|\begin{array}{cc|c} 1 & -1 & x_1 \\ 0 & 2 & x_2 - 2x_1 \\ 0 & 2 & x_3 \\ 0 & 2 & x_4 - x_1 \end{array}\right| \longrightarrow \left|\begin{array}{cc|c} 1 & -1 & x_1 \\ 0 & 2 & x_2 - 2x_1 \\ 0 & 0 & x_3 - x_2 + 2x_1 \\ 0 & 0 & x_4 - x_2 + x_1 \end{array}\right|;$$

quindi la Proposizione 6.4 ci dice che equazioni cartesiane per L sono date da

$$\begin{cases} 2x_1 - x_2 + x_3 = 8, \\ x_1 - x_2 + x_4 = 4. \end{cases}$$

Osservazione 6.9

Per descrivere una retta nello spazio servono 2 equazioni cartesiane. Infatti, perché equazioni della forma $Bx = b$, con $x \in \mathbb{R}^3$ e $B \in M_{p,3}(\mathbb{R})$, descrivano una retta, occorre che si abbia $1 = \dim \operatorname{Ker} B = 3 - \operatorname{rg} B$, cioè $\operatorname{rg} B = 2$. Quindi (dopo aver eliminato le equazioni inutili) $p = 2$. Analogamente, per descrivere una retta in $\mathbb{R}^n$ servono $n - 1$ equazioni cartesiane.

Esercizi 6

6.1 Studia (cioè vedi se ammettono soluzioni, e in tal caso trovale) i sistemi:

$$\begin{cases} 4x+y+z+2v+3w=0, \\ 14x+2y+2z+7v+11w=0, \\ 15x+3y+3z+6v+10w=0; \end{cases} \qquad \begin{cases} 5x+4y+7z=3, \\ x+2y+3z=1, \\ x-y-z=0, \\ 3x+3y+5z=2. \end{cases}$$

6.2 Studia i sistemi lineari:

$$\begin{cases} x+2y-z=0, \\ 4y-z-2w=0, \\ \sqrt{5}x+7y-2z-3w=0; \end{cases} \qquad \begin{cases} 2x_1+x_2-x_3-x_4=3, \\ 3x_1-x_2-x_3=8, \\ x_1-x_3+x_4=4. \end{cases}$$

6.3 Studia i sistemi lineari:

$$\begin{cases} 2x+y-z=0, \\ x-y-z=1, \\ 2x+z=1, \\ \sqrt{2}x+3y-z=1; \end{cases} \qquad \begin{cases} x_1+x_2+x_3=0, \\ 5x_1-x_2+x_3=0, \\ x_1+x_2-x_3=6, \\ 4x_1+x_2+x_3=3. \end{cases}$$

6.4 Studia i sistemi lineari a coefficienti complessi:

$$\begin{cases} x+iy+z=1, \\ 5x-(2-i)y+iz=2, \\ (1+i)x+y-2iz=4-3i; \end{cases} \qquad \begin{cases} x_1+ix_2+(2-3i)x_3=1+i, \\ x_1-2ix_2+2x_3=4, \\ (1-3i)x_1+ix_2-(4+9i)x_3=i-3. \end{cases}$$

6.5 Studia i sistemi lineari a coefficienti complessi:

$$\begin{cases} ix-y+2z-w=1-i, \\ x+(1-2i)y-2z+iw=5i, \\ 3ix+iz-(1-2i)w=3+5i; \end{cases} \qquad \begin{cases} x_1+2ix_2+x_3=4, \\ 2x_1+4x_2+5ix_3=2i, \\ 3x_1+(1+2i)x_2-(5+i)x_3=5i-2, \\ 3x_1+(3+2i)x_2+(3i-1)x_3=2+3i. \end{cases}$$

6.6 Studia al variare dei parametri $k, h \in \mathbb{R}$ i seguenti sistemi:

$$\begin{cases} x+(k-1)y+z=1, \\ 2x+ky+kz=k, \\ kx+2y+(2k-2)z=4-k; \end{cases} \qquad \begin{cases} 3kx+3y+(k+2)z=h, \\ x+2y+2z=k+1, \\ x+ky+kz=1; \end{cases} \qquad \begin{cases} kx+y+z=1, \\ x+ky+z=h, \\ x+y+kz=h^2. \end{cases}$$

6.7 Studia al variare di $a, k \in \mathbb{C}$ (e con $x, y, z \in \mathbb{C}$) il sistema lineare

$$\begin{cases} x+2y+kz=0, \\ ky-z=1, \\ kx+z=a. \end{cases}$$

6.8 Al variare di $k, h \in \mathbb{C}$ studia i sistemi lineari:

$$\begin{cases} x + ky + (4k+1)z = 2, \\ 2x + 2y + 6z = 4, \\ kx + \dfrac{1}{4}y + (1+k)z = 2k + h; \end{cases} \qquad \begin{cases} x_1 + kx_2 + k^2x_3 = 2, \\ x_1 + x_2 + 3x_3 = h, \\ x_1 + 2x_2 + (k+1)x_3 = 2k + h. \end{cases}$$

6.9 Se A è una matrice quadrata di ordine n, dimostra che il rango di A è uguale al numero di pivot (nel senso del Paragrafo 3.3) *non nulli* ottenuti con una qualunque eliminazione di Gauss di A.

6.10 Determina dimensione e base del nucleo e dell'immagine delle seguenti matrici e delle loro trasposte:

$$\begin{vmatrix} 1 & -1 & 4 \\ 1 & 10 & 1 \\ 1 & 3 & 2 \end{vmatrix}, \qquad \begin{vmatrix} 3 & 1 & 1 \\ 2 & 1 & 0 \\ 2 & 1 & 7 \end{vmatrix}.$$

6.11 Determina dimensione e base del nucleo e dell'immagine delle seguenti matrici e delle loro trasposte:

$$\begin{vmatrix} 0 & 1 & 3 \\ 2 & 3 & -5 \\ 7 & 1 & 1 \\ 3 & 1 & 2 \end{vmatrix}, \qquad \begin{vmatrix} 2 & -1 & -1 \\ 3 & -7 & 4 \end{vmatrix}.$$

6.12 Trova dimensione e base del nucleo e dell'immagine di A, B, A^T e B^T, dove

$$A = \begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & -1 & -2 & 2 \\ 5 & 8 & 1 & 4 \end{vmatrix}, \qquad B = \begin{vmatrix} 2 & 1 \\ 1 & 1 \\ 1 & \pi \end{vmatrix}.$$

6.13 Siano $A \in M_{3,3}(\mathbb{R})$ e $b_k \in \mathbb{R}^3$ dati da

$$A = \begin{vmatrix} 0 & 1 & 1 \\ 2 & 3 & 7 \\ 1 & -3 & -1 \end{vmatrix}, \qquad b_k = \begin{vmatrix} k \\ 5 \\ -2 \end{vmatrix}.$$

Trova dimensione e base dei sottospazi $\operatorname{Ker} A$, $\operatorname{Im} A$, $\operatorname{Ker} A^T$ e $\operatorname{Im} A^T$, e studia il sistema $Ax = b_k$ al variare del parametro $k \in \mathbb{R}$, e al variare del parametro $k \in \mathbb{C}$.

6.14 Sia $T: \mathbb{R}^3 \to \mathbb{R}^4$ data da

$$T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} y + z \\ 2x + y - z \\ z \\ 2x + 2y \end{vmatrix}.$$

Determina immagine e nucleo di T, e, se U e W sono i sottospazi di $\mathbb{R}^3$ dati da

$$U = \operatorname{Span}\left(\begin{vmatrix} 2 \\ 1 \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ -1 \\ 2 \end{vmatrix}\right) \quad \text{e} \quad W = \operatorname{Span}\left(\begin{vmatrix} -1 \\ 0 \\ 1 \end{vmatrix}, \begin{vmatrix} 0 \\ 1 \\ 0 \end{vmatrix}\right),$$

trova dimensione e base per $T(U) \cap T(W)$.

6.15 Sia $A_\alpha \in M_{3,3}(\mathbb{R})$ la matrice

$$A_\alpha = \begin{vmatrix} \alpha & \alpha & \alpha+1 \\ -1 & 0 & 2 \\ 2 & 0 & \alpha \end{vmatrix}.$$

Determina, al variare di $\alpha \in \mathbb{R}$, le dimensioni di $\operatorname{Ker} A_\alpha$ e $\operatorname{Im} A_\alpha$, e trova per quali valori di α si ha $\mathbb{R}^3 = \operatorname{Ker} A_\alpha \oplus \operatorname{Im} A_\alpha$.

6.16 Considera i due sottospazi di $\mathbb{R}^4$ dati da

$$U = \{x \in \mathbb{R}^4 \mid 2x_2 - x_3 = 0\} \text{ e } W = \{x \in \mathbb{R}^4 \mid 2x_1 + x_2 - x_4 = 0, 3x_3 - \frac{5}{2}x_4 = 0\}.$$

Trova dimensione, base, ed equazioni cartesiane e parametriche per $U + W$ e $U \cap W$.

6.17 Considera i due sottospazi di $\mathbb{R}^4$ dati da

$$U = \operatorname{Span}\left(\begin{vmatrix} 1 \\ 2 \\ -2 \\ 3 \end{vmatrix}, \begin{vmatrix} 2 \\ 1 \\ 1 \\ -2 \end{vmatrix}, \begin{vmatrix} 4 \\ 5 \\ 1 \\ 1 \end{vmatrix} \right) \text{ e } W = \operatorname{Span}\left(\begin{vmatrix} 0 \\ 3 \\ 1 \\ 1 \end{vmatrix}, \begin{vmatrix} 1 \\ 1 \\ 1 \\ -2 \end{vmatrix}, \begin{vmatrix} 2 \\ 5 \\ 3 \\ 0 \end{vmatrix} \right).$$

Trova dimensione, base ed equazioni cartesiane e parametriche per $U + W$ e $U \cap W$.

6.18 Considera i due sottospazi di $\mathbb{R}^4$ dati da $U = \operatorname{Span}\left(\begin{vmatrix} 6 \\ 1 \\ 3 \\ 2 \end{vmatrix}, \begin{vmatrix} 1 \\ 4 \\ -5 \\ 1 \end{vmatrix}, \begin{vmatrix} 2 \\ 2 \\ 1 \\ 7 \end{vmatrix} \right)$ e $W = \{x \in \mathbb{R}^4 \mid x_1 + 2x_3 - x_4 = 0,\ 2x_1 + 3x_2 - 4x_3 + 2x_4 = 0\}$. Trova dimensione, base ed equazioni cartesiane e parametriche per $U + W$ e $U \cap W$.

6.19 Scrivi equazioni cartesiane per il sottospazio di $\mathbb{R}^4$ di equazioni parametriche

$$x = t_1 - 2t_2 + t_3, \quad y = t_1 + t_3, \quad z = t_1 + 4t_2 - 5t_3, \quad w = t_2 - t_3.$$

Da tali equazioni passa poi nuovamente a equazioni parametriche: in questo modo ottieni le equazioni di partenza oppure no? Perché?

6.20 Scrivi equazioni cartesiane per il sottospazio di $\mathbb{R}^4$ di equazioni parametriche

$$x = t_1 + 2t_2 - 3t_3, \quad y = 4t_1 - 5t_2 + t_3$$
$$z = 6t_1 - 2t_2 - 4t_3, \quad w = -3t_1 + 2t_2 + t_3.$$

Da tali equazioni passa poi nuovamente a equazioni parametriche: in questo modo ottieni le equazioni di partenza oppure no? Perché?

6.21 Scrivi equazioni parametriche per il sottospazio di $\mathbb{R}^4$ di equazioni cartesiane

$$x + y + z - w = 0, \quad 2x - 3y + 7z + 13w = 0, \quad x - y + 3z + 5w = 0.$$

Da tali equazioni passa poi nuovamente a equazioni cartesiane: in questo modo ottieni le equazioni di partenza oppure no? Perché?

6.22 Considera i sottospazi:

$$V_t = \text{Span}\left(\begin{vmatrix}0\\1\\t\end{vmatrix}, \begin{vmatrix}1\\t\\1\end{vmatrix}, \begin{vmatrix}t\\0\\0\end{vmatrix}\right) \subseteq \mathbb{R}^3, \quad W_t = \text{Span}\left(\begin{vmatrix}0\\t\\1\\-1\end{vmatrix}, \begin{vmatrix}1\\0\\t\\1\end{vmatrix}, \begin{vmatrix}0\\t\\0\\1\end{vmatrix}, \begin{vmatrix}1\\0\\0\\t\end{vmatrix}\right) \subseteq \mathbb{R}^4$$

al variare di $t \in \mathbb{R}$. Determina dimensione e base di $\bigcap_{t\in\mathbb{R}} V_t$ e di $\bigcap_{t\in\mathbb{R}} W_t$.

6.23 Trova equazioni cartesiane dei sottospazi di $\mathbb{R}^4$ di equazioni parametriche:

$$\begin{cases} x_1 = t - 7, \\ x_2 = -t + 2, \\ x_3 = -t - 6, \\ x_4 = t - 5; \end{cases} \qquad \begin{cases} x_1 = t_1 + 3t_2, \\ x_2 = t_2 + t_3, \\ x_3 = 2t_1 + t_3, \\ x_4 = t_1 + t_2 + t_3. \end{cases}$$

6.24 Al variare di $t \in \mathbb{R}$, trova equazioni parametriche e cartesiane per:

$$V_t = \text{Span}\left(\begin{vmatrix}1\\1\\t\\1\end{vmatrix}, \begin{vmatrix}0\\t\\0\\1\end{vmatrix}, \begin{vmatrix}0\\0\\t\\1\end{vmatrix}, \begin{vmatrix}0\\2t\\1\\0\end{vmatrix}\right) \subset \mathbb{R}^4.$$

6.25 Siano W_1 e W_2 sottospazi di uno spazio vettoriale V, e $v_1, v_2 \in V$. Dimostra che $v_1 + W_1 = v_2 + W_2$ se e solo se $W_1 = W_2$ e $v_2 - v_1 \in W_1 = W_2$. In particolare, $v_1 + W_1$ è un sottospazio vettoriale se e solo se $v_1 \in W_1$.

6.26 Scrivi equazioni cartesiane per il sottospazio affine di $\mathbb{R}^4$ di equazioni parametriche

$$x = 2t_1 + t_2 - 4, \quad y = 2t_1 - t_2 + 5, \quad z = 3t_1 - t_2 + 2t_3 - 7, \quad w = -3t_1 + t_2 - 4t_3.$$

6.27 Scrivi equazioni parametriche per il sottospazio affine di $\mathbb{R}^4$ di equazioni cartesiane

$$x + 3y + 2z - 3w = 5, \quad 2x - 2y + 2z + 4w = 1, \quad 5x - 2y - 3z + 6w = 11.$$

6.28 Definisci i concetti di equazioni parametriche e cartesiane lineari per un sottospazio affine di uno spazio vettoriale V qualunque. (*Suggerimento:* utilizza applicazioni lineari da V a $\mathbb{R}^n$ o viceversa.)

6.29 Siano $L_1 = v_1 + W_1$ ed $L_2 = v_2 + W_2$ due sottospazi affini di uno spazio vettoriale V, e poniamo $W_0 = W_1 + W_2 + \text{Span}(v_2 - v_1)$. Dimostra che $L_0 = v_1 + W_0$ è il più piccolo sottospazio affine di V contenente sia L_1 sia L_2. Il sottospazio affine L_0 viene chiamato *somma* di L_1 ed L_2, e indicato con $L_1 + L_2$.

6.30 Siano L_1, L_2 sottospazi affini di uno spazio vettoriale V. Dimostra il *Teorema di Grassmann per sottospazi affini:*

$$\dim(L_1 + L_2) + \dim(L_1 \cap L_2) \leq \dim L_1 + \dim L_2,$$

dove per convenzione si pone $\dim(L_1 \cap L_2) = -1$ se $L_1 \cap L_2 = \emptyset$. Dimostra che l'uguaglianza vale se e solo se $L_1 \cap L_2 \neq \emptyset$ oppure $L_1 \cap L_2 = \emptyset$ e $W_1 \cap W_2 = \{O\}$.

Matrici e applicazioni lineari

7

Sommario

Nel Capitolo 5 abbiamo introdotto le applicazioni lineari; in questo capitolo ne riprendiamo lo studio concentrandoci sulle operazioni che possiamo effettuare con esse. Vedremo come sommarle e moltiplicarle per uno scalare, in modo che l'insieme delle applicazioni lineari fra due spazi vettoriali dati risulti essere a sua volta uno spazio vettoriale. Studieremo la composizione di applicazioni lineari, e questo ci condurrà alla fondamentale nozione di isomorfismo fra spazi vettoriali. Vedremo anche che tutte e sole le applicazioni lineari fra $\mathbb{R}^n$ ed $\mathbb{R}^m$ sono quelle della forma L_A dove A è una matrice $m \times n$, e quindi definiremo il prodotto di matrici tramite la composizione delle corrispondenti applicazioni lineari. Infine ci soffermeremo sul caso delle matrici quadrate, dando diversi criteri per stabilire quando una matrice è invertibile e descrivendo un algoritmo per il calcolo dell'inversa. Per semplicità, anche in questo capitolo lavoreremo solo su $\mathbb{R}$, ma tutto quanto diremo varrà anche per applicazioni lineari su un campo $\mathbb{K}$ qualsiasi.

7.1 Composizione e isomorfismi

Il nostro primo obiettivo è definire una struttura di spazio vettoriale sull'insieme delle applicazioni lineari fra due spazi vettoriali dati. Per farlo è sufficiente imitare quanto visto nell'Esempio 4.2.

Definizione 7.1

Siano V e W due spazi vettoriali, $S, T\colon V \to W$ due applicazioni lineari e $\lambda \in \mathbb{R}$. Definiamo le applicazioni $S + T\colon V \to W$ e $\lambda T\colon V \to W$ ponendo

$$(S+T)(v) = S(v) + T(v) \qquad \text{e} \qquad (\lambda T)(v) = \lambda\bigl(T(v)\bigr)$$

per ogni $v \in V$. Le applicazioni $S + T$ e λT sono ancora lineari, come si verifica subito: per esempio

$$(S+T)(\lambda v) = S(\lambda v) + T(\lambda v) = \lambda S(v) + \lambda T(v) = \lambda\bigl(S(v) + T(v)\bigr) = \lambda(S+T)(v),$$

e così via (controllare, prego). L'insieme $\mathcal{L}(V, W)$ delle applicazioni lineari da V in W con queste operazioni diviene uno spazio vettoriale (esercizio). In particolare, lo spazio $\mathcal{L}(V, \mathbb{R})$ delle applicazioni lineari da V in $\mathbb{R}$ si chiama *spazio duale* di V e si indica con V' (o V^*).

Ora, vi è un'altra operazione naturale fra applicazioni lineari: la composizione.

Definizione 7.2

Siano $S\colon U \to V$ e $T\colon V \to W$ due applicazioni lineari; allora la *composizione* di S e T è l'applicazione $T \circ S\colon U \to W$ data da

$$\forall u \in U \qquad (T \circ S)(u) = T\bigl(S(u)\bigr).$$

Proposizione 7.1

Siano $S\colon U \to V$ e $T\colon V \to W$ applicazioni lineari fra gli spazi vettoriali U, V e W. Allora anche la composizione $T \circ S\colon U \to W$ è lineare.

Dimostrazione

Si tratta di una semplice verifica. Presi $u_1, u_2 \in U$ si ha

$$\begin{aligned}(T \circ S)(u_1 + u_2) &= T\bigl(S(u_1 + u_2)\bigr) = T\bigl(S(u_1) + S(u_2)\bigr)\\ &= T\bigl(S(u_1)\bigr) + T\bigl(S(u_2)\bigr) = (T \circ S)(u_1) + (T \circ S)(u_2).\end{aligned}$$

In modo analogo si dimostra che T è omogenea (esercizio). □

Osservazione 7.1

Nel caso in cui $U = V = W$, cioè quando sia T sia S sono endomorfismi, risultano definiti sia $T \circ S$ sia $S \circ T$. In generale, però, la composizione *non* è commutativa, cioè $T \circ S \neq S \circ T$.

Esempio 7.1

Definiamo $S, T\colon \mathbb{R}^3 \to \mathbb{R}^3$ ponendo

$$S\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x \\ y \\ 0 \end{vmatrix} \qquad \text{e} \qquad T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x \\ z \\ 0 \end{vmatrix}.$$

Allora

$$(S \circ T)\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x \\ z \\ 0 \end{vmatrix} \neq \begin{vmatrix} x \\ 0 \\ 0 \end{vmatrix} = (T \circ S)\begin{vmatrix} x \\ y \\ z \end{vmatrix}.$$

Ovviamente, esistono endomorfismi che commutano; per esempio, se $S\colon V \longrightarrow V$ è un endomorfismo qualunque, allora $S \circ \mathrm{id}_V = \mathrm{id}_V \circ S = S$ e $S \circ O = O \circ S = O$.

La composizione, come nel caso di funzioni qualunque, è sempre associativa

$$R \circ (S \circ T) = (R \circ S) \circ T$$

non appena R, S e T sono applicazioni lineari tali che $S \circ T$ e $R \circ S$ siano definite (da cui segue automaticamente che anche $R \circ (S \circ T)$ e $(R \circ S) \circ T$ sono definite; perché?). È facile verificare (esercizio) che la composizione è distributiva rispetto alla somma e al prodotto per scalari, cioè che

$$(S_1 + S_2) \circ T = S_1 \circ T + S_2 \circ T, \qquad S \circ (T_1 + T_2) = S \circ T_1 + S \circ T_2,$$
$$(\lambda S) \circ T = \lambda(S \circ T) = S \circ (\lambda T),$$

non appena tutte le applicazioni lineari coinvolte sono definite. In particolare, quindi, su $\mathcal{L}(V, V)$ la somma e la composizione sono operazioni che verificano le proprietà (1)–(7) definite nel Paragrafo 1.3 (l'elemento neutro rispetto alla somma è l'applicazione nulla O e quello rispetto alla composizione è l'identità id_V); si dice anche che $\mathcal{L}(V, V)$ con tali operazioni è un *anello non commutativo.*

Definizione 7.3

Diremo che un'applicazione lineare $T\colon V \longrightarrow W$ è *invertibile* se esiste un'applicazione *lineare* $S\colon W \longrightarrow V$, l'*inversa* di T, tale che $T \circ S = \mathrm{id}_W$ e $S \circ T = \mathrm{id}_V$. L'inversa di T, se esiste, si indica con T^{-1}.

Osservazione 7.2

L'inversa, se esiste, è unica: se S ed S' sono due inverse di un'applicazione T, abbiamo

$$S' = S' \circ \mathrm{id}_W = S' \circ (T \circ S) = (S' \circ T) \circ S = \mathrm{id}_V \circ S = S.$$

Nella definizione di inversa abbiamo richiesto esplicitamente che l'inversa sia lineare: stiamo lavorando in spazi vettoriali con applicazioni lineari, pertanto una inversa non lineare (che non rispetta le operazioni) sarebbe del tutto inutile. In realtà, non appena $T\colon V \longrightarrow W$ è invertibile come funzione (cioè non appena è bigettiva), allora l'inversa è automaticamente lineare.

