Capitolo 6

Algebra matriciale

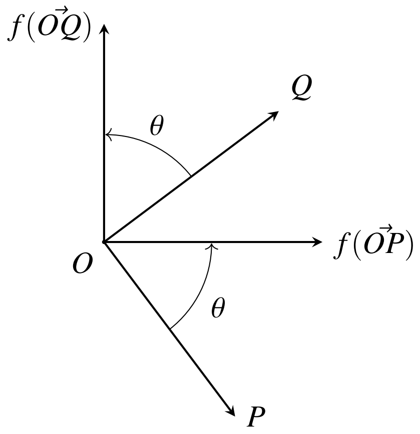
In questo capitolo vedremo che le matrici, oltre a rappresentare sistemi di equazioni lineari, formano una particolare classe di funzioni che comprende molte importanti trasformazioni geometriche. Questo ci consentirà di definire una operazione tra matrici, che chiameremo prodotto, e della quale studieremo le proprietà algebriche.

6.1 Matrici e trasformazioni

Mostriamo in questa sezione come alcune importanti trasformazioni geometriche possano essere rappresentate mediante una matrice.

Più precisamente, supponiamo di aver fissato nel piano una base ortonormale $\{\vec{OP}_1, \vec{OP}_2\}$. Come abbiamo visto nel Capitolo 4, a ogni punto P del piano risulta associata allora una coppia (x_1, x_2) di numeri reali (le sue coordinate rispetto al sistema di riferimento fissato) definiti dalla condizione che $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$.

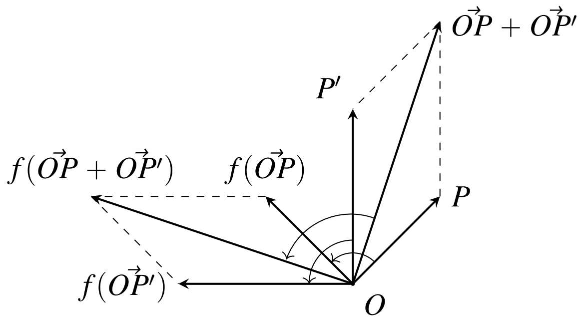
Supponiamo ora di voler applicare ai punti del piano la trasformazione f che consiste in una rotazione di un certo angolo fissato θ attorno all'origine O. Dal momento che ruotare il punto P attorno a O equivale a ruotare il vettore $\vec{OP}$ dello stesso angolo, possiamo considerare questa rotazione come una funzione sui vettori anziché sui punti. Possiamo quindi descrivere la funzione f che va dall'insieme dei vettori applicati in se stesso e che associa a ogni vettore $\vec{OP}$ il vettore $f(\vec{OP})$ che si ottiene ruotando $\vec{OP}$ dell'angolo θ.



Ora ci poniamo l'obiettivo di rappresentare tale trasformazione in coordinate: se (x_1, x_2) sono le coordinate di $\vec{OP}$, vogliamo determinare le coordinate di $f(\vec{OP})$. A questo scopo, osserviamo preliminarmente che la rotazione soddisfa le seguenti due proprietà.

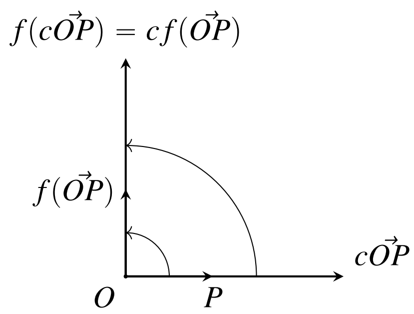
(1) Siano $\vec{OP}$ e $\vec{OP'}$ due vettori qualunque, Allora sommarli e successivamente ruotare il vettore risultante è equivalente a ruotarli prima e successivamente sommare i vettori così ottenuti. In formule

$$f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'}) \tag{6.1}$$



(2) Sia $\vec{OP}$ un qualunque vettore e $c \in \mathbb{R}$. Allora moltiplicare il vettore $\vec{OP}$ per c e successivamente ruotarlo equivale a ruotarlo prima ed in seguito moltiplicarlo per c. In formule