Proposizione 7.2

Sia $T\colon V \longrightarrow W$ un'applicazione lineare. Allora T è invertibile se e solo se è iniettiva e surgettiva.

Dimostrazione

Se T è invertibile è chiaramente bigettiva. Viceversa, supponiamo T iniettiva e surgettiva; allora esiste la funzione inversa $S\colon W \longrightarrow V$. Dobbiamo solo dimostrare che S è lineare.

Prendiamo w_1 e $w_2 \in W$; essendo T surgettiva, esistono $v_1, v_2 \in V$ tali che si abbia $T(v_1) = w_1$ e $T(v_2) = w_2$. Allora $T(v_1 + v_2) = w_1 + w_2$; applicando S otteniamo $S(w_1 + w_2) = v_1 + v_2 = S(w_1) + S(w_2)$.

Analogamente si dimostra (esercizio) che $S(\lambda w_1) = \lambda S(w_1)$. □

Corollario 7.3

Sia $T: V \to W$ un'applicazione lineare. Se $\dim V = \dim W$, allora le tre affermazioni seguenti sono equivalenti: (1) T è invertibile; (2) T è iniettiva; (3) T è surgettiva.

Dimostrazione

Da un lato, T invertibile implica T iniettiva e surgettiva; poi, per il Corollario 5.8.(3) T iniettiva (rispettivamente, surgettiva) implica T surgettiva (rispettivamente, iniettiva) e le due cose insieme implicano T invertibile. □

In particolare, un endomorfismo $T: V \to V$ è invertibile se e solo se è iniettivo (o surgettivo); quindi per vedere se è invertibile, basta controllare se $\operatorname{Ker} T = \{O\}$.

L'esistenza di un'applicazione $T: V \to W$ lineare invertibile fra due spazi vettoriali V e W implica $\dim V = \dim W$ (perché? Confronta con l'Esercizio 7.8). Di più: ci dice che V e W sono *essenzialmente lo stesso spazio;* qualunque cosa accada in V possiamo, tramite T, trasferirla in W e leggerla lì; e anche viceversa, tramite T^{-1}. Per questo motivo introduciamo la seguente definizione.

Definizione 7.4

Due spazi vettoriali V e W sono *isomorfi* (e si scrive $V \cong W$) se esiste un'*isomorfismo* fra V e W, cioè un'applicazione lineare invertibile $T: V \to W$.

Esempio 7.2

Lo spazio $M_{m,n}(\mathbb{R})$ è isomorfo a $M_{n,m}(\mathbb{R})$. Un isomorfismo è dato dalla trasposizione $^T: M_{m,n}(\mathbb{R}) \to M_{n,m}(\mathbb{R})$; infatti, poiché $(A^T)^T = A$ per qualunque matrice A, si ha che l'inversa della trasposizione è ancora la trasposizione (perché?).

Esempio 7.3

Lo spazio $M_{m,n}(\mathbb{R})$ è anche isomorfo a $\mathbb{R}^{mn}$, tramite l'applicazione $T: M_{m,n}(\mathbb{R}) \to \mathbb{R}^{mn}$ che prende le colonne della matrice e le mette una sotto l'altra

$$T\left(\begin{vmatrix} A^1 & \cdots & A^n \end{vmatrix}\right) = \begin{vmatrix} A^1 \\ \vdots \\ A^n \end{vmatrix}.$$

L'applicazione T è ovviamente lineare e bigettiva, per cui è un isomorfismo.

Esempio 7.4

Consideriamo uno spazio vettoriale reale V di dimensione n; prendiamo una sua base $\mathcal{B} = \{v_1, \dots, v_n\}$, e sia $F_{\mathcal{B}}: V \to \mathbb{R}^n$ l'applicazione lineare definita nell'Esempio 5.7. Allora $F_{\mathcal{B}}$ è un isomorfismo; l'inversa è l'applicazione $S: \mathbb{R}^n \to V$ data da

$$\forall a = \begin{vmatrix} a_1 \\ \vdots \\ a_n \end{vmatrix} \in \mathbb{R}^n \qquad S(a) = a_1 v_1 + \cdots + a_n v_n.$$

Osservazione 7.3

L'esempio precedente ci dice che ogni spazio vettoriale su $\mathbb{R}$ di dimensione n è isomorfo a $\mathbb{R}^n$; quindi perché studiamo gli spazi vettoriali in generale invece di limitarci più semplicemente a $\mathbb{R}^n$? Per almeno due motivi. Prima di tutto, l'isomorfismo sopra descritto non è, come si dice, *naturale* o *canonico*; non è univocamente definito una volta dato lo spazio V; dipende dalla *scelta* di una base. Una volta fissata una base $\mathcal{B}$ hai l'isomorfismo e puoi trasformare tutti i problemi su V in problemi su $\mathbb{R}^n$; ma chi dice che sia la base giusta? Che con un'altra base non sia tutto più semplice? E poi, di solito, uno desidera delle risposte che non dipendono dalla base scelta; l'isomorfismo $F_{\mathcal{B}}$ può essere (e lo è, come vedremo nel prossimo capitolo) molto utile come strumento di calcolo, ma poi alla fine i risultati li vogliamo in V, non in $\mathbb{R}^n$, e li vogliamo indipendenti dalla base scelta. L'isomorfismo dell'Esempio 7.2 è invece di tutt'altro genere: non dipende da alcuna scelta, e ci permette a tutti (o quasi) gli effetti di identificare $M_{m,n}(\mathbb{R})$ e $M_{n,m}(\mathbb{R})$ come spazi vettoriali.

Ma non basta: due spazi vettoriali isomorfi possono essere considerati lo stesso spazio *solo quando ci interessano unicamente la struttura e le operazioni di spazio vettoriale*. Se sono in gioco altre strutture, i due spazi devono rimanere accuratamente distinti. Per esempio, $\mathbb{R}_3[t]$ è isomorfo a $\mathbb{R}^4$. Ma coi polinomi possiamo fare tante cose che su $\mathbb{R}^4$ non hanno molto senso: possiamo dire "quanto vale il polinomio $p(t)$ nel punto $t = 27$" mentre "quanto vale il vettore $v \in \mathbb{R}^4$ nel punto $t = 27$" non significa proprio nulla. Un esempio anche più importante è $\mathcal{L}(V, V)$: mentre la composizione di applicazioni lineari è un'operazione assolutamente naturale e quasi inevitabile (e senza la quale non saremmo qui a porci questo problema), la "composizione" di vettori di $\mathbb{R}^n$ è un concetto un po' forzato.

La morale di tutto ciò è che gli isomorfismi $F_{\mathcal{B}}$ sono degli ottimi e importanti strumenti di calcolo, ma non ci devono far dimenticare che una cosa sono gli spazi vettoriali in generale, e un'altra gli spazi $\mathbb{R}^n$ (o che una cosa sono le applicazioni lineari e un'altra le matrici).

Dopo tutto questo bel discorso vediamo invece due isomorfismi naturali, cioè che non dipendono da alcuna scelta arbitraria.

Proposizione 7.4

(1) *Per ogni $n, m \in \mathbb{N}$ lo spazio $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è isomorfo allo spazio delle matrici $M_{m,n}(\mathbb{R})$. In particolare, lo spazio duale $(\mathbb{R}^n)'$ è isomorfo allo spazio dei vettori riga $M_{1,n}(\mathbb{R})$.*

(2) *Sia V uno spazio vettoriale. Allora lo spazio $\mathcal{L}(\mathbb{R}, V)$ è isomorfo a V.*

Dimostrazione

(1) Sia $L: M_{m,n}(\mathbb{R}) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ l'applicazione che associa alla matrice $A \in M_{m,n}(\mathbb{R})$ l'applicazione lineare $L_A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$; l'idea è che L è l'isomorfismo cercato. Prima di tutto, dobbiamo dimostrare che L è lineare, cioè che $L_{A+B} \equiv L_A + L_B$ e che $L_{\lambda A} \equiv \lambda L_A$. Infatti

$$\begin{aligned} L_{A+B}(x) &= (A+B)x = x_1(A^1 + B^1) + \cdots + x_n(A^n + B^n) \\ &= (x_1 A^1 + \cdots + x_n A^n) + (x_1 B^1 + \cdots + x_n B^n) \\ &= Ax + Bx = L_A(x) + L_B(x) = (L_A + L_B)(x) \end{aligned}$$

per tutti gli $x \in \mathbb{R}^n$, per cui $L_{A+B} \equiv L_A + L_B$. Analogamente,

$$L_{\lambda A}(x) = x_1(\lambda A^1) + \cdots + x_n(\lambda A^n) = \lambda(x_1 A^1 + \cdots + x_n A^n) = \lambda\big(L_A(x)\big) = (\lambda L_A)(x)$$

qualunque sia $\lambda \in \mathbb{R}$, e ci siamo.

Per dimostrare che L è un isomorfismo, vogliamo costruirne l'inversa. Noi sappiamo che un'applicazione lineare è completamente determinata dai valori che assume su una base (Proposizione 5.2); quindi ogni $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è completamente descritta dai vettori $T(e_1), \ldots, T(e_n)$. Consideriamo allora l'applicazione

$$J: \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}(\mathbb{R})$$

che associa all'applicazione lineare $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ la matrice che ha come prima colonna $T(e_1)$, come seconda colonna $T(e_2)$ e così via:

$$J(T) = |T(e_1) \; \cdots \; T(e_n)|.$$

Vogliamo dimostrare che J è l'inversa di L. Grazie alla Proposizione 7.2, ci basta far vedere che $L \circ J$ e $J \circ L$ sono l'identità degli spazi appropriati, senza perder tempo a dimostrare che J è lineare (anche se è facile verificarlo).

Per far vedere che $J \circ L = \mathrm{id}_{M_{m,n}(\mathbb{R})}$ basta dimostrare che la j-esima colonna della matrice $J(L_A)$ è esattamente A^j, quali che siano $A \in M_{m,n}(\mathbb{R})$ e $j = 1, \ldots, n$. E infatti la j-esima colonna di $J(L_A)$ è proprio $L_A(e_j) = A^j$, grazie a (5.4).

Viceversa, prendiamo $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$; grazie al Corollario 5.3, per dimostrare che $(L \circ J)(T) \equiv T$ ci basta far vedere che $L_{J(T)}(e_j) = T(e_j)$ per $j = 1, \ldots, n$. E infatti, di nuovo (5.4) ci dice che $L_{J(T)}(e_j)$ è la j-esima colonna di $J(T)$, ovvero, per definizione, $T(e_j)$.

(2) L'idea della dimostrazione è semplice: un'applicazione lineare $T \in \mathcal{L}(\mathbb{R}, V)$ è completamente determinata dal valore su una base; $\{1\}$ è una base di $\mathbb{R}$, e quindi $T(1)$ determina completamente T. Dunque viene naturale tentare di dimostrare che l'applicazione $\Phi: \mathcal{L}(\mathbb{R}, V) \to V$ definita da $\Phi(T) = T(1)$ è l'isomorfismo cercato.

L'applicazione Φ è chiaramente lineare (esercizio); dimostriamo che è iniettiva e surgettiva. È iniettiva: se $\Phi(T) = O$ abbiamo

$$\forall \lambda \in \mathbb{R} \qquad T(\lambda) = \lambda T(1) = \lambda O = O,$$

per cui $T \equiv O$, cioè $\operatorname{Ker} \Phi = \{O\}$, e l'iniettività segue dalla Proposizione 5.5.(4).

Infine, è surgettiva: preso $v \in V$, definiamo $T_v \in \mathcal{L}(\mathbb{R}, V)$ ponendo $T_v(\lambda) = \lambda v$. Allora

$$\Phi(T_v) = T_v(1) = v,$$

per cui $v \in \operatorname{Im} \Phi$, ed è fatta. □

In particolare, siccome l'applicazione $L: M_{m,n}(\mathbb{R}) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ definita nella dimostrazione della Proposizione 7.4 è bigettiva, *tutte le applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^m$ sono del tipo L_A* per un'opportuna matrice $A \in M_{m,n}(\mathbb{R})$.

Esempio 7.5

La matrice associata all'identità di $\mathbb{R}^n$ deve avere come j-esima colonna $\mathrm{id}_{\mathbb{R}^n}(e_j) = e_j$; quindi otteniamo la matrice *identica*

$$I_n = \begin{vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{vmatrix},$$

dove gli spazi bianchi vanno riempiti di zeri; a volte scriveremo I al posto di I_n. Analogamente, l'applicazione nulla viene associata alla *matrice nulla*

$$O = \begin{vmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{vmatrix}.$$

Osservazione 7.4

Indicheremo gli elementi della matrice identica con δ_{ij}, dove

$$\delta_{ij} = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j. \end{cases}$$

Il simbolo δ_{ij} è chiamato *delta di Kronecker.*

7.2 Prodotto di matrici

Abbiamo quindi visto che le applicazioni lineari fra $\mathbb{R}^n$ ed $\mathbb{R}^m$ sono essenzialmente la stessa cosa delle matrici $m \times n$; quindi deve esistere per le matrici un analogo della composizione fra applicazioni lineari. Per trovarlo, consideriamo due matrici $A \in M_{m,n}(\mathbb{R})$ e $B \in M_{n,p}(\mathbb{R})$, e le relative applicazioni lineari $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ ed $L_B\colon \mathbb{R}^p \to \mathbb{R}^n$. I domini e i codomini sono stati scelti in modo tale da permetterci di considerare la composizione $L_A \circ L_B\colon \mathbb{R}^p \to \mathbb{R}^m$, che è dunque un'applicazione lineare appartenente a $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^m)$. Per quanto visto nel paragrafo precedente, deve quindi esistere una (e una sola) matrice $C \in M_{m,p}(\mathbb{R})$ tale che $L_C = L_A \circ L_B$.

Definizione 7.5

Date due matrici $A \in M_{m,n}(\mathbb{R})$ e $B \in M_{n,p}(\mathbb{R})$, diremo che la matrice $C \in M_{m,p}(\mathbb{R})$ tale che $L_A \circ L_B = L_C$ è il *prodotto (righe per colonne)* di A e B, e scriveremo $C = AB$ (o, talvolta, $C = A \cdot B$). In particolare si ha

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

Prima di controllare se questo prodotto ha le proprietà necessarie per essere considerato tale, vediamo come si calcola esplicitamente. La colonna j-esima C^j della matrice prodotto $C = AB$ dev'essere l'immagine tramite $L_A \circ L_B$ del vettore e_j della base canonica; quindi

$$\begin{vmatrix} c_{1j} \\ \vdots \\ c_{mj} \end{vmatrix} = C^j$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} \begin{vmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{vmatrix} = \begin{vmatrix} a_{11}b_{1j} + \cdots + a_{1n}b_{nj} \\ \vdots \\ a_{m1}b_{1j} + \cdots + a_{mn}b_{nj} \end{vmatrix},$$

per $j = 1, \ldots, p$. In altri termini, l'elemento di posto (i, j) della matrice AB è

$$(AB)_{ij} = c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} = A_iB^j, \tag{7.1}$$

che è il motivo per cui si parla di prodotto *righe per colonne.*

Osservazione 7.5

Perché AB sia definita occorre che il numero di *colonne* di A sia uguale al numero di *righe* di B, in modo che la composizione $L_A \circ L_B$ e la formula (7.1) abbiano senso.

Esempio 7.6

Più che un esempio è un esercizio: verifica che i seguenti prodotti di matrici siano giusti

$$\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} \begin{vmatrix} 3 & 6 & 1 \\ -1 & -1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 3 & 1 \\ 2 & 8 & 2 \end{vmatrix}; \qquad \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix};$$

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ 3 & 10 \end{vmatrix} \neq \begin{vmatrix} 7 & 10 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}.$$

In particolare, quindi, il prodotto di matrici non è commutativo, e moltiplicando due matrici non nulle possiamo ottenere la matrice nulla (vedi anche l'Esercizio 7.23).

Osservazione 7.6

Preso un vettore $x \in \mathbb{R}^n$ considerato come vettore colonna, si vede subito che il prodotto di una matrice $A \in M_{m,n}(\mathbb{R})$ per la colonna $x \in M_{n,1}(\mathbb{R})$ è esattamente quanto abbiamo da sempre indicato con Ax, per cui non c'è rischio di confusione.

Vediamo ora se il prodotto di matrici gode delle proprietà necessarie per giustificare il nome "prodotto". Nell'enunciato seguente, dire che delle matrici hanno "le dimensioni giuste" significa che hanno dimensioni tali da rendere possibili tutte le operazioni (somme e prodotti) considerate.

Proposizione 7.5

Siano A, B e C matrici delle dimensioni giuste, e $\lambda \in \mathbb{R}$. Allora:

(1) $A(B+C) = AB + AC$ e $(A+B)C = AC + BC$;

(2) $(\lambda A)B = \lambda(AB)$;

(3) $(AB)C = A(BC)$;

(4) $AI = IA = A$ e $AO = OA = O$;

(5) $(AB)^T = B^T A^T$ *e (se A e B sono matrici complesse)* $(AB)^H = B^H A^H$.

Dimostrazione

(1), (2) e (3) si dimostrano tutte allo stesso modo; dimostriamo per esempio (3), lasciando le altre due per esercizio (Esercizio 7.13). Siccome l'applicazione L che associa a una matrice A l'applicazione lineare L_A è un isomorfismo (e in particolare è bigettiva), per dimostrare che $(AB)C = A(BC)$ basta far vedere che $L_{(AB)C} \equiv L_{A(BC)}$. Questo segue subito dalla definizione di prodotto e dalle proprietà della composizione di applicazioni lineari: infatti,

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

La (4) segue subito dal fatto che $L_I = \mathrm{id}$ e $L_O = O$; rimane quindi solo da dimostrare (5). Nel caso in cui A sia un vettore riga $a \in M_{1,n}(\mathbb{R})$ e B un vettore colonna $b \in M_{n,1}(\mathbb{R})$ abbiamo

$$(ab)^T = ab = a_1 b_1 + \cdots + a_n b_n = b^T a^T,$$

dove la prima uguaglianza vale in quanto ab è un numero reale[1], per cui la (5) in questo caso è verificata.

---

[1] Anche ba e $a^T b^T$ hanno senso, ma sono delle matrici $n \times n$, l'una trasposta dell'altra.

Vediamo ora il caso generale; poniamo, per semplicità, $C = (AB)^T$ e $D = B^T A^T$. Allora usando (5.2), la definizione di prodotto di matrici e il caso particolare appena visto otteniamo

$$d_{ij} = (B^T)_i (A^T)^j = (B^i)^T (A_j)^T = A_j B^i = (AB)_{ji} = c_{ij},$$

per cui $D = C$, come voluto. La dimostrazione di $(AB)^H = B^H A^H$ è identica. □

7.3 Matrici invertibili

Contrariamente al caso delle matrici rettangolari qualsiasi, il prodotto di due matrici quadrate dello stesso ordine n, cioè appartenenti a $M_{n,n}(\mathbb{R})$, si può sempre fare, e dà un elemento di $M_{n,n}(\mathbb{R})$. La Proposizione 7.5 allora ci dice che con il prodotto righe per colonne e la solita somma l'insieme $M_{n,n}(\mathbb{R})$ diviene[(2)] un anello non commutativo, in cui l'elemento neutro per il prodotto è la matrice identica I_n. L'insieme $M_{n,n}(\mathbb{R})$ però non è mai un campo se $n > 1$, in quanto contiene sempre delle matrici non nulle prive di inversa. Per esempio, se $E_{11} \in M_{n,n}(\mathbb{R})$ è la matrice con 1 al posto $(1, 1)$ e 0 altrove, allora $E_{11}A \neq I_n$ (perché?) per ogni matrice $A \in M_{n,n}(\mathbb{R})$.

Le matrici che ammettono un inverso hanno un nome e proprietà specifiche.

Definizione 7.6

Diremo che una matrice $A \in M_{n,n}(\mathbb{R})$ è *invertibile* se esiste una matrice $B \in M_{n,n}(\mathbb{R})$ tale che $AB = BA = I_n$; la matrice B è l'*inversa* di A. L'inversa, se esiste, è unica: se B' è un'altra inversa abbiamo

$$B' = B'I_n = B'(AB) = (B'A)B = I_n B = B.$$

L'inversa di A, se esiste, si indica con A^{-1}. L'insieme delle matrici invertibili di ordine n si denota con $GL_n(\mathbb{R})$, dove GL è l'abbreviazione di "gruppo lineare".

Proposizione 7.6

Siano $A, B \in GL_n(\mathbb{R})$ due matrici invertibili. Allora A^{-1}, A^T e AB sono invertibili, e si ha

$$(A^{-1})^{-1} = A, \qquad (A^T)^{-1} = (A^{-1})^T \qquad \text{e} \qquad (AB)^{-1} = B^{-1}A^{-1}.$$

Dimostrazione

Il fatto che

$$A^{-1}A = AA^{-1} = I$$

ci dice che A è l'inversa di A^{-1}. Trasponendo la precedente uguaglianza troviamo che $(A^{-1})^T$ è l'inversa di A^T. Infine,

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I = (AB)(B^{-1}A^{-1}),$$

e $B^{-1}A^{-1}$ è l'inversa di AB. □

In particolare, $GL_n(\mathbb{R})$ è un gruppo (rispetto al prodotto, non alla somma: vedi l'Esercizio 7.22) *non commutativo*, come notato nell'Esempio 7.6.

(2) Cosa che non dovrebbe sorprenderti, visto che lo stesso risultato vale per $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ con la somma e la composizione.

Vi è ora un problema fondamentale: quando una matrice $A \in M_{n,n}(\mathbb{R})$ è invertibile? La risposta è contenuta nel seguente teorema e nel successivo corollario.

Teorema 7.7

Sia $A \in M_{n,n}(\mathbb{R})$. Allora le seguenti affermazioni sono equivalenti:

(1) *A è invertibile;*

(2) *L_A è invertibile;*

(3) *L_A è iniettiva;*

(4) *L_A è surgettiva;*

(5) $\operatorname{rg} A = n$;

(6) *le colonne di A sono linearmente indipendenti;*

(7) *le righe di A sono linearmente indipendenti;*

(8) *il sistema omogeneo $Ax = O$ ha come unica soluzione $x = O$;*

(9) *per ogni $b \in \mathbb{R}^n$ il sistema $Ax = b$ ha come unica soluzione $x = A^{-1}b$;*

(10) *i pivot di A, comunque determinati, sono non nulli (cioè A è non singolare).*

Osservazione 7.7

Per dimostrare l'equivalenza di queste affermazioni faremo vedere che la (1) implica la (9), che a sua volta implica la (10), che a sua volta implica la (8), e così via fino a tornare al punto di partenza, chiudendo il cerchio. Di conseguenza ciascuna di queste affermazioni implicherà tutte le altre, per cui saranno equivalenti.

Dimostrazione

(1) $\Longrightarrow$ (9). Se A è invertibile, allora il vettore $v^o = A^{-1}b$ è soluzione del sistema: infatti abbiamo $Av^o = AA^{-1}b = Ib = b$. Inoltre è l'unica soluzione: se v è soluzione, si ha $Av = b$ e quindi, applicando A^{-1} a entrambi i membri, $v = A^{-1}b = v^o$.

(9) $\Longrightarrow$ (10). Teorema 3.3.

(10) $\Longrightarrow$ (8). Segue di nuovo dal Teorema 3.3, applicato al sistema $Ax = O$.

(8) $\Longrightarrow$ (6). Proposizione 4.4.

(6) $\Longrightarrow$ (7). Proposizione 5.11.(2).

(7) $\Longrightarrow$ (5). Di nuovo Proposizione 5.11.(2).

(5) $\Longrightarrow$ (4). Corollario 5.8.(2).

(4) $\Longrightarrow$ (3). Corollario 5.8.(3).

(3) $\Longrightarrow$ (2). Corollario 7.3.

(2) $\Longrightarrow$ (1). L'inversa di L_A dev'essere della forma L_B per qualche $B \in M_{n,n}(\mathbb{R})$. È allora chiaro (perché?) che B è l'inversa di A. □

Corollario 7.8

Sia $A \in M_{n,n}(\mathbb{R})$. Allora le seguenti affermazioni sono equivalenti:

(1) *A è invertibile;*

(2) *esiste una matrice $B_1 \in M_{n,n}(\mathbb{R})$ tale che $B_1A = I_n$;*

(3) *esiste una matrice $B_2 \in M_{n,n}(\mathbb{R})$ tale che $AB_2 = I_n$.*

Dimostrazione

(1) $\Longrightarrow$ (2), (3). Basta prendere $B_1 = B_2 = A^{-1}$.

(2) $\Longrightarrow$ (1). Supponiamo che $v \in \mathbb{R}^n$ sia tale che $L_A(v) = O$. Allora (2) ci dà

$$O = L_{B_1}(O) = L_{B_1}\big(L_A(v)\big) = L_{B_1A}(v) = L_{I_n}(v) = v,$$

per cui $\operatorname{Ker} L_A = \{O\}$, cioè L_A è iniettiva e A è invertibile.

(3) $\Longrightarrow$ (1). Prendiamo $v \in \mathbb{R}^n$, e sia $w = L_{B_2}(v)$. Allora

$$L_A(w) = L_A\big(L_{B_2}(v)\big) = L_{AB_2}(v) = L_{I_n}(v) = v,$$

per cui L_A è surgettiva, e A è invertibile. □

Osservazione 7.8

Nel Capitolo 9 vedremo che tutte queste affermazioni equivalgono anche a richiedere che il determinante della matrice sia diverso da zero.

Rimane il problema, una volta determinato che $A \in M_{n,n}(\mathbb{R})$ è invertibile, di come calcolarne l'inversa. Il caso $n = 2$ è trattato nell'Esercizio 7.19; per $n > 2$ giunge in aiuto l'eliminazione di Gauss. Prima di tutto, grazie al Corollario 7.8 la matrice A è invertibile se e solo se esiste una matrice $X \in M_{n,n}(\mathbb{R})$ tale che $AX = I_n$. Ora, dire che $AX = I_n$ equivale a dire che $AX^1 = e_1, \ldots, AX^n = e_n$; quindi per trovare X dobbiamo risolvere simultaneamente i sistemi lineari quadrati

$$Ax = e_1, \quad \ldots, \quad Ax = e_n, \tag{7.2}$$

che hanno tutti la stessa matrice dei coefficienti. La soluzione del primo sistema sarà la prima colonna di A^{-1}, la soluzione del secondo sistema sarà la seconda colonna di A^{-1} e così via.

L'eliminazione di Gauss ci fornisce una tecnica ideale per risolvere un problema di questo genere; infatti, le operazioni necessarie per ridurre in forma triangolare superiore i sistemi (7.2) dipendono solo dalla matrice A, e quindi possono essere effettuate una volta sola per tutti i sistemi considerati. Vediamo questo procedimento in azione su un esempio, in cui sarà indicato anche come concludere il calcolo dell'inversa senza bisogno di risolvere all'indietro i sistemi (7.2) ma usando in modo opportuno operazioni elementari.

Esempio 7.7

Vogliamo determinare l'inversa (se esiste) della matrice

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

Dobbiamo risolvere i tre sistemi $Ax = e_1$, $Ax = e_2$ e $Ax = e_3$. Applichiamo l'eliminazione di Gauss ad A e alla matrice identica (che contiene i termini noti dei sistemi che ci interessano)

$$\left|\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 & 1 & 0 \\ -2 & 2 & 1 & 0 & 0 & 1 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & 3 & 2 & 1 & 0 & 1 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & 0 & -4 & -5 & 3 & 1 \end{array}\right|.$$

Quindi la matrice A è non singolare, e in particolare è invertibile. A questo punto, invece di risolvere all'indietro i tre sistemi triangolari superiori che abbiamo ottenuto, agiamo con delle operazioni elementari in modo analogo a quanto si fa nell'eliminazione di Gauss, ma partendo dalla terza riga in basso e salendo

$$\left|\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & 0 & -4 & -5 & 3 & 1 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 2 & 1 & 0 & -1/4 & 3/4 & 1/4 \\ 0 & -1 & 0 & 1/2 & -1/2 & -1/2 \\ 0 & 0 & -4 & -5 & 3 & 1 \end{array}\right|$$

$$\longrightarrow \left|\begin{array}{ccc|ccc} 2 & 0 & 0 & 1/4 & 1/4 & -1/4 \\ 0 & -1 & 0 & 1/2 & -1/2 & -1/2 \\ 0 & 0 & -4 & -5 & 3 & 1 \end{array}\right| = \left| S \,\middle|\, T \right|.$$

La matrice diagonale S così ottenuta è la matrice dei coefficienti di tre sistemi ancora equivalenti ai tre di partenza; i termini noti di questi nuovi sistemi sono contenuti nella matrice T. Quindi se dividiamo ciascuna riga di T per il corrispondente pivot otteniamo le soluzioni dei tre sistemi di partenza, per cui la matrice inversa è data da

$$A^{-1} = \begin{vmatrix} 1/8 & 1/8 & -1/8 \\ -1/2 & 1/2 & 1/2 \\ 5/4 & -3/4 & -1/4 \end{vmatrix}.$$

Controlla, per esercizio, che $A \cdot A^{-1} = A^{-1} \cdot A = I_3$.

L'Esercizio 9.18 descriverà un altro modo per calcolare l'inversa di una matrice.

Esercizi 7

7.1 Siano $S: U \to V$ e $T: V \to W$ due applicazioni lineari fra spazi vettoriali. Dimostra che $\operatorname{rg}(T \circ S) \leq \min\{\operatorname{rg} S, \operatorname{rg} T\}$.

7.2 Siano $T: \mathbb{R}^3 \to \mathbb{R}_2[t]$ ed $S: \mathbb{R}_2[t] \to \mathbb{R}^2$ le applicazioni lineari date da

$$T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = x + (2y - x)t + zt^2 \qquad \text{e} \qquad S(p) = \begin{vmatrix} p(0) \\ p(1) \end{vmatrix}$$

Calcola $S \circ T: \mathbb{R}^3 \to \mathbb{R}^2$ e trova $\operatorname{Ker}(S \circ T)$ e $\operatorname{Im}(S \circ T)$.

7.3 Sia $T: V \to V$ un endomorfismo tale che $T \circ T = O$. Dimostra che $T + \mathrm{id}_V$ è invertibile.

7.4 Scrivi un'applicazione lineare da $M_{3,3}(\mathbb{R})$ a $M_{2,2}(\mathbb{R})$. Trova un isomorfismo fra $M_{3,4}(\mathbb{R})$ e $M_{6,2}(\mathbb{R})$, se esiste.

7.5 Sia $T_a: \mathbb{R}_2[t] \to \mathbb{R}^3$ l'applicazione lineare data da

$$\forall p \in \mathbb{R}_2[t] \qquad\qquad T_a(p) = \begin{vmatrix} p(0) \\ p(a) \\ p(1) \end{vmatrix},$$

dove $a \in \mathbb{R}$. Trova per quali valori di a l'applicazione T_a è un isomorfismo.

7.6 Dimostra che un sistema lineare quadrato $Ax = b$ ammette soluzione quale che sia b se e solo se il sistema omogeneo $Ax = O$ ammette solo la soluzione $x = O$.

7.7 Un endomorfismo $P \in \mathcal{L}(V, V)$ tale che $P \circ P = P$ si chiama *proiezione*. Dimostra che se P è una proiezione allora $V = \operatorname{Ker} P \oplus \operatorname{Im} P$. (*Suggerimento:* prova a scrivere $v = \big(v - P(v)\big) + P(v)$.)

7.8 Dimostra che due spazi vettoriali di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione.

7.9 Siano U e W due sottospazi di uno spazio vettoriale V tali che $U \cap W = \{O\}$. Dimostra che $U \oplus W$ è isomorfo al prodotto cartesiano $U \times W$ (vedi l'Esercizio 4.26).

7.10 Dato $\lambda \in \mathbb{R}$, quale matrice viene associata all'applicazione $T_\lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ definita da $T_\lambda(x) = \lambda x$?

7.11 Sia $T: V \to W$ un'applicazione lineare, e siano $\mathcal{B}$ una base di V e $\mathcal{C}$ una base di W. Dimostra che per ogni $v \in V$ le coordinate di $T(v)$ rispetto alla base $\mathcal{C}$ si esprimono come polinomi di primo grado senza termini noti nelle coordinate di v rispetto alla base $\mathcal{B}$. (*Suggerimento:* considera $F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1}$.)

7.12 Verifica che se $a \in \mathbb{R}^n$ l'applicazione $\varphi_a\colon \mathbb{R}^n \to \mathbb{R}$ definita nell'Esempio 5.9 è data da $\varphi_a(x) = a^T \cdot x$. Utilizza questa formula per ridimostrare il Lemma 5.10.

7.13 Dimostra la Proposizione 7.5.(1) e (2).

7.14 Calcola tutti i prodotti possibili fra le matrici seguenti:

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix}, \quad \begin{vmatrix} -2 & 3 \\ 1 & 7 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{vmatrix}.$$

7.15 Data la matrice:

$$A = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix},$$

trova tutte le matrici $B \in M_{2,2}(\mathbb{R})$ tali che $AB = BA$.

7.16 Data la matrice

$$A = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 0 & 1 \end{vmatrix} \in M_{3,3}(\mathbb{R}),$$

calcola la potenza n-esima di A (cioè il prodotto di A per se stessa n volte, che si indica con A^n) per ogni $n \in \mathbb{N}$.

7.17 Data la matrice:

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

considera l'insieme $W = \{X \in M_{2,2}(\mathbb{R}) \mid AX = O\}$. Dimostra che W è un sottospazio vettoriale di $M_{2,2}(\mathbb{R})$, calcolane la dimensione e trovane una base.

7.18 Trova per quali $k \in \mathbb{R}$ esiste una matrice $X \in M_{3,2}(\mathbb{R})$ tale che:

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

7.19 Sia

$$A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \in M_{2,2}(\mathbb{R}).$$

Dimostra che A è invertibile se e solo se $ad - bc \neq 0$, e che l'inversa è data da

$$A^{-1} = \frac{1}{ad - bc} \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}.$$

(*Suggerimento:* vedi l'Esercizio 4.15.)

7.20 Calcola l'inversa (se esiste) delle seguenti matrici:

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

7.21 Trova per quali $h \in \mathbb{R}$ la seguente matrice:

$$\begin{vmatrix} h & 1 & 0 \\ 1 & h & h \\ 0 & 1 & 2 \end{vmatrix}$$

è invertibile, e calcolane quando possibile l'inversa.

7.22 Trova due matrici quadrate invertibili la cui somma sia una matrice non invertibile (per cui $GL_n(\mathbb{R})$ non è un gruppo rispetto alla somma).

7.23 Sia $A \in M_{n,n}(\mathbb{R})$. Dimostra che A non è invertibile se e solo se esiste una matrice $B \in M_{n,n}(\mathbb{R})$ non nulla tale che $AB = O$.

7.24 Sia V un sottospazio di $\mathbb{R}^n$ di equazioni parametriche $v = At$, dove $t \in \mathbb{R}^k$ per cui $A \in M_{n,k}(\mathbb{R})$, e sia $B \in M_{n-\mathrm{rg}A,n}(\mathbb{R})$ una matrice le cui righe (trasposte) sono una base di $\operatorname{Ker} A^T$. Dimostra che $Bx = O$ sono equazioni cartesiane per V.

7.25 Siano $A, B \in M_{n,n}(\mathbb{R})$; dimostra che $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, dove $\mathrm{tr}\,A$ indica la traccia della matrice A (definita nell'Esercizio 5.10), e deduci che $\mathrm{tr}\,(B^{-1}AB) = \mathrm{tr}\,A$ per ogni $A \in M_{n,n}(\mathbb{R})$ e $B \in GL_n(\mathbb{R})$.

7.26 Dimostra che date due matrici $A, B \in M_{n,n}(\mathbb{R})$ si ha sempre $AB - BA \neq I_n$. (*Suggerimento*: usa l'esercizio precedente.)

7.27 Dimostra che la proiezione definita nell'Esempio 5.8 è una proiezione anche nel senso dell'Esercizio 7.7.

Cambiamenti di base

8

Sommario

In questo capitolo studieremo più in dettaglio l'isomorfismo $F_{\mathcal{B}}: V \to \mathbb{K}^n$ fra uno spazio vettoriale V di dimensione n sul campo $\mathbb{K}$ e lo spazio $\mathbb{K}^n$, isomorfismo che associa a ogni vettore $v \in V$ le sue coordinate rispetto a una base $\mathcal{B}$. Vedremo come muta $F_{\mathcal{B}}$ cambiando la base $\mathcal{B}$; come si associa una matrice a un'applicazione lineare $T: V \to W$ qualsiasi una volta fissate una base di V e una di W; e vedremo come muta questa matrice cambiando le basi in V e W. Come conseguenza ricaveremo delle tecniche efficienti per risolvere in qualunque spazio vettoriale (di dimensione finita) i problemi che nel Paragrafo 6.3 avevamo risolto in $\mathbb{K}^n$.

8.1 Matrice di cambiamento di base

Nel Paragrafo 6.3 abbiamo risolto tutta una serie di problemi di calcolo per $\mathbb{K}^n$; gli stessi problemi rimangono però ancora aperti in spazi vettoriali qualunque. In particolare, se V è uno spazio vettoriale e $v_1, \ldots, v_k \in V$, come si trovano dimensione e base del sottospazio $\mathrm{Span}(v_1, \ldots, v_k)$? E dimensione e base di somma e intersezione di sottospazi? Se $T: V \to W$ è un'applicazione lineare, come si trovano il rango di T e una base di $\mathrm{Im}\, T$? E dimensione e base di $\mathrm{Ker}\, T$?

La risposta consiste nel "ricondursi al problema precedente" (frase proverbialmente associata ai matematici). L'idea è che, come discusso nell'Osservazione 7.3, ogni spazio vettoriale V di dimensione n sul campo $\mathbb{K}$ è isomorfo a $\mathbb{K}^n$; quindi per risolvere qualunque problema su V che riguardi solo la struttura di spazio vettoriale possiamo trasferire il problema in $\mathbb{K}^n$, risolverlo lì e ritrasportare indietro la soluzione, in modo da avere la risposta in V. Questo trasferimento si fa sostituendo ai vettori le loro coordinate rispetto a una base fissata $\mathcal{B}$, cioè usando l'isomorfismo $F_\mathcal{B}: V \to \mathbb{K}^n$ definito nell'Esempio 5.7. Un paio di esempi chiariranno le idee.

Esempio 8.1

Sia $V = \mathbb{R}_2[t]$, e consideriamo $p_1(t) = t - 1$, $p_2(t) = t^2 - t$ e $p_3(t) = 2t^2 + 4t - 6$. Vogliamo dimensione e base di $W = \mathrm{Span}(p_1, p_2, p_3) \subseteq V$. Il procedimento sarà: fissata una base $\mathcal{B}$ dello spazio vettoriale V, troveremo la dimensione d e una base $\{v_1, \ldots, v_d\}$ di $\mathrm{Span}(F_\mathcal{B}(p_1), F_\mathcal{B}(p_2), F_\mathcal{B}(p_3)) \subseteq \mathbb{R}^3$. Allora, essendo $F_\mathcal{B}$ un isomorfismo, la dimensione di W sarà d, e una base sarà $\{(F_\mathcal{B})^{-1}(v_1), \ldots, (F_\mathcal{B})^{-1}(v_d)\}$. Nel nostro caso, scegliamo come base $\mathcal{B} = \{1, t, t^2\}$. Allora

$$F_\mathcal{B}(p_1) = \begin{vmatrix} -1 \\ 1 \\ 0 \end{vmatrix}, \qquad F_\mathcal{B}(p_2) = \begin{vmatrix} 0 \\ -1 \\ 1 \end{vmatrix}, \qquad F_\mathcal{B}(p_3) = \begin{vmatrix} -6 \\ 4 \\ 2 \end{vmatrix}$$

sono le coordinate di p_1, p_2 e p_3 nella base $\mathcal{B}$. Con le tecniche viste nel Paragrafo 6.3 troviamo subito che la dimensione di $\mathrm{Span}(F_\mathcal{B}(p_1), F_\mathcal{B}(p_2), F_\mathcal{B}(p_3))$ è 2, e che una sua base è $\{v_1 = F_\mathcal{B}(p_1), v_2 = F_\mathcal{B}(p_2)\}$; quindi la dimensione di $\mathrm{Span}(p_1, p_2, p_3) \subset V$ è 2, e una base è data da $\{p_1, p_2\}$. ⊔⊓

Esempio 8.2

Sia adesso $V = \mathbb{R}_3[t]$, e consideriamo i polinomi $q_1(t) = t^2 - 1$ e $q_2(t) = t^3 + t$. Poniamo $U = \mathrm{Span}(q_1, q_2)$ e $W = \mathrm{Span}(p_1, p_2, p_3)$, dove p_1, p_2 e p_3 sono i polinomi dell'esempio precedente. Vogliamo dimensione e una base di $U + W$ e di $U \cap W$. Utilizziamo la base $\mathcal{B} = \{1, t, t^2, t^3\}$; rispetto a questa base, i polinomi che ci interessano hanno coordinate

$$F_\mathcal{B}(p_1) = \begin{vmatrix} -1 \\ 1 \\ 0 \\ 0 \end{vmatrix}, \quad F_\mathcal{B}(p_2) = \begin{vmatrix} 0 \\ -1 \\ 1 \\ 0 \end{vmatrix}, \quad F_\mathcal{B}(p_3) = \begin{vmatrix} -6 \\ 4 \\ 2 \\ 0 \end{vmatrix},$$

$$F_\mathcal{B}(q_1) = \begin{vmatrix} -1 \\ 0 \\ 1 \\ 0 \end{vmatrix}, \quad F_\mathcal{B}(q_2) = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \end{vmatrix}.$$

Sappiamo che una base di W è $\{p_1, p_2\}$, per cui possiamo tralasciare p_3. Quindi, siccome grazie all'Esercizio 8.1 si ha $F_{\mathcal{B}}(U+W) = F_{\mathcal{B}}(U) + F_{\mathcal{B}}(W)$, per studiare $U+W$ e $U \cap W$ dobbiamo considerare la matrice

$$A = \begin{vmatrix} -1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Una veloce riduzione a scala mostra che $\operatorname{rg} A = 3$ e che una base di $\operatorname{Im} A$ è data da $\{A^1, A^2, A^4\}$; quindi le tecniche studiate nel Paragrafo 6.3 ci dicono subito che $\dim(F_{\mathcal{B}}(U) + F_{\mathcal{B}}(W)) = 3$, $\dim(F_{\mathcal{B}}(U) \cap F_{\mathcal{B}}(W)) = 1$, e che $\{F_{\mathcal{B}}(p_1), F_{\mathcal{B}}(p_2), F_{\mathcal{B}}(q_2)\}$ è una base di $F_{\mathcal{B}}(U) + F_{\mathcal{B}}(W)$. Ne segue che $\dim(U+W) = 3$, che $\dim U \cap W = 1$ e che $\{p_1, p_2, q_2\}$ è una base di $U + W$. Infine sempre le tecniche del Paragrafo 6.3 mostrano che $\{F_{\mathcal{B}}(q_1)\}$ è una base di $F_{\mathcal{B}}(U) \cap F_{\mathcal{B}}(W)$, per cui $\{q_1\}$ è una base di $U \cap W$. ⊓⊔

Dunque adesso sappiamo come risolvere i primi due problemi (calcolo di dimensione e base di uno span, e dell'intersezione e/o somma di sottospazi) in qualunque spazio vettoriale di dimensione finita. Ovviamente, i conti che devono essere effettuati dipendono dalla scelta di una base.

Esempio 8.3

Sia $V = \mathbb{R}_2[t]$ e $p_1, p_2, p_3 \in V$ gli stessi polinomi dell'Esempio 8.1; questa volta prendiamo come base $\mathcal{B}' = \{1, t-1, 2t^2+4t-6\}$. Ovviamente, bisogna prima verificare che $\mathcal{B}'$ sia effettivamente una base di $\mathbb{R}_2[t]$; ma è sufficiente osservare che $F_{\mathcal{B}}(1)$, $F_{\mathcal{B}}(t-1)$ e $F_{\mathcal{B}}(2t^2+4t-6)$ sono linearmente indipendenti in $\mathbb{R}^3$ (esercizio). Le coordinate dei tre polinomi rispetto a $\mathcal{B}'$ sono date da

$$F_{\mathcal{B}'}(p_1) = \begin{vmatrix} 0 \\ 1 \\ 0 \end{vmatrix}, \qquad F_{\mathcal{B}'}(p_2) = \begin{vmatrix} 0 \\ -3 \\ 1/2 \end{vmatrix}, \qquad F_{\mathcal{B}'}(p_3) = \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix},$$

per cui senza bisogno di altri conti vediamo subito che $\dim \operatorname{Span}(p_1, p_2, p_3) = 2$ e che una base è data da $\{p_1, p_3\}$. ⊓⊔

Diventa quindi interessante vedere che cosa succede in generale alle coordinate di un vettore cambiando base. Sia V uno spazio vettoriale di dimensione n sul campo $\mathbb{K}$, e fissiamo due basi $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{v'_1, \ldots, v'_n\}$ di V. Questo ci fornisce due isomorfismi $F_{\mathcal{B}}, F_{\mathcal{B}'} \colon V \to \mathbb{K}^n$, ciascuno dei quali associa a un vettore $v \in V$ le sue coordinate rispetto alla corrispondente base. Abbiamo quindi la seguente situazione:

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \qquad & v & \longmapsto & v \\ {\scriptstyle F_{\mathcal{B}'}}\downarrow & & \downarrow{\scriptstyle F_{\mathcal{B}}} & & \downarrow & & \downarrow \\ \mathbb{K}^n & & \mathbb{K}^n & & \text{Coordinate di } v \text{ rispetto a } \mathcal{B}' & & \text{Coordinate di } v \text{ rispetto a } \mathcal{B} \end{array}$$

A questo punto viene la tentazione di chiudere il quadrato tracciando una freccia nel lato inferiore; si tratta di trovare un'applicazione lineare da $\mathbb{K}^n$ in $\mathbb{K}^n$ che trasformi le coordinate rispetto a $\mathcal{B}'$ in coordinate rispetto a $\mathcal{B}$. Tale applicazione esiste, ed è unica: si tratta per forza di

$$\Psi = F_{\mathcal{B}} \circ (F_{\mathcal{B}'})^{-1}.$$

Il diagramma completo

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle F_{\mathcal{B}'}}\downarrow & & \downarrow{\scriptstyle F_{\mathcal{B}}} \\ \mathbb{K}^n & \xrightarrow{\Psi} & \mathbb{K}^n \end{array}$$

è un esempio di *diagramma commutativo*: in qualunque ordine si seguono le frecce, si ottiene lo stesso risultato. In questo caso vuol dire che (partendo dall'angolo in alto a sinistra) andare prima a destra e poi scendere è lo stesso di prima scendere e poi andare a destra; infatti

$$F_{\mathcal{B}} \circ \mathrm{id}_V = F_{\mathcal{B}} = \Psi \circ F_{\mathcal{B}'}.$$

Ora, $\Psi \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$; per la Proposizione 7.4, deve esistere una matrice quadrata B tale che $\Psi = L_B$; essendo Ψ invertibile (l'inversa è $F_{\mathcal{B}'} \circ (F_{\mathcal{B}})^{-1}$), anche la matrice B è invertibile.