$$f(c\vec{OP}) = cf(\vec{OP}) \tag{6.2}$$



Le due proprietà appena osservate ci permettono di rappresentare la rotazione in coordinate. Infatti, se P ha coordinate (x_1, x_2), per definizione si ha $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$. Ma allora per il vettore ruotato $f(\vec{OP})$ si ha

$$f(\vec{OP}) = f(x_1\vec{OP}_1 + x_2\vec{OP}_2) =$$

(applicando la proprietà (6.1))

$$= f(x_1\vec{OP}_1) + f(x_2\vec{OP}_2) =$$

(applicando la proprietà (6.2) a entrambi gli addendi)

$$f(\vec{OP}) = x_1 f(\vec{OP}_1) + x_2 f(\vec{OP}_2). \tag{6.3}$$

Ora, essendo $f(\vec{OP}_1)$ e $f(\vec{OP}_2)$ vettori del piano, essi potranno essere scritti come combinazione dei vettori $\vec{OP}_1$ e $\vec{OP}_2$ della base fissata. Cioè si hanno le uguaglianze

$$f(\vec{OP}_1) = a_{11}\vec{OP}_1 + a_{21}\vec{OP}_2 \tag{6.4}$$

$$f(\vec{OP}_2) = a_{12}\vec{OP}_1 + a_{22}\vec{OP}_2 \tag{6.5}$$

per certi numeri reali $a_{11}, a_{21}, a_{12}, a_{22}$. Ma allora, sostituendo (6.4) e (6.5) nella (6.3), si ottiene

$$f(\vec{OP}) = x_1(a_{11}\vec{OP}_1 + a_{21}\vec{OP}_2) + x_2(a_{12}\vec{OP}_1 + a_{22}\vec{OP}_2)$$

e, raccogliendo i coefficienti di $\vec{OP}_1$ e $\vec{OP}_2$ secondo le proprietà (4.1)-(4.8), si ottiene

$$f(\vec{OP}) = (a_{11}x_1 + a_{12}x_2)\vec{OP}_1 + (a_{21}x_1 + a_{22}x_2)\vec{OP}_2. \tag{6.6}$$

Da quest'ultima uguaglianza, vediamo che le coordinate (sempre rispetto alla base fissata $\vec{OP}_1, \vec{OP}_2$) del vettore $f(\vec{OP})$ ottenuto ruotando il vettore $\vec{OP}$ di coordinate (x_1, x_2) sono date dalla coppia $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$.
Quindi conoscere la rotazione in coordinate è equivalente a conoscere i coefficienti

$a_{11}, a_{21}, a_{12}, a_{22}$. Disponendo tali numeri nella matrice

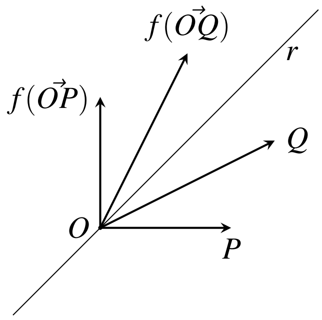
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

possiamo dire che questa rappresenta la rotazione, nel senso che la funzione

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \tag{6.7}$$

che associa a ogni coppia[1] (x_1, x_2) la coppia $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$ fornisce le coordinate di un vettore ruotato a partire dalle coordinate del vettore di partenza.

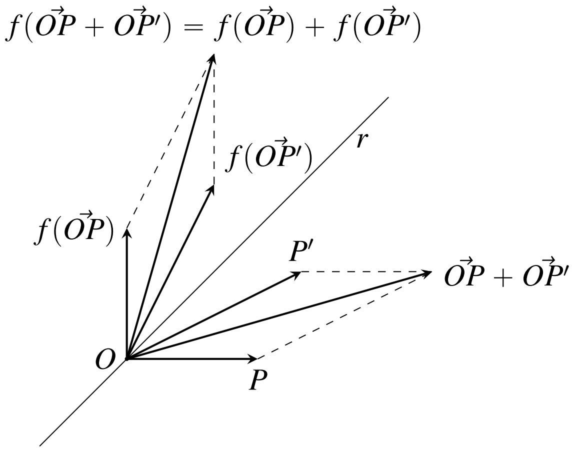
Lo stesso ragionamento fatto per arrivare alla matrice che rappresenta una rotazione nel piano può essere ripetuto anche per altre importanti trasformazioni geometriche che soddisfano le proprietà (6.1) e (6.2). Ad esempio la riflessione rispetto a una retta per O, che manda ogni vettore $\vec{OP}$ nel vettore simmetrico rispetto alla retta, come nel seguente disegno



ha esattamente le stesse proprietà (6.1) e (6.2) viste nel caso della rotazione. Infatti