Dunque ogni volta che abbiamo due basi $\mathcal{B}$ e $\mathcal{B}'$ di uno spazio vettoriale V troviamo una matrice invertibile B che trasforma le coordinate rispetto a $\mathcal{B}'$ nelle coordinate rispetto a $\mathcal{B}$. In concreto, preso un vettore $v \in V$, siano $x = F_{\mathcal{B}}(v) \in \mathbb{K}^n$ le sue coordinate rispetto a $\mathcal{B}$, e $x' = F_{\mathcal{B}'}(v) \in \mathbb{K}^n$ le sue coordinate rispetto a $\mathcal{B}'$. Allora si ha

$$x = Bx', \tag{8.1}$$

cioè

$$\begin{cases} x_1 = b_{11}x'_1 + \cdots + b_{1n}x'_n, \\ \qquad\vdots \\ x_n = b_{n1}x'_1 + \cdots + b_{nn}x'_n. \end{cases}$$

Nota che

$$x' = B^{-1}x,$$

per cui la matrice B^{-1} ci permette di recuperare le coordinate rispetto a $\mathcal{B}'$ partendo da quelle rispetto a $\mathcal{B}$.

Si tratta ora di vedere come si trova la matrice B. La formula (5.4) ci dice che la colonna h-esima di B è data da

$$B^h = L_B(e_h) = \Psi(e_h) = \big(F_{\mathcal{B}} \circ (F_{\mathcal{B}'})^{-1}\big)(e_h) = F_{\mathcal{B}}\big((F_{\mathcal{B}'})^{-1}(e_h)\big) = F_{\mathcal{B}}(v'_h),$$

cioè è data dalle coordinate di v'_h rispetto alla base $\mathcal{B}$. In altre parole, la matrice B *contiene per colonne le coordinate dei vettori della nuova base $\mathcal{B}'$ rispetto alla vecchia base $\mathcal{B}$*. In simboli,

$$B = \left| F_{\mathcal{B}}(v'_1) \cdots F_{\mathcal{B}}(v'_n) \right|.$$

Definizione 8.1

La matrice B viene chiamata *matrice di cambiamento di base* (o *di passaggio*) *da* $\mathcal{B}$ *a* $\mathcal{B}'$. Chiaramente, B^{-1} è la matrice di cambiamento di base da $\mathcal{B}'$ a $\mathcal{B}$, cioè nella direzione inversa.

Osservazione 8.1

Dire che B contiene per colonne le coordinate dei vettori della nuova base $\mathcal{B}'$ rispetto alla vecchia base $\mathcal{B}$ è equivalente (perché?) a scrivere

$$\left|v_1' \cdots v_n'\right| = \left|v_1 \cdots v_n\right| B, \tag{8.2}$$

dove il prodotto è il solito prodotto righe per colonne. Quindi la matrice B trasforma i vettori della *vecchia* base $\mathcal{B}$ nei vettori della *nuova* base $\mathcal{B}'$, ed è per questo che la chiamiamo matrice di passaggio *da $\mathcal{B}$ a $\mathcal{B}'$*. Ricordati, però, che la matrice B trasforma invece le *coordinate* rispetto alla base $\mathcal{B}'$ nelle coordinate rispetto alla base $\mathcal{B}$, grazie a (8.1); i vettori vanno in un verso, le coordinate nel verso opposto.

Esempio 8.4

Prendiamo $V = \mathbb{R}_2[t]$, $\mathcal{B} = \{1, t, t^2\}$ e $\mathcal{B}' = \{1, t-1, 2t^2 + 4t - 6\}$. Allora la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$ contiene per colonne le coordinate dei polinomi di $\mathcal{B}'$ rispetto alla base $\mathcal{B}$, cioè

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

Le matrici di cambiamento di base compaiono, ovviamente, anche nel caso di sottospazi vettoriali di $\mathbb{R}^n$.

Esempio 8.5

Sia $V \subset \mathbb{R}^3$ lo spazio descritto nell'Esempio 4.12, e consideriamo le due basi $\mathcal{B} = \{v_1, v_2\}$ e $\mathcal{B}' = \{e_1, e_2\}$, dove i vettori $v_1, v_2 \in V$ sono dati da

$$v_1 = \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix} \quad \text{e} \quad v_2 = \begin{vmatrix} -1 \\ -1 \\ 0 \end{vmatrix}.$$

Allora

$$\begin{cases} e_1 = \frac{1}{2} v_1 + \frac{1}{2} v_2, \\ e_2 = -\frac{1}{2} v_1 - \frac{3}{2} v_2, \end{cases}$$

per cui la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$ è data da

$$B = \begin{vmatrix} 1/2 & -1/2 \\ 1/2 & -3/2 \end{vmatrix}.$$

Analogamente,

$$\begin{cases} v_1 = 3e_1 + e_2, \\ v_2 = -e_1 - e_2, \end{cases}$$

per cui la matrice di cambiamento di base da $\mathcal{B}'$ a $\mathcal{B}$ è data da

$$B' = \begin{vmatrix} 3 & -1 \\ 1 & -1 \end{vmatrix}.$$

È facile verificare che $BB' = B'B = I_2$.

In generale, per trovare la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$ bisogna calcolare tutte le coordinate dei vettori di $\mathcal{B}'$ rispetto a $\mathcal{B}$, operazione che richiede di solito la soluzione di tanti sistemi lineari quant'è la dimensione dello spazio considerato.

Ci sono però alcuni casi particolari in cui la situazione è più semplice. Per esempio, se siamo in $\mathbb{K}^n$ e la base rispetto alla quale vogliamo trovare le coordinate è la base canonica, o siamo in $M_{m,n}(\mathbb{K})$ e la base è quella data nell'Esercizio 4.17. In tali casi, infatti, il calcolo delle coordinate si riduce (perché?) a trascrivere in colonna gli elementi del vettore o della matrice in un ordine opportuno. Analogamente, se lo spazio vettoriale che stiamo considerando è $\mathbb{R}_n[t]$ e la base rispetto alla quale calcoliamo le coordinate è $\{1, t, \ldots, t^n\}$, per trovare le coordinate basta scrivere ordinatamente i coefficienti del polinomio in una colonna di $\mathbb{R}^{n+1}$. Insomma, queste basi ci piacciono molto perché rispetto a esse il calcolo delle coordinate è particolarmente facile.

E se lo spazio vettoriale V è uno di questi, ma le basi $\mathcal{B}$ e $\mathcal{B}'$ sono altre, che facciamo? Ormai dovresti conoscere abbastanza la mentalità dei matematici da immaginare che, ancora una volta, "ci riconduciamo al caso precedente". Questo vuol dire che scegliamo una via che sembrerebbe complicarci la vita, mentre in realtà ci fornisce un metodo rapido per il calcolo della matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$.

Consideriamo una terza base $\mathcal{B}_0$ (che nella nostra mente sarà sempre una base rispetto a cui il calcolo delle coordinate sia facile) e proviamo a spezzare il calcolo della matrice di passaggio in due momenti successivi: passiamo prima da $\mathcal{B}$ a $\mathcal{B}_0$, e poi da $\mathcal{B}_0$ a $\mathcal{B}'$. Utilizzando di nuovo i diagrammi commutativi possiamo sintetizzare questo procedimento nel seguente schema

$$\begin{array}{ccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle F_{\mathcal{B}'}}\big\downarrow & & \big\downarrow{\scriptstyle F_{\mathcal{B}_0}} & & \big\downarrow{\scriptstyle F_{\mathcal{B}}} \\ \mathbb{K}^n & \xrightarrow{L_C} & \mathbb{K}^n & \xrightarrow{L_D} & \mathbb{K}^n \end{array}, \qquad \begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle F_{\mathcal{B}'}}\big\downarrow & & \big\downarrow{\scriptstyle F_{\mathcal{B}}} \\ \mathbb{K}^n & \xrightarrow{L_B} & \mathbb{K}^n \end{array}$$

dove C è la matrice di passaggio da $\mathcal{B}_0$ a $\mathcal{B}'$, D quella da $\mathcal{B}$ a $\mathcal{B}_0$, e B quella da $\mathcal{B}$ a $\mathcal{B}'$. Il diagramma a sinistra descrive il passaggio da $\mathcal{B}$ a $\mathcal{B}_0$ seguito dal passaggio da $\mathcal{B}_0$ a $\mathcal{B}'$, mentre il diagramma a destra descrive il passaggio diretto da $\mathcal{B}$ a $\mathcal{B}'$. Cerchiamo adesso di capire qual è (se esiste) il legame fra le tre matrici B, C e D.

L'applicazione in basso nel diagramma a destra, L_B, è l'*unica* che può completare il diagramma; ma se mettiamo insieme le applicazioni in basso nel diagramma a sinistra ne otteniamo una seconda (nota che gli altri tre lati del diagramma a destra corrispondono esattamente ai tre lati esterni del diagramma a sinistra). L'unica possibilità è dunque che le due applicazioni coincidano: quindi $L_B = L_D \circ L_C$, ovvero

$$B = DC.$$

Ricordiamoci che abbiamo scelto *ad hoc* la base $\mathcal{B}_0$ in modo che la matrice C si scriva immediatamente. Sfortunatamente, non è altrettanto semplice scrivere D, in quanto stavolta $\mathcal{B}_0$ è la base di arrivo e non quella di partenza. Sappiamo però scrivere facilmente l'inversa A di D, che è la matrice di passaggio da $\mathcal{B}_0$ a $\mathcal{B}$.

Riassumendo, per trovare B possiamo allora adottare la procedura seguente: scriviamo le matrici C e $A = D^{-1}$ che sono rispettivamente le matrici di passaggio da $\mathcal{B}_0$ a $\mathcal{B}'$ e da $\mathcal{B}_0$ a $\mathcal{B}$, e calcoliamo $B = A^{-1}C$.

L'Esempio 7.7 ci suggerisce come possiamo procedere in modo efficiente. Per calcolare $B = A^{-1}C$ basta risolvere i sistemi lineari $AB^1 = C^1, \ldots, AB^n = C^n$, dove $C^1, \ldots, C^n$ sono le colonne di C. Basta quindi scrivere le matrici A e C una accanto all'altra, $|A|C|$, e operare con una riduzione a scala e con operazioni elementari, come già visto nell'Esempio 7.7, fino a ottenere $|I_n|A^{-1}C|$, trovando quindi a destra la matrice $B = A^{-1}C$ cercata.

Illustriamo questo metodo con un esempio, anche per convincerci che tanta astrazione non è inutile ma permette davvero di fare i conti più rapidamente.

Esempio 8.6

Sia $V = \mathbb{R}_2[t]$ e siano $p_1(t) = t$, $p_2(t) = t^2 + 1$, $p_3(t) = 1 + t$, $q_1(t) = t^2$, $q_2(t) = t^2 + t$, e $q_3(t) = 1 - t$. Vogliamo verificare che $\mathcal{B} = \{p_1, p_2, p_3\}$ e $\mathcal{B}' = \{q_1, q_2, q_3\}$ sono basi di $\mathbb{R}_2[t]$, e trovare la matrice B di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$ utilizzando il metodo sopra descritto. Cominciamo scrivendo le coordinate dei vettori di $\mathcal{B}$ e $\mathcal{B}'$ rispetto alla base $\mathcal{B}_0 = \{1, t, t^2\}$ di $\mathbb{R}_2[t]$ ottenendo le matrici A di cambiamento di base da $\mathcal{B}_0$ a $\mathcal{B}$ e C di cambiamento di base da $\mathcal{B}_0$ a $\mathcal{B}'$

$$A = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \quad \text{e} \quad C = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{vmatrix}.$$

La riduzione a scala mostra facilmente che sia A sia C hanno rango 3, e quindi $\mathcal{B}$ e $\mathcal{B}'$ sono entrambe basi di $\mathbb{R}_2[t]$. Per trovare B procediamo come annunciato

$$\left|\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & -1 \end{array}\right|$$

$$\longrightarrow \left|\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & -2 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 1 & -1 \end{array}\right| \longrightarrow \left|\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & -2 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right|,$$

per cui

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

8.2 Matrice associata a un'applicazione lineare

Torniamo ora ai problemi posti all'inizio; dobbiamo vedere come risolvere quelli relativi alle applicazioni lineari. Siano V, W spazi vettoriali, con $\dim V = n$ e $\dim W = m$, e $T: V \to W$ un'applicazione lineare; vogliamo associare a T un'applicazione lineare di $\mathbb{K}^n$ in $\mathbb{K}^m$ (ovvero una matrice) da usare per rispondere a tutti i quesiti che potremmo porci su T. L'idea è che la matrice deve fare alle coordinate dei vettori ciò che T fa ai vettori stessi.

Fissiamo allora una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V e una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ di W. Questa volta abbiamo il seguente diagramma:

$$\begin{array}{ccccccc} V & \xrightarrow{T} & W & & v & \longmapsto & T(v) \\ {\scriptstyle F_B}\downarrow & & \downarrow{\scriptstyle F_C} & & \downarrow & & \downarrow \\ \mathbb{K}^n & & \mathbb{K}^m & & \text{Coordinate di } v & & \text{Coordinate di } T(v) \\ & & & & \text{rispetto a } \mathcal{B} & & \text{rispetto a } \mathcal{C} \end{array}$$

Anche questa volta possiamo chiudere il quadrato: se poniamo

$$\Phi = F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1}, \tag{8.3}$$

il diagramma

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle F_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle F_{\mathcal{C}}} \\ \mathbb{K}^n & \xrightarrow{\Phi} & \mathbb{K}^m \end{array}$$

diventa commutativo, cioè

$$F_{\mathcal{C}} \circ T = \Phi \circ F_{\mathcal{B}}.$$

Ora, $\Phi \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$; quindi la Proposizione 7.4 ci dice che deve esistere una (e una sola) matrice $A \in M_{m,n}(\mathbb{K})$ tale che $\Phi = L_A$.

Definizione 8.2

La matrice A è la matrice *associata* a T (o che *rappresenta* T) *rispetto alle basi* $\mathcal{B}$ e $\mathcal{C}$.

Dunque se per ogni $v \in V$ indichiamo con $x \in \mathbb{K}^n$ le coordinate di v rispetto a $\mathcal{B}$, il vettore $y = Ax \in \mathbb{K}^m$ deve contenere le coordinate di $T(v)$ rispetto alla base $\mathcal{C}$. Per calcolare esplicitamente la matrice A, basta notare che, siccome le coordinate di v_j rispetto a $\mathcal{B}$ sono date dal vettore e_j, abbiamo

$$A^j = L_A(e_j) = F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1}(e_j) = F_{\mathcal{C}}\big(T(v_j)\big),$$

per cui la j-esima colonna A^j di A contiene le coordinate di $T(v_j)$ rispetto a $\mathcal{C}$. In altri termini, la matrice A *contiene per colonne le coordinate rispetto alla base di arrivo* $\mathcal{C}$ *dei trasformati secondo* T *dei vettori della base di partenza* $\mathcal{B}$. In simboli

$$\big|T(v_1) \cdots T(v_n)\big| = |w_1 \cdots w_m| A.$$

Esempio 8.7

Prendiamo $V = \mathbb{R}_2[t]$, $W = \mathbb{R}_3[t]$ e consideriamo l'applicazione lineare $T\colon V \longrightarrow W$ data da $p(t) \mapsto t^2 p'(t+1)$, dove l'apice indica la derivata rispetto a t. Per esempio, se $p(t) = t^2$, allora $p'(t) = 2t$, $p'(t+1) = 2(t+1)$ e

$$[T(p)](t) = t^2 p'(t+1) = 2t^2(t+1) = 2t^3 + 2t^2.$$

Prendiamo come basi le solite $\mathcal{B} = \{1, t, t^2\}$ e $\mathcal{C} = \{1, t, t^2, t^3\}$; essendo $\dim V = 3$ e $\dim W = 4$, la matrice A associata a T sarà una matrice 4×3. Per trovarla, calcoliamo i trasformati dei vettori di $\mathcal{B}$

$$T(1) = 0, \qquad T(t) = t^2, \qquad T(t^2) = 2t^3 + 2t^2.$$

Scrivendo le coordinate dei trasformati rispetto alla base $\mathcal{C}$ otteniamo

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

Esempio 8.8

Prendiamo $A \in M_{m,n}(\mathbb{R})$. Allora si vede facilmente che A è la matrice che rappresenta l'applicazione $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ rispetto alle basi canoniche di $\mathbb{R}^n$ ed $\mathbb{R}^m$ (perché?). Se invece prendiamo basi non canoniche, L_A può essere rappresentata da una matrice diversa. Per esempio, sia $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ data da

$$T\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} 2x + 2z \\ x - y \end{vmatrix}.$$

Allora $T = L_A$, dove

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

Prendiamo come basi

$$\mathcal{B}' = \left\{ \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}, \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}, \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix} \right\} \quad \text{e} \quad \mathcal{C}' = \left\{ \begin{vmatrix} 1 \\ 1 \end{vmatrix}, \begin{vmatrix} 2 \\ 3 \end{vmatrix} \right\}.$$

Allora

$$T\begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix} = \begin{vmatrix} 4 \\ 1 \end{vmatrix} = 10\begin{vmatrix} 1 \\ 1 \end{vmatrix} - 3\begin{vmatrix} 2 \\ 3 \end{vmatrix}, \qquad T\begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix} = \begin{vmatrix} 0 \\ -2 \end{vmatrix} = 4\begin{vmatrix} 1 \\ 1 \end{vmatrix} - 2\begin{vmatrix} 2 \\ 3 \end{vmatrix},$$

$$T\begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \end{vmatrix} = -2\begin{vmatrix} 1 \\ 1 \end{vmatrix} + \begin{vmatrix} 2 \\ 3 \end{vmatrix},$$

per cui la matrice associata a T rispetto alle basi $\mathcal{B}'$ e $\mathcal{C}'$ è

$$A' = \begin{vmatrix} 10 & 4 & -2 \\ -3 & -2 & 1 \end{vmatrix}.$$

A questo punto abbiamo gli strumenti necessari per rispondere a tutte le domande iniziali. Infatti, $F_{\mathcal{C}} \circ T = L_A \circ F_{\mathcal{B}}$ ci dice esattamente (perché?) che

$$\operatorname{Im} T = (F_{\mathcal{C}})^{-1}(\operatorname{Im} A) \qquad \text{e} \qquad \operatorname{Ker} T = (F_{\mathcal{B}})^{-1}(\operatorname{Ker} A);$$

quindi possiamo recuperare tutte le informazioni che cerchiamo esaminando A.

Esempio 8.9

Siano V, W e $T: V \to W$ come nell'Esempio 8.7; vogliamo dimensioni e basi di $\operatorname{Ker} T$ e $\operatorname{Im} T$. Prima di tutto, si vede subito che il rango della matrice A che rappresenta T rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$ è 2; quindi, essendo $F_{\mathcal{C}}$ e $F_{\mathcal{B}}$ isomorfismi

$$\operatorname{rg} T = \dim \operatorname{Im} T = \dim \operatorname{Im} A = \operatorname{rg} A = 2 \quad \text{e} \quad \dim \operatorname{Ker} T = \dim \operatorname{Ker} A = 3 - 2 = 1.$$

Siccome le ultime due colonne di A sono indipendenti, esse formano una base di $\operatorname{Im} A$; dunque i trasformati degli ultimi due elementi di $\mathcal{B}$, cioè $\{t^2, 2t^3 + 2t^2\}$, compongono una base di $\operatorname{Im} T$. Infine, la prima colonna di A è nulla, cioè $Ae_1 = O$; quindi il primo vettore di $\mathcal{B}$ forma una base di $\operatorname{Ker} T$, la base $\{1\}$.

Possiamo formalizzare quanto ottenuto con la seguente proposizione.

Proposizione 8.1

Siano V e W spazi vettoriali sul campo $\mathbb{K}$, con $\dim V = n$ e $\dim W = m$. Fissiamo una base $\mathcal{B}$ di V e una base $\mathcal{C}$ di W. Allora l'applicazione $T \mapsto A$ che associa a un'applicazione lineare $T: V \to W$ la matrice che la rappresenta rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$ è un isomorfismo fra $\mathcal{L}(V, W)$ e $M_{m,n}(\mathbb{K})$. In particolare,

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

Dimostrazione

Basta dimostrare che l'associazione $T \mapsto L_A$ è un isomorfismo fra $\mathcal{L}(V, W)$ ed $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$; infatti la Proposizione 7.4 ci dice già che $L_A \mapsto A$ è un isomorfismo fra $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ e $M_{m,n}(\mathbb{K})$, e la composizione di isomorfismi è ancora un isomorfismo (esercizio).

Definiamo $\Phi: \mathcal{L}(V, W) \to \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ ponendo

$$\Phi(T) = L_A = F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1}.$$

L'applicazione Φ è chiaramente lineare; per esempio

$$\Phi(T + S) = F_{\mathcal{C}} \circ (T + S) \circ (F_{\mathcal{B}})^{-1} = F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1} + F_{\mathcal{C}} \circ S \circ (F_{\mathcal{B}})^{-1} = \Phi(T) + \Phi(S).$$

Rimane da dimostrare che Φ è invertibile. L'inversa $\Psi: \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) \to \mathcal{L}(V, W)$ è

$$\Psi(L_A) = (F_{\mathcal{C}})^{-1} \circ L_A \circ F_{\mathcal{B}}.$$

Infatti,

$$\Psi \circ \Phi(T) = (F_{\mathcal{C}})^{-1} \circ \Phi(T) \circ F_{\mathcal{B}} = (F_{\mathcal{C}})^{-1} \circ F_{\mathcal{C}} \circ T \circ (F_{\mathcal{B}})^{-1} \circ F_{\mathcal{B}} = T,$$

e analogamente $\Phi \circ \Psi(L_A) = L_A$. □

Ovviamente, la matrice associata a un'applicazione lineare dipende fortemente dalle basi scelte; cambiando basi, cambia la matrice.