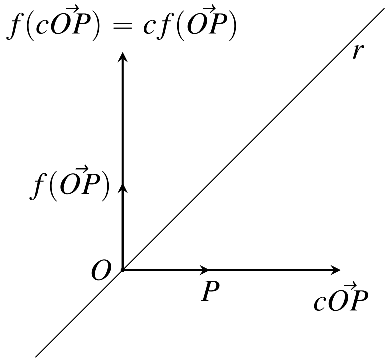
(1) dati due vettori $\vec{OP}$ e $\vec{OP'}$, sommarli e poi riflettere il vettore risultante oppure prima rifletterli e poi sommare i vettori riflessi è equivalente

[1]Qui e nel seguito indicheremo spesso le n-uple in colonna anziché in riga. Questo ci è comodo per questioni di spazio ma, sopratutto, è giustificato dall'utilizzo del prodotto di matrici che vedremo nel seguito del capitolo.



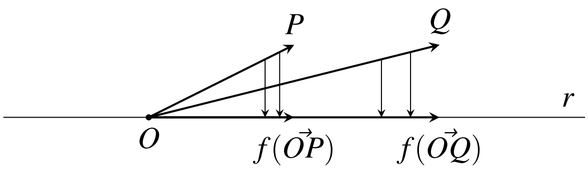
(cioè $f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'})$).

(2) dato un vettore $\vec{OP}$ e un numero reale c, moltiplicare il vettore per c e poi rifletterlo oppure prima rifletterlo e poi moltiplicarlo per c è equivalente

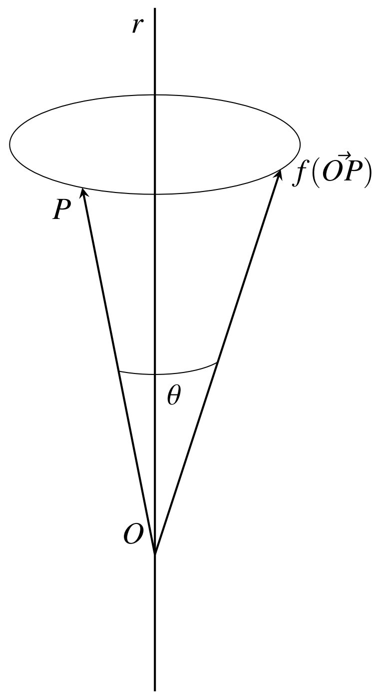


(cioè $f(c\vec{OP}) = cf(\vec{OP})$).

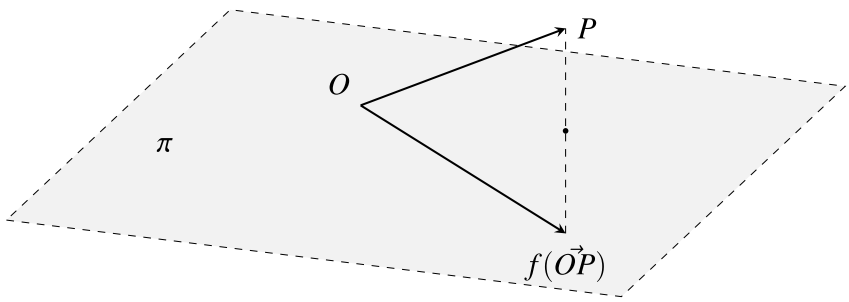
Quindi possiamo ripetere esattamente gli stessi passaggi visti sopra per la rotazione e rappresentare anche la riflessione mediante una matrice con due righe e due colonne.
Come terzo esempio di importante trasformazione geometrica che soddisfa le proprietà (6.1) e (6.2), e quindi può essere rappresentata da una matrice, citiamo la proiezione su una retta fissata passante per O.



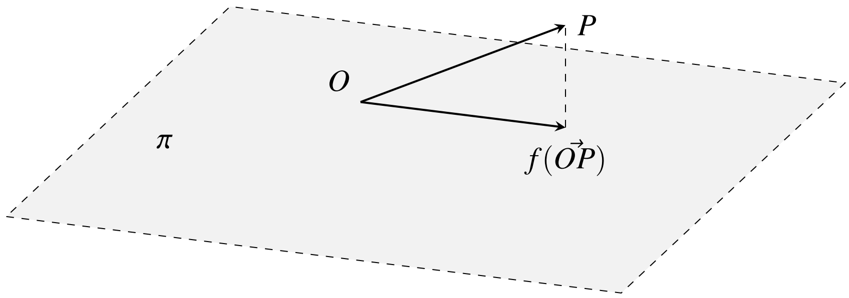
Quanto visto fino ad ora sul piano risulta vero anche per alcune importanti trasformazioni geometriche dello spazio tridimensionale. Ad esempio, non è difficile vedere che le proprietà (6.1) e (6.2) sono soddisfatte anche dalla rotazione di un angolo fissato attorno a una retta fissata passante per O (detta *asse della rotazione*)



dalla riflessione rispetto a un piano passante per O



e dalla proiezione ortogonale su un piano passante per l'origine O.



In tutti e tre questi esempi, essendo soddisfatte le proprietà $f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'})$ e $f(c\vec{OP}) = cf(\vec{OP})$, possiamo ripetere gli stessi passaggi visti nel caso della rotazione nel piano, con la differenza che stavolta avremo una base $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$ formata da tre vettori. Per ogni vettore $\vec{OP}$ avremo $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$ e, con ragionamenti simili a quelli visti sopra per ottenere la (6.6), si otterrà

$$\begin{aligned} f(\vec{OP}) = &(a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\vec{OP}_1 \\ &+ (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)\vec{OP}_2 \\ &+ (a_{31}x_1 + a_{32}x_2 + a_{33}x_3)\vec{OP}_3. \end{aligned}$$

Ovvero le coordinate (rispetto alla base fissata $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$) del vettore $f(\vec{OP})$ ottenuto trasformando il vettore $\vec{OP}$ di coordinate (x_1, x_2, x_3) sono date dalla terna

$$(a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_1 + a_{22}x_2 + a_{23}x_3, a_{31}x_1 + a_{32}x_2 + a_{33}x_3). \tag{6.8}$$

Quindi la trasformazione data si traduce in coordinate in una funzione del tipo

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} \tag{6.9}$$

In questo caso, conoscere la trasformazione considerata in coordinate equivale a conoscere la matrice

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

con tre righe e tre colonne che contiene i coefficienti che appaiono nella (6.8).

Definizione 6.1. Dati due spazi vettoriali V e W (si veda l'Osservazione 4.1), una funzione $f : V \to W$ si dice **funzione lineare** (o **applicazione lineare**) se soddisfa le proprietà

(1) $f(v + v') = f(v) + f(v')$ per qualunque $v, v' \in V$

(2) $f(cv) = cf(v)$ per ogni $v \in V$ e ogni scalare c

Le formule (6.7) e (6.9), che ci danno la rappresentazione in coordinate delle trasformazioni geometriche viste sopra, sono casi particolari di funzioni del tipo

$$f : \mathbb{K}^n \to \mathbb{K}^m, \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \tag{6.10}$$

(dove $\mathbb{K}$ è un campo, per esempio $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) che risultano più in generale determinate dalla matrice con m righe e n colonne

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

Chiameremo l'applicazione (6.10) associata a tale matrice l'**applicazione lineare determinata da** A e la denoteremo con L_A.
Nel paragrafo successivo concentreremo la nostra attenzione su tali funzioni. Usando la nozione di composizione di funzioni vista nel Capitolo 2 definiremo un'operazione di prodotto tra matrici. Potremo quindi considerare l'insieme delle matrici, analogamente a quanto fatto per gli insiemi numerici, come un insieme sul quale è definita un'operazione della quale studiare le proprietà. In particolare vedremo che certi insiemi di matrici sono anelli.
Introduciamo alcune notazioni preliminari.