Esempio 8.10

Siano V, W e T come nell'Esempio 8.7, ma stavolta prendiamo come base di V l'insieme $\mathcal{B}' = \{t + 1, t^2 + 1, t^2 + t\}$; lasciamo invariata la base di W. Questa volta i trasformati della base di partenza sono

$$T(t + 1) = t^2, \qquad T(t^2 + 1) = 2t^3 + 2t^2, \qquad T(t^2 + t) = 2t^3 + 3t^2,$$

per cui la matrice associata è

$$A' = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & 3 \\ 0 & 2 & 2 \end{vmatrix}.$$

Vediamo di capire come la matrice associata dipende dalle basi. Siano V e W due spazi vettoriali con $\dim V = n$ e $\dim W = m$. Scegliamo due basi $\mathcal{B}, \mathcal{B}'$ di V, due basi $\mathcal{C}, \mathcal{C}'$ di W, e sia $T: V \to W$ un'applicazione lineare. Indichiamo con B la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$, con C la matrice di cambiamento di base da $\mathcal{C}$ a $\mathcal{C}'$ (in modo che C^{-1} sia la matrice di cambiamento di base da $\mathcal{C}'$ a $\mathcal{C}$), con A la matrice che rappresenta T rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$, e con A' la matrice che rappresenta T rispetto alle basi $\mathcal{B}'$ e $\mathcal{C}'$. Mettendo insieme quanto visto finora, otteniamo i seguenti diagrammi commutativi

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{T} & W & \xrightarrow{\mathrm{id}_W} & W \\ {\scriptstyle F_{\mathcal{B}'}}\downarrow & & {\scriptstyle F_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle F_{\mathcal{C}}} & & \downarrow{\scriptstyle F_{\mathcal{C}'}} \\ \mathbb{K}^n & \xrightarrow{L_B} & \mathbb{K}^n & \xrightarrow{L_A} & \mathbb{K}^n & \xrightarrow{(L_C)^{-1}} & \mathbb{K}^m \end{array}, \qquad \begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle F_{\mathcal{B}'}}\downarrow & & \downarrow{\scriptstyle F_{\mathcal{C}'}} \\ \mathbb{K}^n & \xrightarrow{L_{A'}} & \mathbb{K}^m \end{array}.$$

Come abbiamo già visto nel paragrafo precedente, l'applicazione in basso nel diagramma a destra, $L_{A'}$, è l'*unica* che può completare il diagramma; ma se mettiamo insieme le applicazioni in basso nel diagramma a sinistra ne otteniamo una seconda (nota che gli altri tre lati del diagramma a destra corrispondono esattamente ai tre lati esterni del diagramma a sinistra). L'unica possibilità è dunque che le due applicazioni coincidano: quindi si deve avere $L_{A'} = (L_C)^{-1} \circ L_A \circ L_B$, ovvero

$$A' = C^{-1}AB. \tag{8.4}$$

Questo è il modo con cui cambia la matrice associata a T al cambiare delle basi.

Questa tecnica consistente nel procedere con i diagrammi è molto efficace (abbiamo finito senza fare alcun conto), ma forse poco intuitiva. Vediamo allora di ottenere lo stesso risultato in un altro modo. Prendiamo $v \in V$, e indichiamo con $x, x' \in \mathbb{K}^n$ le sue coordinate rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$. Analogamente, indichiamo con $y, y' \in \mathbb{K}^m$ le coordinate di $T(v)$ rispetto alle basi $\mathcal{C}$ e $\mathcal{C}'$. Allora noi sappiamo che

$$x = Bx', \qquad y = Ax, \qquad y' = A'x', \qquad y = Cy'.$$

Quindi

$$A'x' = y' = C^{-1}y = C^{-1}Ax = C^{-1}ABx';$$

siccome questo deve valere per ogni $x' \in \mathbb{K}^n$, otteniamo di nuovo (8.4).

Esempio 8.11

Verifichiamo la teoria nell'esempio che abbiamo usato più volte in questo paragrafo (vedi gli Esempi 8.7, 8.9 e 8.10). Con le notazioni introdotte abbiamo

$$A = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{vmatrix}; \quad A' = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & 3 \\ 0 & 2 & 2 \end{vmatrix}, \quad B = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix}, \quad C = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix};$$

è facile quindi verificare che $A' = C^{-1}AB$.

Esempio 8.12

Rivediamo ora l'Esempio 8.8. In questo caso $\mathcal{B}$ e $\mathcal{C}$ sono le basi canoniche di $\mathbb{R}^3$ ed $\mathbb{R}^2$, per cui

$$A = \begin{vmatrix} 2 & 0 & 2 \\ 1 & -1 & 0 \end{vmatrix}, \quad A' = \begin{vmatrix} 10 & 4 & -2 \\ -3 & -2 & 1 \end{vmatrix}, \quad B = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{vmatrix}, \quad C = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix},$$

e di nuovo $A' = C^{-1}AB$.

Un caso particolarmente interessante è $V = W$, ovvero quando $T\colon V \to V$ è un endomorfismo. In tal caso, tanto vale prendere la stessa base $\mathcal{B}$ in partenza e in arrivo, per cui si parla di *matrice associata* a T (o che *rappresenta* T) *rispetto alla base* $\mathcal{B}$. In particolare, se A è associata a T rispetto alla base $\mathcal{B}$, A' rispetto alla base $\mathcal{B}'$, e B è la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$, abbiamo

$$A' = B^{-1}AB. \tag{8.5}$$

Definizione 8.3

Due matrici quadrate $A, A' \in M_{n,n}(\mathbb{K})$ dello stesso ordine si dicono *simili* se esiste una matrice $B \in GL_n(\mathbb{K})$ tale che $A' = B^{-1}AB$ (e quindi $A = BA'B^{-1} = (B^{-1})^{-1}A'B^{-1}$).

Proposizione 8.2

Sia $T\colon V \to V$ un endomorfismo, $\mathcal{B}$ una base di V e A la matrice associata a T rispetto a $\mathcal{B}$. Allora:

(1) *se $\mathcal{B}'$ è un'altra base di V, la matrice A' associata a T rispetto a $\mathcal{B}'$ è simile ad A;*

(2) *viceversa, se A' è una matrice simile ad A, allora esiste una base $\mathcal{B}'$ di V tale che A' rappresenta T rispetto a $\mathcal{B}'$.*

Dimostrazione

(1) È l'uguaglianza (8.5).

(2) Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ e prendiamo $B \in GL_n(\mathbb{K})$ tale che $A' = B^{-1}AB$. Definiamo $\mathcal{B}' = \{v_1', \ldots, v_n'\}$ tramite la

$$|v_1' \cdots v_n'| = |v_1 \cdots v_n|B,$$

ovvero

$$\begin{cases} v_1' = b_{11}v_1 + \cdots + b_{n1}v_n, \\ \quad\vdots \\ v_n' = b_{1n}v_1 + \cdots + b_{nn}v_n. \end{cases}$$

Se $\mathcal{B}'$ è una base di V ci siamo: infatti in tal caso B è proprio la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$, e quindi A' rappresenta T rispetto a $\mathcal{B}'$.

Per dimostrare che $\mathcal{B}'$ è una base, basta dimostrare che $v_1', \ldots, v_n'$ sono linearmente indipendenti. Ma infatti da

$$O = \alpha_1 v_1' + \cdots + \alpha_n v_n' = |v_1' \cdots v_n'| \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{vmatrix} = |v_1 \cdots v_n| B \begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{vmatrix}$$

deduciamo

$$B\begin{vmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{vmatrix} = O,$$

in quanto $v_1, \ldots, v_n$ sono linearmente indipendenti, e quindi $\alpha_1 = \cdots = \alpha_n = 0$, perché B è invertibile. □

Definizione 8.4

La *classe di similitudine* $\mathcal{O}_A$ di una matrice $A \in M_{n,n}(\mathbb{K})$ è

$$\mathcal{O}_A = \{B^{-1}AB \mid B \in GL_n(\mathbb{K})\}.$$

Dunque a ogni endomorfismo $T\colon V \to V$ possiamo associare in maniera unica una classe di similitudine che contiene tutte e sole le matrici che possono rappresentare T rispetto a una qualche base di V. I Capitoli 13 e 14 saranno dedicati, fra le altre cose, al problema di trovare in questa classe di similitudine una matrice particolarmente semplice (per esempio diagonale), da cui sia molto facile leggere le proprietà di T.

Osservazione 8.2

Due matrici A, $A' \in M_{n,n}(\mathbb{K})$ sono simili se e solo se il sistema lineare omogeneo $BA' - AB = O$, le cui incognite sono le n^2 componenti della matrice B, ammette una soluzione in cui B risulta invertibile (perché?).

Esercizi 8

8.1 Sia $T\colon V \to W$ un'applicazione lineare, e U_1, U_2 due sottospazi vettoriali di V. Dimostra che $T(U_1 + U_2) = T(U_1) + T(U_2)$.

8.2 Considera i tre vettori

$$v_1 = \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}, \quad v_2 = \begin{vmatrix} 1 \\ -1 \\ 0 \end{vmatrix}, \quad v_3 = \begin{vmatrix} 3 \\ 1 \\ 5 \end{vmatrix}.$$

Dimostra che $\mathcal{B} = \{v_1, v_2, v_3\}$ è una base di $\mathbb{R}^3$, e trova la matrice di cambiamento di base dalla base $\mathcal{B}$ alla base canonica.

8.3 Considera i polinomi $p_1(t) = t^2 - 2t, p_2(t) = 1 + 2t, p_3(t) = 2 - t^2, q_1(t) = -1 + t$, $q_2(t) = -1 + t - t^2$ e $q_3(t) = 2t + 2t^2$. Dimostra che $\mathcal{B} = \{p_1, p_2, p_3\}$ e $\mathcal{C} = \{q_1, q_2, q_3\}$ sono basi di $\mathbb{R}_2[t]$, e trova la matrice di passaggio da $\mathcal{B}$ a $\mathcal{C}$.

8.4 Considera le matrici $X_1 = \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix}$, $X_2 = \begin{vmatrix} 4 & 3 \\ 4 & 3 \end{vmatrix}$, $X_3 = \begin{vmatrix} 5 & 1 \\ -2 & 3 \end{vmatrix}$, $X_4 = \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix}$, $Y_1 = \begin{vmatrix} 1 & 1 \\ 3 & 0 \end{vmatrix}$, $Y_2 = \begin{vmatrix} 2 & 4 \\ 0 & 1 \end{vmatrix}$, $Y_3 = \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix}$, e $Y_4 = \begin{vmatrix} 7 & 0 \\ 1 & 2 \end{vmatrix}$. Dimostra che $\mathcal{B} = \{X_1, X_2, X_3, X_4\}$ e $\mathcal{C} = \{Y_1, Y_2, Y_3, Y_4\}$ sono basi di $M_{2,2}(\mathbb{R})$ e trova la matrice di passaggio da $\mathcal{B}$ a $\mathcal{C}$.

8.5 Considera i polinomi $p_1(t) = -2it + t^2, p_2(t) = -1 + it^2, p_3(t) = 3i - (1 + i)t + t^2$, $q_1(t) = 2i - t^2$, $q_2(t) = 2 - i + it + (1 + 3i)t^2$ e $q_3(t) = i + (3 - 2i)t + t^2$. Dimostra che $\mathcal{B} = \{p_1, p_2, p_3\}$ e $\mathcal{C} = \{q_1, q_2, q_3\}$ sono basi di $\mathbb{C}_2[t]$ e trova la matrice di passaggio da $\mathcal{B}$ a $\mathcal{C}$.

8.6 Considera $v_1 = \begin{vmatrix} 1 \\ 2 \\ 4 \end{vmatrix}$, $v_2 = \begin{vmatrix} 1 \\ 5 \\ 5 \end{vmatrix}$, $v_3 = \begin{vmatrix} 3 \\ 3 \\ 1 \end{vmatrix}$, $u_1 = \begin{vmatrix} 1 \\ 5 \\ 6 \end{vmatrix}$, $u_2 = \begin{vmatrix} -3 \\ 7 \\ 2 \end{vmatrix}$, $u_3 = \begin{vmatrix} 1 \\ 4 \\ 1 \end{vmatrix}$. Dimostra che $\mathcal{B} = \{v_1, v_2, v_3\}$ e $\mathcal{C} = \{u_1, u_2, u_3\}$ sono basi di $\mathbb{R}^3$ e trova la matrice di passaggio da $\mathcal{B}$ a $\mathcal{C}$.

8.7 Presi $v_1 = \begin{vmatrix} 1+i \\ 2-i \\ 3 \end{vmatrix}$, $v_2 = \begin{vmatrix} i \\ 1 \\ 1+3i \end{vmatrix}$, $v_3 = \begin{vmatrix} 1 \\ i \\ -i \end{vmatrix}$, $u_1 = \begin{vmatrix} 1 \\ 1+2i \\ 1 \end{vmatrix}$, $u_2 = \begin{vmatrix} i \\ 3 \\ i \end{vmatrix}$, $u_3 = \begin{vmatrix} 1 \\ 4-i \\ 2-5i \end{vmatrix}$, dimostra che $\mathcal{B} = \{v_1, v_2, v_3\}$ e $\mathcal{C} = \{u_1, u_2, u_3\}$ sono basi di $\mathbb{C}^3$ e trova la matrice di passaggio da $\mathcal{B}$ a $\mathcal{C}$.

8.8 Considera i polinomi $p_1(t) = 3, p_2(t) = 2 + t^3, p_3(t) = t - t^2 - 4t^3, p_4(t) = t^2 - t^3$, e $p_5(t) = t + 2t^2$. Estrai, se possibile, da $\{p_1, \ldots, p_5\}$ una base di $\mathbb{R}_3[t]$.

8.9 Sia $W \subseteq M_{2,2}(\mathbb{R})$ il sottospazio

$$W = \operatorname{Span}\left(\begin{vmatrix}1 & 1\\ 0 & 0\end{vmatrix}, \begin{vmatrix}0 & 1\\ 0 & 1\end{vmatrix}, \begin{vmatrix}0 & 0\\ 1 & 1\end{vmatrix}, \begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix}\right).$$

Calcola la dimensione di W e trovane una base.

8.10 Considera i sottospazi $U = \operatorname{Span}(p_1, p_2)$ e $W = \operatorname{Span}(q_2, q_3)$ di $\mathbb{R}_2[t]$, dove i polinomi p_1, p_2, q_2 e q_3 sono gli stessi dell'Esercizio 8.3. Calcola dimensione e base di $U + W$ e $U \cap W$.

8.11 Sia $S\colon \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare rappresentata rispetto alle basi canoniche dalla matrice

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

Trova la matrice $A' \in M_{2,3}(\mathbb{R})$ che rappresenta S rispetto alle basi

$$\mathcal{C}_1 = \left\{\begin{vmatrix}1\\0\\0\end{vmatrix}, \begin{vmatrix}1\\1\\0\end{vmatrix}, \begin{vmatrix}1\\1\\1\end{vmatrix}\right\} \subset \mathbb{R}^3 \qquad \text{e} \qquad \mathcal{C}_2 = \left\{\begin{vmatrix}1\\0\end{vmatrix}, \begin{vmatrix}1\\1\end{vmatrix}\right\} \subset \mathbb{R}^2.$$

8.12 Siano A, $B \in M_{3,3}(\mathbb{R})$ le matrici

$$A = \begin{vmatrix}0 & 1 & 1\\ 2 & 3 & 7\\ 1 & -3 & -1\end{vmatrix}, \qquad B = \begin{vmatrix}2 & 0 & 2\\ -1 & 3 & -1\\ 1 & -1 & 3\end{vmatrix}.$$

Scrivi le matrici associate agli endomorfismi L_A, $L_B\colon \mathbb{R}^3 \to \mathbb{R}^3$ rispetto alle due basi seguenti:

$$\mathcal{B}_1 = \left\{\begin{vmatrix}-3\\3\\1\end{vmatrix}, \begin{vmatrix}4\\2\\-1\end{vmatrix}, \begin{vmatrix}0\\1\\-1\end{vmatrix}\right\}, \qquad \mathcal{B}_2 = \left\{\begin{vmatrix}1\\1\\0\end{vmatrix}, \begin{vmatrix}0\\1\\1\end{vmatrix}, \begin{vmatrix}1\\0\\1\end{vmatrix}\right\}.$$

8.13 Sia $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorfismo che è rappresentato rispetto alla base canonica dalla matrice

$$A = \begin{vmatrix}1 & 3\\ 3 & 5\end{vmatrix}.$$

Trova la matrice che rappresenta T rispetto alla base

$$\mathcal{B} = \left\{\begin{vmatrix}1\\3\end{vmatrix}, \begin{vmatrix}3\\5\end{vmatrix}\right\}.$$

8.14 Siano $v_1, v_2, v_3, w_1, w_2, w_3 \in \mathbb{R}^3$ i vettori

$$v_1 = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}, \quad v_2 = \begin{vmatrix} 0 \\ 1 \\ -1 \end{vmatrix}, \quad v_3 = \begin{vmatrix} 0 \\ 0 \\ 2 \end{vmatrix}, \quad w_1 = \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix}, \quad w_2 = \begin{vmatrix} -1 \\ 0 \\ 2 \end{vmatrix}, \quad w_3 = \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}.$$

Dimostra che $\mathcal{B} = \{v_1, v_2, v_3\}$ è una base di $\mathbb{R}^3$, verifica che esiste un unic endomorfismo $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ tale che $T(v_j) = w_j$ per $j = 1, 2, 3$, e trova la matric associata a T rispetto alla base $\mathcal{B}$ e rispetto alla base canonica.

8.15 Sia $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ la base di $M_{2,2}(\mathbb{R})$ descritta nell'Esercizio 4.17, e si $T: M_{2,2}(\mathbb{R}) \to M_{2,2}(\mathbb{R})$ l'endomorfismo dato da

$$T(E_{11}) = T(E_{22}) = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}, \qquad T(E_{12}) = T(E_{21}) = \begin{vmatrix} -1 & -1 \\ -1 & -1 \end{vmatrix}.$$

Calcola dimensione e base di $\operatorname{Ker} T$ e $\operatorname{Im} T$.

8.16 Sia $T: \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare data da

$$T \begin{vmatrix} x \\ y \end{vmatrix} = \begin{vmatrix} x - y \\ y - x \end{vmatrix}.$$

Trova $A \in M_{2,2}(\mathbb{R})$ che *non* rappresenta T rispetto ad alcuna base di $\mathbb{R}^2$.

8.17 Trova una base $\mathcal{B}$ del sottospazio vettoriale $W \subset \mathbb{R}^4$ descritto dall'equazione cartesiana $x_1 + 2x_2 - x_3 = 0$. Sia poi $T: W \to \mathbb{R}^4$ l'applicazione lineare data da

$$T \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{vmatrix} = \begin{vmatrix} x_1 + x_2 + 2x_3 \\ x_1 - x_3 + x_4 \\ 3x_1 + x_2 + 2x_4 \\ x_2 - x_4 \end{vmatrix}.$$

Dimostra che $\operatorname{Im} T \subseteq W$, per cui possiamo considerare T come un endomorfismo di W; trova la matrice che rappresenta questo endomorfismo rispetto alla base $\mathcal{B}$.

8.18 Data la matrice

$$A = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix},$$

considera l'applicazione $T: M_{2,2}(\mathbb{R}) \to M_{2,2}(\mathbb{R})$ data da $T(X) = AX - XA$. Dimostra che T è lineare, calcola $\operatorname{Ker} T$, $\operatorname{Im} T$ e dimostra che $M_{2,2}(\mathbb{R}) = \operatorname{Ker} T \oplus \operatorname{Im} T$.

8.19 Data la matrice $A = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}$, sia $T: M_{2,2}(\mathbb{R}) \to M_{2,2}(\mathbb{R})$ l'endomorfismo dato da $T(X) = AX - XA$. Calcola il rango di T, e trova dimensione e base di $\operatorname{Ker} T$ e $\operatorname{Im} T$.

8.20 Verifica che

$$\left\{\begin{vmatrix}1 & 0\\1 & i\end{vmatrix}, \begin{vmatrix}2 & 0\\1 & 0\end{vmatrix}, \begin{vmatrix}3 & 0\\1 & 1+\sqrt{5}i\end{vmatrix}, \begin{vmatrix}0 & 2+i\\4 & 2-i\end{vmatrix}\right\}$$

è una base $\mathcal{B}$ di $M_{2,2}(\mathbb{C})$. Presa $A = \begin{vmatrix}2 & -1\\-3i & 1+3i\end{vmatrix} \in M_{2,2}(\mathbb{C})$, considera l'endomorfismo lineare S da $M_{2,2}(\mathbb{C})$ in sé dato da $S(X) = AX - X^T$. Scrivi la matrice associata a S rispetto alla base $\mathcal{B}$ e trova nucleo e immagine di S.

8.21 Data la matrice

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

considera l'endomorfismo $T\colon M_{3,3}(\mathbb{R}) \to M_{3,3}(\mathbb{R})$ dato da $T(X) = XA$. Trova il rango di T e base e dimensione del nucleo di T.

8.22 Sia $T\colon \mathbb{R}_2[t] \to \mathbb{R}_2[t]$ l'applicazione lineare data da $T(p) = tp' - 2p$, dove l'apice indica la derivata del polinomio. Trova la matrice che rappresenta T rispetto alle basi $\mathcal{B}$, $\mathcal{C}$ dell'Esercizio 8.3, e rispetto alla base $\mathcal{D} = \{1, t, t^2\}$; trova poi dimensione e base di $\operatorname{Ker} T$ e $\operatorname{Im} T$.

8.23 Considera l'applicazione lineare $T : \mathbb{R}_2[t] \to \mathbb{R}^3$ data da

$$T(p) = \begin{vmatrix}p'(0)\\p(1)-p(2)\\p(4)-p(6)\end{vmatrix},$$

dove l'apice indica la derivata prima. Scrivi la matrice associata a T rispetto alla base $\{1, t, t^2\}$ di $\mathbb{R}_2[t]$ e alla base canonica di $\mathbb{R}^3$ e trova nucleo e immagine di T.

8.24 Considera l'applicazione lineare $T : \mathbb{C}_3[t] \to \mathbb{C}_3[t]$ data da

$$T(p) = \begin{vmatrix}ip'(t)\\(1+i)p(0)t - p''(0)t^2\\p(t) - tp'(t)\\(2-7i)p''(t^2)\end{vmatrix},$$

dove l'apice indica la derivata prima e il doppio apice la derivata seconda. Scrivi la matrice associata a T rispetto alla base $\{1, t, t^2, t^3\}$ di $\mathbb{C}_3[t]$ e trova nucleo e immagine di T.

8.25 Scopri se le seguenti coppie di matrici sono composte da matrici simili:

$$\begin{vmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 3\end{vmatrix} \quad \text{e} \quad \begin{vmatrix}1 & 0 & 0\\0 & 3 & 0\\0 & 0 & 2\end{vmatrix}; \qquad \begin{vmatrix}1 & 2 & 3\\2 & 3 & 1\\3 & 1 & 2\end{vmatrix} \quad \text{e} \quad \begin{vmatrix}1 & 2 & 3\\2 & 3 & 4\\3 & 5 & 7\end{vmatrix}.$$

$\boxed{8.26}$ Sia $T: V \to W$ un'applicazione lineare di rango r. Dimostra che esistono una base $\mathcal{B}$ di V e una base $\mathcal{C}$ di W tali che la matrice associata a T rispetto a queste basi è della forma

$$\begin{vmatrix} I_r & O \\ O & O \end{vmatrix}.$$

$\boxed{8.27}$ Trova un endomorfismo $T: \mathbb{R}^2 \to \mathbb{R}^2$ tale che la matrice

$$\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}$$

non può essere associata a T rispetto ad alcuna base di $\mathbb{R}^2$.

$\boxed{8.28}$ Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale V. Supponiamo che esista una base $\mathcal{B}$ di V tale che la matrice associata a T rispetto a $\mathcal{B}$ sia la matrice identica. Dimostra che $T = \mathrm{id}_V$.

$\boxed{8.29}$ Sia $P: V \to V$ una proiezione (nel senso dell'Esercizio 7.7) di rango r. Dimostra che esiste una base di V rispetto a cui P è rappresentata dalla matrice

$$\begin{vmatrix} I_r & O \\ O & O \end{vmatrix}.$$

$\boxed{8.30}$ Sia $V = \{A \in M_{2,2}(\mathbb{C}) \mid BA = O\}$, dove $B = \begin{vmatrix} 1 & i \\ i & -1 \end{vmatrix}$. Trova una base $\mathcal{B}$ di V. Indicata con $T: M_{2,2}(\mathbb{C}) \to M_{2,2}(\mathbb{C})$ l'applicazione che scambia le due colonne, dimostra che $T(V) = V$ e trova la matrice che rappresenta $T|_V$ rispetto a $\mathcal{B}$.

Determinanti

9

Sommario

In questo capitolo introduciamo uno strumento fondamentale per lo studio dell'Algebra Lineare: il determinante di una matrice quadrata, una funzione che dice esattamente quando le colonne (o le righe) di una matrice quadrata sono linearmente indipendenti. Invece di definirlo direttamente con una formula, elencheremo le proprietà che vorremmo possedesse, e faremo vedere che esiste al più una funzione che soddisfa le nostre richieste. L'esistenza sarà poi dimostrata tramite gli sviluppi di Laplace, uno dei due metodi principali per il calcolo del determinante (l'altro è un'applicazione dell'eliminazione di Gauss). Faremo anche vedere che il determinante del prodotto di matrici è uguale al prodotto dei determinanti (Teorema di Binet). Vedremo infine, nel Teorema degli orlati, come calcolare il rango di matrici qualunque usando il determinante di sottomatrici quadrate. Come al solito, tutto quanto diremo varrà per matrici a coefficienti in un campo $\mathbb{K}$ qualsiasi.

9.1 Esistenza e unicità del determinante

Nell'Esercizio 7.19 abbiamo visto che una matrice

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \in M_{2,2}(\mathbb{R})$$

è invertibile se e solo se $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Pertanto le righe A_1 e A_2 (o le colonne A^1 e A^2) di A sono linearmente indipendenti se e solo se $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Abbiamo dunque trovato una funzione $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ tale che i vettori $v, w \in \mathbb{R}^2$ sono linearmente indipendenti se e solo se $d(v, w) \neq 0$: è la funzione

$$d\left(\begin{vmatrix} v_1 \\ v_2 \end{vmatrix}, \begin{vmatrix} w_1 \\ w_2 \end{vmatrix} \right) = v_1 w_2 - v_2 w_1 .$$

Osservazione 9.1

Esiste un'altra interpretazione della funzione $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$. Prendiamo due vettori $v, w \in \mathbb{R}^2$, indichiamo le loro coordinate rispettivamente con (v_1, v_2) e (w_1, w_2), e consideriamo il parallelogramma P di vertici O, v, w e $v + w$. Per semplicità, supponiamo che tutte le coordinate di v e w siano positive (vedi la Figura 9.1).

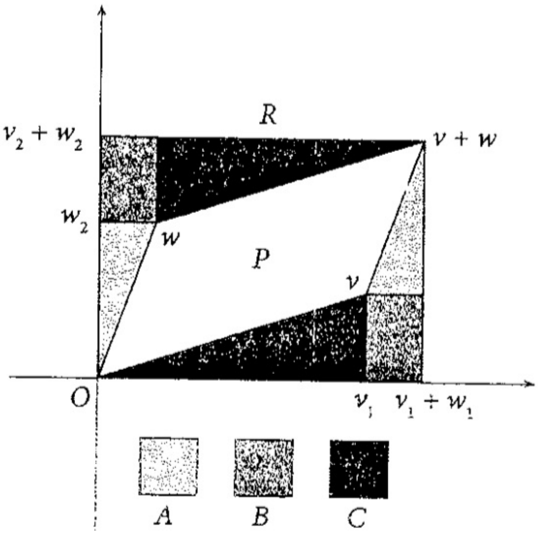

Figura 9.1 - Il determinante come area orientata.

Se indichiamo con R il rettangolo di vertici opposti O e $v + w$, vediamo dalla figura che R può venire suddiviso in sette parti: un parallelogramma (P), due rettangoli congruenti (B) e due coppie di triangoli congruenti (A e C). Quindi

$$\operatorname{Area}(P) = \operatorname{Area}(R) - 2\ \operatorname{Area}(A) - 2\ \operatorname{Area}(B) - 2\ \operatorname{Area}(C).$$

Siccome si ha $\operatorname{Area}(R) = (v_1 + w_1)(v_2 + w_2)$, $\operatorname{Area}(B) = v_2 w_1$, $\operatorname{Area}(A) = w_1 w_2/2$ e $\operatorname{Area}(C) = v_1 v_2/2$, allora $\operatorname{Area}(P) = v_1 w_2 - v_2 w_1 = d(v, w)$. La funzione d coincide così con l'area del parallelogramma generato da v e w. O meglio, con l'*area orientata* del parallelogramma: $d(v, w) = \operatorname{Area}(P)$ se per andare da v a w si ruota in senso antiorario, come nella Figura 9.1; $d(v, w) = -\operatorname{Area}(P)$ altrimenti (verifica che questo è effettivamente quanto accade). In particolare, ritroviamo nuovamente che d si annulla se e solo se v e w sono linearmente dipendenti: infatti in tal caso il "parallelogramma" da essi generato ha area nulla – è un segmento.

Scopo di questo paragrafo è trovare un analogo in dimensione n, cioè una funzione che dica esattamente quando n vettori di $\mathbb{R}^n$ sono linearmente indipendenti; vedremo che si potrà anche usare per stabilire se p vettori di $\mathbb{R}^n$ sono linearmente indipendenti. Per trovare questa funzione, il *determinante*, cominciamo col fare un elenco delle proprietà che desideriamo abbia; l'idea è che partendo da esse saremo poi in grado di costruire la funzione cercata. Ovviamente, le proprietà che indicheremo sono state selezionate con la certezza – dovuta a più di un secolo di esperienza – che ci porteranno alla giusta conclusione; altre scelte apparentemente altrettanto valide potrebbero condurre a risultati infausti (vedi l'Esempio 9.4).

Un'osservazione preliminare. Invece di considerare n vettori di $\mathbb{R}^n$ separatamente, possiamo metterli in una matrice A e cercare una funzione $d\colon M_{n,n}(\mathbb{R}) \to \mathbb{R}$ tale che $d(A) = 0$ se e solo se le righe[(1)] di A sono linearmente dipendenti. Quindi considereremo d sia come funzione della matrice A sia come funzione delle sue n righe $A_1, \ldots, A_n$.

Il caso più eclatante in cui le righe di A sono linearmente dipendenti è quando due righe sono uguali. Quindi la prima proprietà che richiediamo alla funzione d è

(A) $$d(A_1, \ldots, A_n) = 0 \qquad \text{se } A_i = A_j \text{ per qualche } i \neq j.$$

In altri termini, *d si annulla se due righe sono uguali.*

Ora, se le n righe $A_1, \ldots, A_n$ sono linearmente dipendenti, chiaramente lo sono anche $\lambda A_1, A_2, \ldots, A_n$ qualunque sia $\lambda \in \mathbb{R}$. Viceversa, se $\lambda \neq 0$ e $A_1, \ldots, A_n$ sono linearmente indipendenti, anche $\lambda A_1, A_2, \ldots, A_n$ lo sono; però, se λ diventa sempre più piccolo in valore assoluto allora i vettori $\lambda A_1, A_2, \ldots, A_n$ sono sempre più vicini a essere dipendenti – diventandolo esattamente nel momento in cui λ si annulla. Un modo abbastanza naturale di rappresentare questo comportamento con la nostra funzione d (e del tutto coerente con l'interpretazione di d come area del parallelogramma vista per $n = 2$) è richiedere che $d(\lambda A_1, A_2, \ldots, A_n)$ sia uguale a λ volte $d(A_1, \ldots, A_n)$. Siccome è evidente che lo stesso ragionamento si poteva applicare a qualsiasi riga, e non solo alla prima, la nostra seconda richiesta allora è

(B) $$d(\ldots, \lambda A_i, \ldots) = \lambda\, d(\ldots, A_i, \ldots) \qquad \text{per qualsiasi } \lambda \in \mathbb{R} \text{ e } i = 1, \ldots, n.$$

In altri termini, *d è omogeneo in ogni riga.*

Abbiamo richiesto che d si comporti bene rispetto al prodotto per scalari. E rispetto alla somma? Nei capitoli precedenti abbiamo studiato a lungo le applicazioni lineari; quindi ci farebbe piacere che anche il determinante fosse in qualche modo lineare. Proviamo quindi ad aggiungere alla nostra lista la proprietà

(C) $$d(\ldots, A_i' + A_i'', \ldots) = d(\ldots, A_i', \ldots) + d(\ldots, A_i'', \ldots) \quad \text{per ogni } i = 1, \ldots, n.$$

In altri termini, richiediamo che d sia *additivo in ogni riga*. Le proprietà (B) e (C) si possono riassumere dicendo che *d è lineare in ogni riga.*

Vi è una funzione banale che soddisfa tutte queste proprietà: la funzione identicamente nulla. Per escluderla possiamo richiedere che d sia non zero su una matrice fissata. Per semplicità (e in analogia con il caso $n = 2$) richiediamo

(D) $$d(I_n) = 1,$$

cioè chiediamo che *d calcolato sulla matrice identica sia uguale a* 1.

[(1)] Utilizzeremo le righe invece delle colonne di A per mantenere una certa compatibilità con l'eliminazione di Gauss.

Per capire se stiamo procedendo nella direzione giusta, vediamo che conseguenze si possono trarre dalle proprietà (A)-(C).

Proposizione 9.1

Supponiamo che la funzione $d: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ goda delle proprietà (A)-(C). Allora:

(1) *se A ha una riga nulla allora $d(A) = 0$;*

(2) *il valore di d non cambia sommando a una riga un multiplo di un'altra, ovvero $d(\ldots, A_i + \lambda A_j, \ldots, A_j, \ldots) = d(\ldots, A_i, \ldots, A_j, \ldots)$ per ogni $i \neq j$ e $\lambda \in \mathbb{R}$;*

(3) *il valore di d cambia segno se si scambiano fra loro due righe qualunque, ovvero $d(\ldots, A_i, \ldots, A_j, \ldots) = -d(\ldots, A_j, \ldots, A_i, \ldots)$, per ogni $i \neq j$;*

(4) *se S è una matrice triangolare superiore ottenuta da A con un'eliminazione di Gauss (o una riduzione a scala) con σ scambi di righe, allora $d(A) = (-1)^\sigma d(S)$;*

(5) *se le righe di A sono linearmente dipendenti allora $d(A) = 0$.*

Dimostrazione

(1) Basta applicare (B) con $\lambda = 0$, dove A_i è la riga nulla.

(2) Utilizzando nell'ordine le proprietà (C), (B) e (A) otteniamo

$$\begin{aligned}
d(\ldots, A_i + \lambda A_j, \ldots, A_j, \ldots) &= d(\ldots, A_i, \ldots, A_j, \ldots) + d(\ldots, \lambda A_j, \ldots, A_j, \ldots) \\
&= d(\ldots, A_i, \ldots, A_j, \ldots) + \lambda\, d(\ldots, A_j, \ldots, A_j, \ldots) \\
&= d(\ldots, A_i, \ldots, A_j, \ldots).
\end{aligned}$$

(3) Questa volta utilizziamo nell'ordine le proprietà (A), (C) e di nuovo (A):

$$\begin{aligned}
0 &= d(\ldots, A_i + A_j, \ldots, A_i + A_j, \ldots) \\
&= d(\ldots, A_i, \ldots, A_i + A_j, \ldots) + d(\ldots, A_j, \ldots, A_i + A_j, \ldots) \\
&= d(\ldots, A_i, \ldots, A_i, \ldots) + d(\ldots, A_i, \ldots, A_j, \ldots) \\
&\quad + d(\ldots, A_j, \ldots, A_i, \ldots) + d(\ldots, A_j, \ldots, A_j, \ldots) \\
&= d(\ldots, A_i, \ldots, A_j, \ldots) + d(\ldots, A_j, \ldots, A_i, \ldots).
\end{aligned}$$

(4) La matrice S si ottiene dalla matrice A tramite due tipi di operazioni: lo scambio di due righe e la sostituzione di una riga A_i con una combinazione lineare del tipo $A_i + \lambda A_j$, con $j \neq i$. Per il punto (3), a ogni scambio di righe il valore di d viene moltiplicato per -1 (cioè cambiato di segno). La seconda operazione, invece, lascia d immutato, grazie al punto (2). Quindi $d(S) = \pm d(A)$, dove il segno dipende dal numero di scambi di righe esattamente come asserito.

(5) Se le righe di A sono linearmente dipendenti, qualunque riduzione a scala S di A ha almeno una riga nulla. Per (1), dunque, $d(S) = 0$, e il punto (4) ci assicura che in tal caso $d(A) = 0$. □

Osservazione 9.2

La Proposizione 9.1 ci dice che le operazioni effettuate in un'eliminazione di Gauss (o in una riduzione a scala) possono al più cambiare di segno d. Ciò non è più vero se alla riga A_i sostituiamo $\lambda A_i + \mu A_j$ con $i \neq j$; in tal caso il valore di d viene infatti moltiplicato per λ, grazie alla Proposizione 9.1.(2) e alla proprietà (B). Questo è il pericolo insito nell'"eliminazione di Gauss generalizzata" di cui avevamo parlato nell'Osservazione 3.6.

Dunque la nostra funzione d, se esiste, si annulla sulle righe linearmente dipendenti. A priori, però, potrebbe annullarsi *anche* su righe linearmente indipendenti. Sarà la condizione (D) ad assicurarci che questo non accade; ma prima, vediamo come calcolare d con le informazioni che disponiamo, almeno in alcuni casi semplici.

Esempio 9.1

Sia $d\colon M_{2,2}(\mathbb{R}) \to \mathbb{R}$ una funzione soddisfacente le proprietà (A)-(C), e prendiamo una matrice

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \in M_{2,2}(\mathbb{R}).$$

Le righe di A si possono scrivere usando la base canonica $\{e_1, e_2\}$ di $\mathbb{R}^2$ come

$$A_1 = a_{11}e_1^T + a_{12}e_2^T \qquad \text{e} \qquad A_2 = a_{21}e_1^T + a_{22}e_2^T$$

(convincitene prima di proseguire), dove $e_1^T = \begin{vmatrix} 1 & 0 \end{vmatrix}$ ed $e_2^T = \begin{vmatrix} 0 & 1 \end{vmatrix}$. Utilizzando, nell'ordine, le proprietà (B) e (C) sulla prima e poi sulla seconda riga, la proprietà (A) e la Proposizione 9.1.(3), otteniamo

$$\begin{aligned} d(A) &= d(a_{11}e_1^T + a_{12}e_2^T, a_{21}e_1^T + a_{22}e_2^T) \\ &= a_{11}\, d(e_1^T, a_{21}e_1^T + a_{22}e_2^T) + a_{12}\, d(e_2^T, a_{21}e_1^T + a_{22}e_2^T) \\ &= a_{11}a_{21}\, d(e_1^T, e_1^T) + a_{11}a_{22}\, d(e_1^T, e_2^T) + a_{12}a_{21}\, d(e_2^T, e_1^T) + a_{12}a_{22}\, d(e_2^T, e_2^T) \\ &= a_{11}a_{22}d(e_1^T, e_2^T) + a_{12}a_{21}d(e_2^T, e_1^T) = (a_{11}a_{22} - a_{12}a_{21})d(e_1^T, e_2^T) \\ &= (a_{11}a_{22} - a_{12}a_{21})\, d(I_2). \end{aligned}$$

In particolare, se vale anche (D) otteniamo

$$d(A) = a_{11}a_{22} - a_{12}a_{21}. \tag{9.1}$$

Abbiamo quindi ricavato una formula per calcolare d su ogni matrice 2×2, utilizzando soltanto le proprietà (A)-(D) e le loro conseguenze. In altre parole, per $n = 2$ se una tale funzione d esiste è necessariamente unica, in quanto dev'essere data dalla formula (9.1). Viceversa, la formula (9.1) definisce davvero una funzione d che soddisfa le proprietà (A)-(D), come avrai cura di verificare (del resto, è proprio la funzione con cui avevamo iniziato il paragrafo). Dunque per $n = 2$ esiste una e una sola funzione che soddisfa le proprietà (A)-(D), il *determinante* delle matrici 2×2.

Osservazione 9.3

Il procedimento utilizzato per ricavare la formula (9.1) può essere usato anche per $n > 2$, e condurre così a una dimostrazione dell'esistenza e unicità del determinante tramite quello che viene detto il gruppo delle *permutazioni*. La formula che si ottiene è complicata da scrivere e praticamente inutile per il calcolo del determinante; quindi preferiamo procedere in altro modo. Per avere un'idea dei fenomeni che si devono affrontare con questo approccio, risolvi l'Esercizio 9.1 e gli Esercizi 9.31-9.32.

Continuiamo con esempi del calcolo della funzione d su matrici semplici.

Esempio 9.2

Sia

$$A = \begin{vmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{vmatrix} \in M_{n,n}(\mathbb{R})$$

una matrice diagonale (dove gli spazi bianchi sono da riempire di zeri). Allora utilizzando ripetutamente la proprietà (B) otteniamo

$$\begin{aligned} d(A) &= d(a_{11}e_1^T, a_{22}e_2^T, \dots, a_{nn}e_n^T) = a_{11}\, d(e_1^T, a_{22}e_2^T, \dots, a_{nn}e_n^T) \\ &= a_{11}a_{22}\, d(e_1^T, e_2^T, \dots, a_{nn}e_n^T) = \cdots = a_{11}a_{22}\cdots a_{nn}\, d(e_1^T, e_2^T, \dots, e_n^T) \\ &= a_{11}a_{22}\cdots a_{nn}\, d(I_n). \end{aligned}$$

In particolare, se vale anche (D) si ha

$$d(A) = a_{11}a_{22}\cdots a_{nn}.$$

Esempio 9.3

Sia

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix} \in M_{n,n}(\mathbb{R})$$

una matrice triangolare superiore; vogliamo trovare $d(A)$. Scriviamo $A_1 = A_1' + A_1''$, dove $A_1' = | a_{11} \quad 0 \quad \cdots \quad 0 |$ e $A_1'' = | 0 \quad a_{12} \quad \cdots \quad a_{1n} |$. Allora

$$\begin{aligned} d(A) &= d(A_1, A_2, \dots, A_n) = d(A_1', A_2, \dots, A_n) + d(A_1'', A_2, \dots, A_n) \\ &= d(A_1', A_2, \dots, A_n) \end{aligned}$$

grazie alla Proposizione 9.1.(5), in quanto le righe $A_1'', A_2, \dots, A_n$ sono linearmente dipendenti (perché?). Procediamo in modo analogo con la seconda riga. Scriviamo pertanto $A_2 = A_2' + A_2''$, dove $A_2' = | 0 \quad a_{22} \quad 0 \quad \cdots \quad 0 |$ e $A_2'' = | 0 \quad 0 \quad a_{23} \quad \cdots \quad a_{2n} |$. Allora

$$\begin{aligned} d(A) &= d(A_1', A_2, \dots, A_n) = d(A_1', A_2', \dots, A_n) + d(A_1', A_2'', \dots, A_n) \\ &= d(A_1', A_2', \dots, A_n) \end{aligned}$$

in quanto di nuovo le righe $A_1', A_2'', \dots, A_n$ sono linearmente dipendenti (perché?). Procedendo in tal modo si arriva a $d(A) = d(A_1', A_2', \dots, A_n') = a_{11}a_{22}\cdots a_{nn}\, d(I_n)$, grazie all'esempio precedente. In particolare, se vale anche la proprietà (D) otteniamo

$$d(A) = d(A_1', A_2', \dots, A_n') = a_{11}a_{22}\cdots a_{nn};$$

in altre parole, *il determinante di una matrice triangolare superiore è uguale al prodotto degli elementi sulla diagonale principale.*

Questo è quanto ci basta per calcolare d applicato a una matrice qualunque.

Proposizione 9.2

Supponiamo che la funzione $d: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ goda delle proprietà (A)-(C). Presa una matrice $A \in M_{n,n}(\mathbb{R})$, siano $p_1, \ldots, p_n \in \mathbb{R}$ i pivot ottenuti tramite un'eliminazione di Gauss effettuata con σ scambi di righe. Allora

$$d(A) = (-1)^\sigma p_1 \cdots p_n \, d(I_n). \tag{9.2}$$

Dimostrazione

Abbiamo già visto che $d(A) = (-1)^\sigma d(S)$, dove S è la matrice triangolare superiore prodotta dall'eliminazione di Gauss considerata; quindi per concludere è sufficiente utilizzare l'Esempio 9.3. □

Corollario 9.3

Sia $d: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ una funzione che gode delle proprietà (A)-(C) e inoltre tale che $d(I_n) \neq 0$. Allora $d(A) = 0$ se e solo se le righe di A sono linearmente dipendenti.

Dimostrazione

Se le righe di A sono linearmente dipendenti, sappiamo già che si deve avere $d(A) = 0$. Viceversa, se sono linearmente indipendenti il rango di A è n, per cui i pivot sono tutti non nulli e dunque (9.2) ci dice che $d(A) \neq 0$. □

Corollario 9.4

Per ogni $n \in \mathbb{N}^*$ esiste al più un'unica funzione $d: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ che gode delle proprietà (A)-(D).

Dimostrazione

Infatti, se d esiste si deve calcolare tramite (9.2). □

Osservazione 9.4

La formula (9.2) *non* basta per dimostrare che la nostra funzione d esiste. Il problema è che, come abbiamo detto più volte, i pivot di una matrice *non* sono univocamente determinati; quindi chi ci dice che il loro prodotto lo sia? Quello che potrebbe succedere è che effettuando un'eliminazione di Gauss la formula (9.2) dia un valore per $d(A)$, mentre effettuando una diversa eliminazione di Gauss la formula dia un valore *differente* per $d(A)$. Se ciò accadesse, saremmo costretti a concludere che una tale funzione d non può esistere, in quanto una funzione non può assumere due valori diversi se calcolata sullo stesso elemento. In altri termini, se ciò accadesse saremmo costretti a concludere che le proprietà (A)-(D) sono *contraddittorie:* non possono verificarsi tutte contemporaneamente.

Esempio 9.4

Se alle proprietà (A)-(D) avessimo aggiunto la

(E) $\quad d(\ldots, A_i, \ldots, A_j, \ldots) = d(\ldots, A_j, \ldots, A_i, \ldots) \qquad \text{se } i \neq j,$

avremmo scoperto che nessuna funzione può soddisfare contemporaneamente tutte le proprietà (A)-(E). Infatti, la Proposizione 9.1.(3), che è conseguenza di (A)-(C) soltanto, assieme alla (E) darebbe $d(A) \equiv 0$, contraddicendo platealmente la (D). □

Dunque (9.2) non basta per dimostrare l'esistenza della funzione d. Per superare questo problema, nel prossimo teorema daremo una formula per calcolare d che non dipende da nessuna scelta arbitraria. Ma prima ci serve una definizione.

Definizione 9.1

Sia $A \in M_{n,n}(\mathbb{R})$. La matrice che si ottiene cancellando la riga i e la colonna j di A si chiama *minore* (i,j) di A, e si indica con $A_{ij} \in M_{n-1,n-1}(\mathbb{R})$.

Esempio 9.5

Sia

$$A = \begin{vmatrix} 1 & 2 & -3 \\ 4 & 6 & 1 \\ 3 & 3 & -2 \end{vmatrix}.$$

Allora

$$A_{11} = \begin{vmatrix} 6 & 1 \\ 3 & -2 \end{vmatrix}, \qquad A_{23} = \begin{vmatrix} 1 & 2 \\ 3 & 3 \end{vmatrix}$$

e così via.

Teorema 9.5

Per ogni $n \in \mathbb{N}^*$ esiste un'unica funzione $d_n: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ che soddisfa le proprietà (A)-(D). Se $n = 1$ essa è data da $d_1(A) = a_{11}$; se $n > 1$ da

$$\begin{aligned} d_n(A) &= a_{11}d_{n-1}(A_{11}) - a_{21}d_{n-1}(A_{21}) + \cdots + (-1)^{n+1}a_{n1}d_{n-1}(A_{n1}) \\ &= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}\, d_{n-1}(A_{i1}). \end{aligned} \tag{9.3}$$

Osservazione 9.5

Il simbolo di *sommatoria* $\sum$ serve ad abbreviare formule in cui vengono sommati tutti i valori di un certo oggetto a_j al variare dell'indice j. Per esempio

$$\sum_{j=1}^{n} a_j = \sum_{1 \leq j \leq n} a_j = a_1 + \cdots + a_n.$$

Osservazione 9.6

La (9.3) è un esempio di *formula ricorsiva* (o *induttiva*); la funzione d_n, che dipende da n, viene calcolata in termini di d_{n-1}, la quale a sua volta viene espressa in termini di d_{n-2} e così via. In questo modo si riconduce il calcolo di d_n al calcolo di d_2 che sappiamo come eseguire, grazie all'Esempio 9.1; e quindi, risalendo pazientemente, otteniamo anche un valore unico per d_n. In altri termini, tramite (9.3) si definisce davvero una funzione $d_n: M_{n,n}(\mathbb{R}) \to \mathbb{R}$; tutto quello che dovremo fare è verificare che soddisfi le proprietà (A)-(D) qualunque sia $n \in \mathbb{N}^*$.

Dimostrazione del Teorema 9.5

Procediamo per induzione su n. Per $n = 2$ sappiamo già che esiste un'unica funzione $d_2: M_{2,2}(\mathbb{R}) \to \mathbb{R}$ che soddisfa le proprietà (A)-(D), la funzione

$$d_2(A) = a_{11}a_{22} - a_{12}a_{21}. \tag{9.4}$$

Siccome $A_{11} = |a_{22}|$, $A_{12} = |a_{21}|$, $A_{21} = |a_{12}|$, $A_{22} = |a_{11}|$, ponendo $d_1(|a|) = a$ per ogni matrice $|a| \in M_{1,1}(\mathbb{R})$ vediamo che (9.4) e (9.3) coincidono.

Supponiamo allora che esista un'unica funzione $d_{n-1}: M_{n-1,n-1}(\mathbb{R}) \to \mathbb{R}$ che soddisfi (A)-(D), e definiamo $d_n: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ tramite la (9.3). Dobbiamo dimostrare che anche questa d_n soddisfa le proprietà (A)-(D).

Cominciamo dalla più facile, la (D). Come si verifica subito, il minore (1,1) della matrice identica I_n è la matrice identica I_{n-1}. Inoltre gli elementi della prima colonna di I_n successivi al primo sono tutti zero; quindi in $d_n(I_n)$ solo il primo addendo della formula (9.3) è non nullo, per cui

$$d_n(I_n) = 1 \cdot d_{n-1}(I_{n-1}) - 0 + \cdots + (-1)^{n+1}\, 0 = d_{n-1}(I_{n-1}) = 1,$$

in quanto, per ipotesi induttiva, d_{n-1} soddisfa la proprietà (D).

Passiamo a (C). Supponiamo che la riga in cui avviene la somma sia la riga j_0. Consideriamo tre matrici:

$$B = \begin{vmatrix} A_1 \\ \vdots \\ A'_{j_0} + A''_{j_0} \\ \vdots \\ A_n \end{vmatrix}, \qquad A' = \begin{vmatrix} A_1 \\ \vdots \\ A'_{j_0} \\ \vdots \\ A_n \end{vmatrix}, \qquad A'' = \begin{vmatrix} A_1 \\ \vdots \\ A''_{j_0} \\ \vdots \\ A_n \end{vmatrix};$$

vogliamo dimostrare che $d_n(B) = d_n(A') + d_n(A'')$. Per far ciò dobbiamo descrivere gli elementi della prima colonna e i minori $(i, 1)$ di B in termini dei corrispondenti oggetti di A' e A''. Prima di tutto, $b_{i1} = a_{i1}$ se $i \neq j_0$ e $b_{i1} = a'_{j_0 1} + a''_{j_0 1}$ se $i = j_0$. Per i minori, invece, avviene un fenomeno complementare. Infatti si vede subito che il minore $B_{j_0 1}$ è uguale al minore $A'_{j_0 1}$ (che a sua volta è uguale al minore $A''_{j_0 1}$), mentre i minori B_{i1} con $i \neq j_0$ differiscono dai minori A'_{i1} e A''_{i1} solo nella riga proveniente dalla riga j_0, che è rimasta una somma. Ricordando che, per ipotesi induttiva, d_{n-1} gode della proprietà (C), otteniamo

$$b_{i1} d_{n-1}(B_{i1}) = \begin{cases} a_{i1}[d_{n-1}(A'_{i1}) + d_{n-1}(A''_{i1})] & \text{se } i \neq j_0; \\ [a'_{j_0 1} + a''_{j_0 1}]\, d_{n-1}(A'_{j_0 1}) = [a'_{j_0 1} + a''_{j_0 1}]\, d_{n-1}(A''_{j_0 1}) & \text{se } i = j_0. \end{cases}$$

Dunque abbiamo

$$\begin{aligned} d_n(B) &= b_{11}\, d_{n-1}(B_{11}) + \cdots + (-1)^{j_0+1} b_{j_0 1}\, d_{n-1}(B_{j_0 1}) \\ &\qquad + \cdots + (-1)^{n+1} b_{n1}\, d_{n-1}(B_{n1}) \\ &= a_{11}[d_{n-1}(A'_{11}) + d_{n-1}(A''_{11})] + \cdots + (-1)^{j_0+1}[a'_{j_0 1} + a''_{j_0 1}]\, d_{n-1}(A'_{j_0 1}) \\ &\qquad + \cdots + (-1)^{n+1}\, a_{n1}[d_{n-1}(A'_{n1}) + d_{n-1}(A''_{n1})] \\ &= [a_{11}\, d_{n-1}(A'_{11}) + \cdots + (-1)^{j_0+1} a'_{j_0 1}\, d_{n-1}(A'_{j_0 1}) \\ &\qquad + \cdots + (-1)^{n+1} a_{n1}\, d_{n-1}(A'_{n1})] \\ &\quad + [a_{11}\, d_{n-1}(A''_{11}) + \cdots + (-1)^{j_0+1} a''_{j_0 1}\, d_{n-1}(A''_{j_0 1}) \\ &\qquad + \cdots + (-1)^{n+1} a_{n1}\, d_{n-1}(A''_{n1})] \\ &= d_n(A') + d_n(A''), \end{aligned}$$

e ci siamo. La dimostrazione di (B) è del tutto analoga, e te la lasciamo come esercizio.

Rimane da dimostrare la proprietà (A) – e ora vedremo a cosa serve l'alternarsi dei segni nella (9.3). Sia A una matrice con due righe uguali, diciamo la i_0 e la j_0; ovvero, supponiamo $A_{i_0} = A_{j_0}$. Dobbiamo dimostrare che $d_n(A) = 0$. Se $i \neq i_0, j_0$ anche il minore A_{i1} ha due righe uguali; per ipotesi induttiva, $d_{n-1}(A_{i1}) = 0$. Allora la (9.3) si riduce a

$$d_n(A) = (-1)^{i_0+1} a_{i_0 1}\, d_{n-1}(A_{i_0 1}) + (-1)^{j_0+1} a_{j_0 1}\, d_{n-1}(A_{j_0 1}).$$

Prima di tutto, essendo le righe i_0 e j_0 uguali, abbiamo $a_{i_0 1} = a_{j_0 1}$. Inoltre, $A_{i_0 1}$ e $A_{j_0 1}$ hanno le stesse righe, disposte in ordine diverso. Per l'esattezza, si ha una situazione di questo genere

$$A_{i_0 1} = \begin{vmatrix} \vdots \\ A'_{i_0-1} \\ A'_{i_0+1} \\ \vdots \\ A'_{j_0-1} \\ A'_{j_0} = A'_{i_0} \\ A'_{j_0+1} \\ \vdots \end{vmatrix}, \qquad A_{j_0 1} = \begin{vmatrix} \vdots \\ A'_{i_0-1} \\ A'_{i_0} = A'_{j_0} \\ A'_{i_0+1} \\ \vdots \\ A'_{j_0-1} \\ A'_{j_0+1} \\ \vdots \end{vmatrix},$$

dove A'_i è la riga A_i di A privata del primo elemento. Dunque, per trasformare $A_{j_0 1}$ in $A_{i_0 1}$ basta scambiare A'_{i_0} con, nell'ordine, A'_{i_0+1}, A'_{i_0+2} e così via fino a A'_{j_0-1}. In questo modo effettuiamo $j_0 - i_0 - 1$ scambi di righe (in quanto siamo partiti da $i_0 + 1$ e siamo arrivati a $i_0 + (j_0 - i_0) - 1$). Dunque la Proposizione 9.1.(3), che per ipotesi induttiva vale per d_{n-1}, ci dà $d_{n-1}(A_{i_0 1}) = (-1)^{j_0-i_0-1} d_{n-1}(A_{j_0 1})$, per cui

$$\begin{aligned} d_n(A) &= (-1)^{i_0+1} a_{i_0 1}\, d_{n-1}(A_{i_0 1}) + (-1)^{j_0+1} a_{j_0 1}\, d_{n-1}(A_{j_0 1}) \\ &= (-1)^{i_0+1}(-1)^{j_0-i_0-1} a_{j_0 1}\, d_{n-1}(A_{j_0 1}) + (-1)^{j_0+1} a_{j_0 1}\, d_{n-1}(A_{j_0 1}) \\ &= (-1)^{j_0}\big(a_{j_0 1}\, d_{n-1}(A_{j_0 1}) - a_{j_0 1}\, d_{n-1}(A_{j_0 1})\big) = 0, \end{aligned}$$

e abbiamo concluso.

Definizione 9.2

La funzione definita dalla formula (9.3) si chiama *determinante*, ed è di solito indicata con $\det\colon M_{n,n}(\mathbb{R}) \to \mathbb{R}$.

Esempio 9.6

Sia

$$A = \begin{vmatrix} 3 & 4 & 2 & 1 \\ 0 & 7 & 6 & 3 \\ 1 & 1 & 2 & 4 \\ 3 & 1 & -2 & 4 \end{vmatrix} \in M_{4,4}(\mathbb{R});$$

vogliamo calcolarne il determinante $\det A$. Effettuando un'eliminazione di Gauss su A otteniamo

$$\begin{vmatrix} 3 & 4 & 2 & 1 \\ 0 & 7 & 6 & 3 \\ 1 & 1 & 2 & 4 \\ 3 & 1 & -2 & 4 \end{vmatrix} \longrightarrow \begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 7 & 6 & 3 \\ 3 & 4 & 2 & 1 \\ 3 & 1 & -2 & 4 \end{vmatrix} \longrightarrow \begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 7 & 6 & 3 \\ 0 & 1 & -4 & -11 \\ 0 & -2 & -8 & -8 \end{vmatrix}$$

$$\longrightarrow \begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & -4 & -11 \\ 0 & 7 & 6 & 3 \\ 0 & -2 & -8 & -8 \end{vmatrix} \longrightarrow \begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & -4 & -11 \\ 0 & 0 & 34 & 80 \\ 0 & 0 & -16 & -30 \end{vmatrix} \longrightarrow \begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & -4 & -11 \\ 0 & 0 & 34 & 80 \\ 0 & 0 & 0 & 130/17 \end{vmatrix},$$

e applicando la Proposizione 9.3 troviamo $\det(A) = (-1)^2 1 \cdot 1 \cdot 34 \cdot (130/17) = 260$, in quanto abbiamo effettuato due scambi di righe. Altrimenti, possiamo utilizzare la formula (9.3) per ottenere nuovamente

$$\begin{aligned}
\det(A) &= 3\det\begin{vmatrix} 7 & 6 & 3 \\ 1 & 2 & 4 \\ 1 & -2 & 4 \end{vmatrix} - 0\det\begin{vmatrix} 4 & 2 & 1 \\ 1 & 2 & 4 \\ 1 & -2 & 4 \end{vmatrix} + \det\begin{vmatrix} 4 & 2 & 1 \\ 7 & 6 & 3 \\ 1 & -2 & 4 \end{vmatrix} - 3\det\begin{vmatrix} 4 & 2 & 1 \\ 7 & 6 & 3 \\ 1 & 2 & 4 \end{vmatrix} \\
&= 3\left[7\det\begin{vmatrix} 2 & 4 \\ -2 & 4 \end{vmatrix} - \det\begin{vmatrix} 6 & 3 \\ -2 & 4 \end{vmatrix} + \det\begin{vmatrix} 6 & 3 \\ 2 & 4 \end{vmatrix}\right] \\
&\quad + \left[4\det\begin{vmatrix} 6 & 3 \\ -2 & 4 \end{vmatrix} - 7\det\begin{vmatrix} 2 & 1 \\ -2 & 4 \end{vmatrix} + \det\begin{vmatrix} 2 & 1 \\ 6 & 3 \end{vmatrix}\right] \\
&\quad - 3\left[4\det\begin{vmatrix} 6 & 3 \\ 2 & 4 \end{vmatrix} - 7\det\begin{vmatrix} 2 & 1 \\ 2 & 4 \end{vmatrix} + \det\begin{vmatrix} 2 & 1 \\ 6 & 3 \end{vmatrix}\right] \\
&= 3[7 \cdot 16 - 30 + 18] + [4 \cdot 30 - 7 \cdot 10 + 0] - 3[4 \cdot 18 - 7 \cdot 6 + 0] \\
&= 300 + 50 - 90 = 260.
\end{aligned}$$

$\sqcap\!\!\!\!\sqcup$

9.2 Sviluppi di Laplace

Abbiamo dunque dimostrato l'esistenza e l'unicità della funzione determinante, e abbiamo dato due modi per calcolarla: tramite l'eliminazione di Gauss, o tramite la formula (9.3). Vediamo ora qualche conseguenza di quanto ottenuto finora.

Prima di tutto, è chiaro (esercizio) che la dimostrazione del Teorema 9.5 può essere adattata in modo da dimostrare anche che

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

qualunque sia $1 \leq j \leq n$. Questa formula si chiama *sviluppo di Laplace del determinante lungo la colonna j-esima*. Vi è uno sviluppo analogo per righe.

Proposizione 9.6

Sia $A \in M_{n,n}(\mathbb{R})$ e $1 \leq i \leq n$. Allora

$$\det(A) = (-1)^{i+1}a_{i1}\det(A_{i1}) + \cdots + (-1)^{i+n}a_{in}\det(A_{in}) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}),$$

(sviluppo di Laplace del determinante lungo la riga i-esima).

Dimostrazione

Sia $A_i = a_{i1}e_1^T + \cdots + a_{in}e_n^T$ la riga i-esima. Allora

$$\begin{aligned}\det(A) &= \det(A_1,\ldots,A_i,\ldots,A_n)\\ &= a_{i1}\det(A_1,\ldots,e_1^T,\ldots,A_n) + \cdots + a_{in}\det(A_1,\ldots,e_n^T,\ldots,A_n)\\ &= a_{i1}\det(B_1) + \cdots + a_{in}\det(B_n),\end{aligned}\tag{9.5}$$

dove

$$B_j = \begin{vmatrix} A_1 \\ \vdots \\ A_{i-1} \\ e_j^T \\ A_{i+1} \\ \vdots \\ A_n \end{vmatrix} \in M_{n,n}(\mathbb{R}).$$

Ora, se nella matrice B_j sottraiamo opportuni multipli della i-esima riga (che è e_j^T) a tutte le altre righe, otteniamo una matrice C_j di uguale determinante e la cui j-esima colonna è e_i. Per l'esattezza

$$C_j = \left|\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{array}\right|.$$

Calcolando ora $\det(C_j)$ con lo sviluppo di Laplace lungo la colonna j-esima, solo un addendo è diverso da zero, per cui

$$\det(B_j) = \det(C_j) = (-1)^{i+j}\det(A_{ij}). \tag{9.6}$$

Mettendo insieme (9.5) e (9.6) abbiamo la tesi. □

Osservazione 9.7

Un buon modo per ricordarsi i segni negli sviluppi di Laplace è di visualizzarli in una scacchiera infinita:

$$\begin{vmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{vmatrix}$$

Quindi partendo dall'angolo in alto a sinistra (segno $+$) si procede in orizzontale o verticale di un elemento alla volta, cambiando segno a ogni passo, fino a giungere all'elemento (e quindi al segno) corrispondente al minore desiderato.

Per calcolare il determinante possiamo dunque sviluppare sia per righe sia per colonne. Del resto, avevamo già visto che il numero di righe linearmente indipendenti di una matrice è uguale al numero di colonne linearmente indipendenti; quindi è ragionevole che una funzione creata per dire se le righe sono linearmente indipendenti dia risultati analoghi per le colonne. Anzi, i risultati sono proprio uguali.

Corollario 9.7

Sia $A \in M_{n,n}(\mathbb{R})$. *Allora* $\det(A^T) = \det(A)$.

Dimostrazione

Procediamo per induzione su n. Per $n = 2$ basta guardare la formula. Supponiamo sia vero per $n - 1$; per dimostrarlo per n, basta osservare (esercizio) che la trasposta del minore A_{1i} è esattamente il minore $(i, 1)$ della trasposta A^T, e che l'elemento di posto $(i, 1)$ di A^T è a_{1i}. Allora se calcoliamo il determinante di A^T sviluppando lungo la prima colonna si ha

$$\begin{aligned}\det A^T &= \sum_{i=1}^{n}(-1)^{i+1}a_{1i}\det\big((A^T)_{i1}\big) = \sum_{i=1}^{n}(-1)^{i+1}a_{1i}\det\big((A_{1i})^T\big)\\ &= \sum_{i=1}^{n}(-1)^{i+1}a_{1i}\det(A_{1i}) = \det A,\end{aligned}$$

grazie all'ipotesi induttiva e alla Proposizione 9.6. □

Quale metodo convenga utilizzare per calcolare un determinante dipende dalla situazione. In generale l'eliminazione di Gauss è più efficiente e richiede un minor numero di conti. Altre volte, invece, può convenire usare gli sviluppi di Laplace. Per esempio questo è il caso per matrici con molti zeri: una scelta saggia delle righe o colonne lungo cui sviluppare può ridurre di molto i calcoli.

Esempio 9.7

Vogliamo calcolare il determinante di

$$A = \begin{vmatrix} 0 & 0 & 0 & 27 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 4 & 1 \\ 2 & 0 & 1 & 0 & \sqrt{2} \end{vmatrix}.$$

La terza colonna ha un solo termine non nullo, al posto $(5, 3)$; quindi sviluppando lungo la terza colonna otteniamo

$$\det A = (-1)^{5+3}\det\begin{vmatrix} 0 & 0 & 27 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 3 & 4 & 1 \end{vmatrix}.$$

Ora possiamo sviluppare lungo la prima colonna, ottenendo

$$\det A = -\det\begin{vmatrix} 0 & 27 & 0 \\ -1 & 0 & 0 \\ 3 & 4 & 1 \end{vmatrix}.$$

Infine, sviluppando lungo la prima riga

$$\det A = -(-1)^{1+2}\, 27 \det \begin{vmatrix} -1 & 0 \\ 3 & 1 \end{vmatrix} = -27,$$

risultato ottenuto essenzialmente senza aver dovuto fare alcun conto.

Esempio 9.8

Con un po' di allenamento s'impara anche a semplificare il calcolo del determinante con qualche saggia applicazione della Proposizione 9.1.(2) e dell'Esercizio 9.4. L'idea è di creare più zeri possibile, in modo da semplificare drasticamente gli sviluppi di Laplace.

Prendiamo, per esempio,

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

Sottraendo la terza colonna dalla prima otteniamo

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

A questo punto possiamo sviluppare lungo la seconda riga ottenendo

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

Sommando la seconda riga alla prima e alla terza otteniamo

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

e quindi, sviluppando lungo la prima colonna, $\det A = -3 \det \begin{vmatrix} 2 & 2 \\ 2 & 4 \end{vmatrix} = -12$. Ovviamente, questo era solo uno dei possibili modi di procedere; sperimentane altri, per verificare che otterrai sempre $\det A = -12$.

Un'altra conseguenza dell'esistenza e unicità del determinante è il fatto, già più volte annunciato, che il prodotto dei pivot di una matrice quadrata è, a meno del segno, indipendente dall'eliminazione di Gauss usata.

Corollario 9.8

Siano $p_1, \dots, p_n \in \mathbb{R}$ i pivot ottenuti tramite un'eliminazione di Gauss di una matrice quadrata $A \in M_{n,n}(\mathbb{R})$. Allora $|p_1 \cdots p_n| = |\det A|$. In particolare, il modulo del prodotto dei pivot non dipende dall'eliminazione di Gauss effettuata, ma solo dalla matrice A.

Dimostrazione

Segue immediatamente dall'esistenza e unicità della funzione determinante e dalla Proposizione 9.2.

□

Osservazione 9.8

Abbiamo visto nell'Osservazione 9.1 che il determinante delle matrici 2×2 misura l'area orientata del parallelogramma generato dalle colonne della matrice. Quando (nel Capitolo 12) parleremo del prodotto vettore in $\mathbb{R}^3$ vedremo che il determinante delle matrici 3×3 misura il volume orientato del solido (un parallelepipedo inclinato) generato dalle colonne della matrice.

9.3 Teorema di Binet

Abbiamo visto che il determinante è lineare in ciascuna riga e dunque, grazie al Corollario 9.7, in ciascuna colonna (perché?). Vediamo ora come si comporta rispetto alle operazioni proprie dello spazio delle matrici $M_{n,n}(\mathbb{R})$.

Gli Esercizi 9.13 e 9.14 mostrano che in generale il determinante della somma di matrici ha poco a che vedere con la somma dei determinanti. Per quanto riguarda il prodotto per scalari, si ha $\det(\lambda A) = \lambda^n \det A$ per ogni $A \in M_{n,n}(\mathbb{R})$ – vedi l'Esercizio 9.15. Ma il determinante rivela tutta la sua potenza in connessione con il prodotto di matrici, come mostra il fondamentale *Teorema di Binet*.

Teorema 9.9 (Binet)

Se $A, B \in M_{n,n}(\mathbb{R})$ allora

$$\det(AB) = (\det A)(\det B). \tag{9.7}$$

Dimostrazione

Cominciamo supponendo che $\det B = 0$; in particolare, grazie al Corollario 9.3, le righe (e quindi le colonne) di B sono linearmente dipendenti. Allora esiste un vettore $v \in \mathbb{R}^n$ non nullo tale che $Bv = O$; quindi $(AB)v = O$, cioè anche la matrice AB è singolare. Ne segue che anche le righe della matrice AB sono linearmente dipendenti, per cui $\det(AB) = 0$, e (9.7) in questo caso è verificata.

Supponiamo allora che $\det B \neq 0$. In questo caso possiamo considerare la funzione $f: M_{n,n}(\mathbb{R}) \to \mathbb{R}$ definita da

$$f(A) = \frac{\det(AB)}{\det(B)}.$$

Se riusciamo a dimostrare che f soddisfa le proprietà (A)-(D), l'unicità del determinante darà $f(A) = \det(A)$, da cui la tesi. Cominciamo col provare che f soddisfa (D). Si ha

$$f(I_n) = \frac{\det(I_n B)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1,$$

come voluto. Passiamo a (C). Supponiamo che la riga j_0-esima di A sia somma di due vettori, cioè che $A_{j_0} = A'_{j_0} + A''_{j_0}$. Allora la riga j_0-esima di AB è pure somma di due righe, $(AB)_{j_0} = A'_{j_0}B + A''_{j_0}B$. Siccome il determinante è additivo rispetto alle righe, $\det(AB)$ si decompone in una somma; quindi altrettanto succede per $f(A)$, e la proprietà (C) è verificata.

La dimostrazione della proprietà (B) è del tutto analoga, per non dire più semplice (esercizio).

Rimane la proprietà (A). Se la matrice A ha due righe uguali, anche la matrice AB le ha, quindi $\det(AB) = 0$ e $f(A) = 0$, come dovevamo dimostrare. □

Corollario 9.10

Sia $A \in M_{n,n}(\mathbb{R})$. Allora A è invertibile se e solo se $\det(A) \neq 0$, e in tal caso

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Dimostrazione

La matrice A è invertibile se e solo se le sue righe sono linearmente indipendenti, il che accade se e solo se $\det A \neq 0$. Infine il Teorema 9.9 ci dice che

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

per cui $\det(A^{-1}) = (\det A)^{-1}$. □

Il determinante ci fornisce anche un altro metodo per risolvere un sistema quadrato di matrice non singolare: il *Teorema di Cramer*.

Corollario 9.11 (Cramer)

Sia $Ax = b$ un sistema quadrato di ordine n con A non singolare. Allora l'unica soluzione $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ del sistema è data da

$$\forall i = 1, \ldots, n \qquad\qquad v_i = \frac{\det B_i}{\det A},$$

dove B_i è la matrice ottenuta sostituendo la colonna b dei termini noti alla colonna i-esima di A, cioè $B_i = \left|A^1 \ \cdots \ A^{i-1} \ b \ A^{i+1} \ \cdots \ A^n\right|$.

Dimostrazione

Indichiamo con X_i la matrice ottenuta sostituendo nella matrice identica alla colonna i-esima il vettore v, cioè $X_i = \left|e_1 \ \cdots \ e_{i-1} \ v \ e_{i+1} \ \cdots \ e_n\right|$. Allora si ha $Av = b$ se e solo se $AX_i = B_i$ per ogni $i = 1, \ldots, n$. Prendendo il determinante di entrambi i membri e applicando il Teorema di Binet troviamo allora $\det B_i = \det(AX_i) = (\det A)v_i$, in quanto $\det X_i = v_i$, come puoi vedere sviluppando il determinante di X_i lungo la i-esima riga, e ci siamo. □

Esempio 9.9

Vogliamo trovare la soluzione del sistema

$$\begin{cases} x + 2z = 1, \\ -x + 3y + z = 0, \\ x + 7z = 1, \end{cases}$$

applicando il Teorema di Cramer appena dimostrato. Si calcola subito $\det A = 15$, per cui la matrice A dei coefficienti del sistema è invertibile. Poi

$$B_1 = \begin{vmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 1 & 0 & 7 \end{vmatrix}, \qquad B_2 = \begin{vmatrix} 1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 7 \end{vmatrix}, \qquad B_3 = \begin{vmatrix} 1 & 0 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 1 \end{vmatrix};$$

quindi $\det B_1 = 15$, $\det B_2 = 5$ e $\det B_3 = 0$, e la soluzione è $x = 1$, $y = 1/3$, $z = 0$. □

Vi è un'altra conseguenza del Teorema di Binet che sarà fondamentale in seguito.

Corollario 9.12

Se $A \in M_{n,n}(\mathbb{R})$ *e* $B \in GL_n(\mathbb{R})$, *si ha* $\det(B^{-1}AB) = \det A$.

Dimostrazione

Infatti

$$\det(B^{-1}AB) = \det(B^{-1})\det A \ \det B = (\det B)^{-1}\det B \ \det A = \det A. \qquad \square$$

Dunque due matrici nella stessa classe di similitudine (vedi la Definizione 8.4) hanno lo stesso determinante. Questo ci permette di dare la seguente definizione.

Definizione 9.3

Sia $T\colon V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione n. Allora il *determinante* di T è dato da $\det A$, dove $A \in M_{n,n}(\mathbb{R})$ è la matrice che rappresenta T rispetto a una qualsiasi base di V.

Per la Proposizione 8.2 e il Corollario 9.12 il determinante di un endomorfismo non dipende dalla base scelta, per cui è ben definito. Vedremo più avanti come, assieme al rango, sia una delle caratteristiche più importanti di un endomorfismo.

9.4 Teorema degli orlati

Abbiamo anticipato che è possibile utilizzare il determinante per vedere quando p vettori di $\mathbb{R}^n$ (con $p \leq n$) sono linearmente dipendenti. In realtà, possiamo usare il determinante per trovare il rango di una matrice qualunque, anche rettangolare. Questo è il contenuto del *Teorema degli orlati*, che enunceremo fra poco.

Definizione 9.4

Sia $A \in M_{m,n}(\mathbb{R})$; una *sottomatrice* (quadrata) di *ordine* p di A è una matrice A' ottenuta considerando solo p righe e colonne fissate di A. Se ad A' aggiungiamo un'altra riga e un'altra colonna di A diremo che stiamo *orlando* A'.

Esempio 9.10

Se

$$A = \begin{vmatrix} 1 & 0 & 2 & 4 \\ 0 & 3 & 1 & 4 \\ 1 & 0 & 7 & 3 \end{vmatrix};$$

allora una sottomatrice di ordine 2 ottenuta considerando solo la prima e la seconda riga, e la seconda e la terza colonna, è

$$A' = \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix}.$$

Possiamo orlare A' aggiungendo la terza riga e, per esempio, la prima colonna; questo ci fornisce la sottomatrice di ordine 3

$$A'' = \begin{vmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 1 & 0 & 7 \end{vmatrix}.$$

Teorema 9.13 (degli orlati)

Sia $A \in M_{m,n}(\mathbb{R})$ una matrice. Allora $\operatorname{rg} A = r$ *se e solo se esiste una sottomatrice A' di ordine r di A non singolare, e tutte le sottomatrici di ordine $r+1$ di A ottenute orlando A' hanno determinante nullo.*

Dimostrazione

Supponiamo $\operatorname{rg} A = r$. Allora tutte le $(r+1)$-uple di colonne di A sono linearmente dipendenti; in particolare, quindi, qualunque sottomatrice di ordine $r+1$ ha determinante nullo. D'altra parte, esiste una sottomatrice quadrata di A di ordine r con determinante non nullo. Infatti, siccome $\operatorname{rg} A = r$ esistono r colonne di A linearmente indipendenti. La matrice formata da queste colonne ha rango r; quindi deve avere anche r righe linearmente indipendenti. Considerando queste righe e queste colonne otteniamo una sottomatrice quadrata di ordine r e rango r, e quindi di determinante non nullo.

Viceversa, supponiamo che esista una sottomatrice A' di ordine r di A non singolare, e tale che tutte le sottomatrici di ordine $r+1$ ottenute orlando A' abbiano determinante nullo. Supponiamo, per assurdo, che il rango di A non sia r; vogliamo giungere a una contraddizione. Prima di tutto, l'esistenza di una sottomatrice di ordine r non singolare ci assicura che A contiene almeno r colonne linearmente indipendenti. In particolare, il rango di A non può essere inferiore a r; dovendo essere diverso da r, è per lo meno uguale a $r+1$.

Siccome $\operatorname{rg} A \geq r+1$, possiamo trovare una $(r+1)$-esima colonna da aggiungere alle r colonne di A' in modo da ottenere ancora un insieme di colonne linearmente indipendenti. La matrice formata da queste colonne ha rango $r+1$; quindi possiamo trovare una riga che aggiunta alle r righe di A' formi un insieme di righe linearmente indipendenti. In questo modo abbiamo identificato una sottomatrice non singolare di ordine $r+1$ ottenuta orlando A', contro l'ipotesi. La contraddizione deriva dall'aver supposto $\operatorname{rg} A \neq r$; quindi si deve necessariamente avere $\operatorname{rg} A = r$, e abbiamo finito. □

Esempio 9.11

Vogliamo trovare il rango della matrice

$$A = \begin{vmatrix} 1 & 0 & 1 & 1 \\ 3 & 1 & -4 & 2 \\ 0 & 7 & 7 & -7 \\ 2 & 1 & 3 & 1 \\ 7 & 2 & 9 & 5 \end{vmatrix}.$$

Siccome non è la matrice nulla, il rango è almeno 1; e infatti la prima riga e la prima colonna ci danno una sottomatrice di ordine 1 non singolare, $|1|$. La sottomatrice di ordine 2 nell'angolo in alto a sinistra ha determinante $1 \neq 0$; quindi il rango è almeno 2. La sottomatrice di ordine 3 nell'angolo in alto a sinistra ha determinante

$$\det \begin{vmatrix} 1 & 0 & 1 \\ 3 & 1 & -4 \\ 0 & 7 & 7 \end{vmatrix} = 56 \neq 0,$$

per cui il rango è almeno 3. Per stabilire se il rango è 3 oppure 4, dobbiamo orlare questa sottomatrice e calcolare i determinanti degli orlati. Ci sono solo due modi per orlarla: aggiungendo la quarta riga e la quarta colonna, o aggiungendo la quinta riga e la quarta colonna. Otteniamo allora $\operatorname{rg} A = 3$, in quanto

$$\det\begin{vmatrix}1 & 0 & 1 & 1\\ 3 & 1 & -4 & 2\\ 0 & 7 & 7 & -7\\ 2 & 1 & 3 & 1\end{vmatrix} = 0 \qquad \text{e} \qquad \det\begin{vmatrix}1 & 0 & 1 & 1\\ 3 & 1 & -4 & 2\\ 0 & 7 & 7 & -7\\ 7 & 2 & 9 & 5\end{vmatrix} = 0.$$

Il teorema degli orlati non è il metodo più efficace per calcolare il rango di una matrice (la riduzione a scala è molto più efficiente), ma può essere utile in situazioni particolari (per esempio matrici con molti zeri; ma vedi anche il Capitolo 10).

Esercizi 9

$\boxed{9.1}$ Dimostra, utilizzando il metodo dell'Esempio 9.1, che esiste una e una sola funzione $d\colon M_{3,3}(\mathbb{R}) \to \mathbb{R}$ che soddisfa le proprietà (A)-(D).

$\boxed{9.2}$ Dimostra che non esiste alcuna matrice $A \in M_{2,2}(\mathbb{R})$ tale che $A^T = A$, $A^2 = I$, $\det A > 0$ e $|a_{11} + a_{22}| \neq 2$.

$\boxed{9.3}$ Calcola il determinante delle seguenti matrici:

$$\begin{vmatrix} 0 & 2 & 7 \\ 0 & 0 & \pi \\ 3 & 0 & 0 \end{vmatrix}, \qquad \begin{vmatrix} -2 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & -3 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 3 \\ 3 & 3 & 4 & 4 \\ 5 & 6 & 7 & 8 \end{vmatrix}.$$

$\boxed{9.4}$ Dimostra l'equivalente per le colonne della Proposizione 9.1.(1)-(3).

$\boxed{9.5}$ Siano $x, y, z \in \mathbb{R}$ tali che

$$\det \begin{vmatrix} x & y & z \\ 1 & 1 & 2 \\ 0 & 1 & 3 \end{vmatrix} = 1.$$

Calcola i determinanti delle seguenti matrici:

$$\begin{vmatrix} 3x & 3y & 3z \\ 1/2 & 1/2 & 1 \\ 0 & 1 & 3 \end{vmatrix}, \qquad \begin{vmatrix} x & y & z \\ 1 & 2 & 5 \\ 2x & 1+2y & 3+2z \end{vmatrix}, \qquad \begin{vmatrix} x & y & z \\ 2+2x & 2+2y & 4+2z \\ 1 & 0 & -1 \end{vmatrix}.$$

$\boxed{9.6}$ Calcola, con più tecniche, i determinanti delle seguenti matrici:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 5 \\ 3 & 0 & 1 & -1 \\ 2 & 1 & 0 & 2 \end{vmatrix}, \qquad \begin{vmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{vmatrix}.$$

$\boxed{9.7}$ Al variare di $k \in \mathbb{R}$ calcola i determinanti delle seguenti matrici:

$$\begin{vmatrix} 2 & k & 1 \\ 1 & 2k & 5 \\ k-1 & 1 & -2 \end{vmatrix}, \qquad \begin{vmatrix} -3 & 1 & k & 0 \\ 4 & k-1 & 5 & 1 \\ 0 & k-1 & 2 & 7 \\ k & 1 & -1 & -2 \end{vmatrix}.$$

9.8 Al variare di $k \in \mathbb{C}$ calcola i determinanti delle seguenti matrici:

$$\begin{vmatrix} k-i & 1 & -1 \\ 0 & 1-ik & 2-3i \\ 1+2i & i-1 & k+1 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 0 & k-1 & 1 \\ k & 1 & 2i & 1-i \\ ik+3 & 0 & 1-i & 1-3i \\ 1 & k & ik & 1 \end{vmatrix}$$

9.9 Verifica che la matrice $\begin{vmatrix} 5k & k+3 \\ k-1 & k \end{vmatrix}$ è invertibile per ogni valore di $k \in \mathbb{R}$. Cosa accade se $k \in \mathbb{C}$?

9.10 Verifica che la matrice $\begin{vmatrix} k & k+i \\ k-i & k \end{vmatrix}$ è invertibile per ogni valore di $k \in \mathbb{C}$.

9.11 Siano $a_1, a_2, a_3, a_4 \in \mathbb{C}$. Calcola i determinanti delle seguenti matrici:

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^4 & a_2^4 & a_3^4 \end{vmatrix},$$

e generalizza a matrici $n \times n$ la formula ottenuta con $n = 3, 4$ nei primi due casi. (*Suggerimento:* moltiplica ogni riga per a_1 e sottraila dalla successiva.)

9.12 Sia $A \in M_{n,n}(\mathbb{R})$ una matrice della forma

$$A = \left|\begin{array}{c|c} B & C \\ \hline O & D \end{array}\right|,$$

con $B \in M_{h,h}(\mathbb{R})$, $C \in M_{h,n-h}(\mathbb{R})$, $D \in M_{n-h,n-h}(\mathbb{R})$ e $h < n$. Dimostra che $\det A = (\det B)(\det D)$. (*Suggerimento:* procedi per induzione su h).

9.13 Trova due matrici A, $B \in M_{2,2}(\mathbb{R})$ tali che $\det A + \det B \neq 0 = \det(A + B)$.

9.14 Trova due matrici $C, D \in M_{2,2}(\mathbb{R})$ non nulle tali che $\det(C + D) = \det(C) + \det(D)$.

9.15 Se $A \in M_{n,n}(\mathbb{R})$ e $\lambda \in \mathbb{R}$, allora dimostra che $\det(\lambda A) = \lambda^n \det(A)$.

9.16 Sia $A \in M_{n,n}(\mathbb{R})$ una matrice antisimmetrica, cioè tale che $A^T = -A$. Dimostra che se n è dispari allora $\det A = 0$.

9.17 Sia $A \in M_{n,n}(\mathbb{R})$ e $1 \leq i, j \leq n$. Dimostra che

$$\sum_{k=1}^{n} (-1)^{j+k} a_{ik} \det(A_{jk}) = \begin{cases} 0 & \text{se } i \neq j, \\ \det A & \text{se } i = j. \end{cases}$$

(*Suggerimento:* se $i \neq j$ è il determinante di una matrice con due righe uguali.)

9.18 Se $A \in M_{n,n}(\mathbb{R})$, la *matrice cofattore* di A è la matrice $\mathrm{Cof}(A)$ che al posto (i, j) ha il numero $(-1)^{j+i} \det(A_{ji})$. Dimostra che

$$A \cdot \mathrm{Cof}(A) = (\det A) I_n.$$

In particolare, se A è invertibile dimostra che $A^{-1} = (\det A)^{-1} \mathrm{Cof}(A)$ e deduci quindi che $\det(\mathrm{Cof}(A)) = (\det A)^{n-1}$. (*Suggerimento:* usa l'esercizio precedente e il Corollario 9.10.)

9.19 Dimostra che la seguente matrice è invertibile e trovane l'inversa:

$$\begin{vmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{vmatrix}.$$

9.20 Risolvi usando il Teorema di Cramer i seguenti sistemi lineari:

$$\begin{cases} x + 7y + 3z = 2, \\ -x + 2z = -1, \\ 3x + y + z = 1; \end{cases} \qquad \begin{cases} x + 2y + 3z = -1, \\ y - 2z = -2, \\ 2x + 3y + 4z = 0; \end{cases} \qquad \begin{cases} 888445x + 887112y = \\ 887112x + 885781y = \end{cases}$$

9.21 Verifica che la matrice incompleta del sistema:

$$\begin{cases} 2kx + (k-2)y = k, \\ (k+1)x + 3ky = k - 1, \end{cases}$$

è invertibile per ogni $k \in \mathbb{R}$ e risolvi il sistema usando il Teorema di Cramer

9.22 Verifica che la matrice incompleta del sistema:

$$\begin{cases} kx - y + kz = 3, \\ kx + y + (k-1)z = 2k - 3, \\ (k-1)x + ky - kz = k + 4, \end{cases}$$

è invertibile per ogni $k \in \mathbb{R}$ e risolvi il sistema usando il Teorema di Cramer.

9.23 Verifica che la matrice incompleta del sistema:

$$\begin{cases} (k-3)x + (ik+1)y = 2k - i + 7, \\ -ikx + (k+3-i)y = 3ik - 2 + i, \end{cases}$$

è invertibile per ogni $k \in \mathbb{C}$ e risolvi il sistema usando il Teorema di Cramer.

9.24 Verifica che la matrice incompleta del sistema:

$$\begin{cases} (k+83)x + 10y + (k+59)z = 2i - 1, \\ 15x + 3y + 8z = 2ik - 3 + 2i, \\ (3k+27)x + (3k+25)z = 2 - (1-i)k, \end{cases}$$

è invertibile per ogni $k \in \mathbb{C}$ e risolvi il sistema usando il Teorema di Cramer

9.25 Sia $A \in M_{n,n}(\mathbb{R})$, e considera le applicazioni lineari $S, T\colon M_{n,n}(\mathbb{R}) \to M_{n,n}(\mathbb{R})$ da $S(X) = AX - XA$ e $T(X) = AX$. Dimostra che $\det S = 0$ e $\det \Gamma = (\det A)^n$

9.26 Sia $T\colon \mathbb{R}_3[t] \to \mathbb{R}_3[t]$ l'endomorfismo $T(p) = p'$. Calcola $\det T$.

9.27 Considera $T : \mathbb{C}_3[t] \to \mathbb{C}_3[t]$ data da $T(p) = (1+i)\cdot p + (2-i)\cdot p'$. Calcola de

9.28 Calcola, al variare di $t \in \mathbb{R}$, il rango delle seguenti matrici.

$$\begin{vmatrix} 1 & -1 & t & -2 \\ 2 & -1 & 1 & -1 \\ 3 & t+1 & 4 & 6 \end{vmatrix}, \quad \begin{vmatrix} 0 & -1 & 2 \\ 1 & 2 & -1 \\ 3 & t+3 & 3 \\ 4 & 6 & t \end{vmatrix}, \quad \begin{vmatrix} t+1 & 1 & 1 \\ t+2 & t^2+1 & 2 \\ 2t & 1 & 1 \end{vmatrix}.$$

9.29 Sia $A \in M_{n,n}(\mathbb{C})$. Dimostra che $\det A^H = \det \overline{A} = \overline{\det A}$. Deducine che se $A^H = A$, allora $\det A$ è un numero reale.

9.30 Calcola il rango della matrice

$$\begin{vmatrix} 1 & -1-i & 3 & 0 & 1 & 4 & 4 \\ 1 & 2+2i & -5 & i-2 & -1 & 0 & -1 \\ 1 & 2+2i & 3 & i-2 & 1 & 2 & 5 \\ 1-i & 4 & -5+5i & 3i-1 & i-1 & 2-2i & 0 \\ 1 & -4-4i & 7 & 2-i & 2 & 1 & 3 \end{vmatrix}$$

usando il teorema degli orlati.

9.31 Una *permutazione* dell'insieme $I_n = \{1, \ldots, n\}$ è una bigezione $\sigma \colon I_n \to I_n$; una *trasposizione* è una permutazione che scambia solo due elementi lasciando fermi tutti gli altri.

(1) Dimostra che ogni permutazione di I_n può essere ottenuta come composizione di trasposizioni. (*Suggerimento*: procedi per induzione su n.)

(2) Se σ è una permutazione di I_n, indichiamo con $\epsilon(\sigma)$ il numero di *inversioni* di σ, cioè la cardinalità dell'insieme $\{(i,j) \in I_n \times I_n \mid i < j \text{ e } \sigma(i) > \sigma(j)\}$. Il *segno* di σ è $\operatorname{sgn}(\sigma) = (-1)^{\epsilon(\sigma)}$. Dimostra che $\operatorname{sgn}(\tau) = -1$ per ogni trasposizione τ, e che $\operatorname{sgn}(\sigma_1 \circ \sigma_2) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$ per ogni coppia di permutazioni σ_1 e σ_2. Deduci che se σ è ottenuta come composizione di m trasposizioni allora $\operatorname{sgn}(\sigma) = (-1)^m$. (*Suggerimento*: dimostra che

$$\prod_{1 \le i < j \le n} \big(\sigma(j) - \sigma(i)\big) = \operatorname{sgn}(\sigma) \prod_{1 \le i < j \le n} (j - i)$$

per ogni permutazione σ.)

9.32 Dimostra che

$$\det A = \sum_{\sigma} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

per ogni matrice $A = (a_{ij}) \in M_{n,n}(R)$, dove la somma è estesa a tutte le permutazioni σ di I_n (vedi l'esercizio precedente). (*Suggerimento*: usa il Corollario 9.4.)