Definizione 6.2. Una matrice con m righe e n colonne è detta **matrice** m **per** n. L'insieme delle matrici m per n le cui entrate appartengono a un campo dato $\mathbb{K}$ (ad esempio il campo $\mathbb{R}$ dei numeri reali, o quello $\mathbb{C}$ dei numeri complessi) si denota $M_{m,n}(\mathbb{K})$.
Nel caso in cui $m = n$, ovvero il numero di righe sia uguale al numero di colonne, si usa scrivere $M_n(\mathbb{K})$ al posto di $M_{n,n}(\mathbb{K})$.
Un elemento di $M_n(\mathbb{K})$ si dice **matrice quadrata di ordine** n.

6.2 Composizione e prodotto di matrici

In questo paragrafo vogliamo calcolare la matrice che rappresenta la composizione di due funzioni del tipo (6.10) visto alla fine del paragrafo precedente. Più precisamente, consideriamo le applicazioni lineari $f = L_A : \mathbb{K}^n \to \mathbb{K}^m$ definita da

$$f \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \tag{6.11}$$

e $g = L_B : \mathbb{K}^p \to \mathbb{K}^n$ data da

$$g \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2p}y_p \\ \vdots \\ b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p \end{pmatrix} \tag{6.12}$$

In base a quanto visto nella Sezione 2.6, la composizione $f \circ g$ può essere calcolata in quanto il codominio di g, ovvero $\mathbb{K}^n$, è anche il dominio di f. Si ha

$$(f \circ g) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = f \left(g \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \right) = f \begin{pmatrix} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2p}y_p \\ \vdots \\ b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{1n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ a_{21}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{2n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ \vdots \\ a_{m1}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{mn}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \end{pmatrix} =$$

(raccogliendo $y_1, y_2, \cdots, y_p$)

$$= \begin{pmatrix} (a_{11}b_{11} + \cdots + a_{1n}b_{n1})y_1 + \cdots + (a_{11}b_{1p} + \cdots + a_{1n}b_{np})y_p \\ (a_{21}b_{11} + \cdots + a_{2n}b_{n1})y_1 + \cdots + (a_{21}b_{1p} + \cdots + a_{2n}b_{np})y_p \\ \vdots \\ (a_{m1}b_{11} + \cdots + a_{mn}b_{n1})y_1 + \cdots + (a_{m1}b_{1n} + \cdots + a_{mn}b_{np})y_p \end{pmatrix}$$

Da quest'ultima espressione vediamo che la composizione $f \circ g$ è ancora una funzione del tipo (6.10), cioè è determinata da una matrice C, e più precisamente

$$C = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1n} + \cdots + a_{mn}b_{np} \end{pmatrix}$$

Ora, notiamo che le entrate c_{ij} di tale matrice sono tutte espressioni del tipo

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} \tag{6.13}$$

in cui il primo indice dell'entrata di A è fisso (sempre uguale a i) e il secondo indice dell'entrata di B è fisso (sempre uguale a j). Invece gli indici "interni" (il secondo dell'entrata di A e il primo dell'entrata di B) variano da 1 a n. Usando la notazione di sommatoria, potremmo quindi scrivere $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. Possiamo quindi definire il prodotto di due matrici A e B come la matrice $C = AB$ che rappresenta la composizione $L_A \circ L_B$ delle funzioni associate ad A e B.

Definizione 6.3. La matrice C le cui entrate sono date dalla (6.13) è detta **prodotto di** A **per** B, e si scrive $C = AB$.

Ricordiamo che la composizione $L_A \circ L_B$ delle due funzioni L_A e L_B può essere fatta solo sotto opportune condizioni (il codominio di $L_B : \mathbb{K}^p \to \mathbb{K}^n$ deve essere uguale al dominio di $L_A : \mathbb{K}^n \to \mathbb{K}^m$). Di conseguenza anche il prodotto di due matrici può essere fatto solo sotto opportune condizioni. Più precisamente, dal momento che la matrice A di L_A ha m righe e n colonne, mentre la matrice B di L_B ha n righe e p colonne, si possono moltiplicare tra loro due matrici A e B (in quest'ordine) se e solo se *il numero di colonne di A è uguale al numero di righe di B*. In tal caso, il risultato AB è una matrice con m righe e p colonne:

$$A \in M_{m,n}(\mathbb{K}),\ B \in M_{n,p}(\mathbb{K}) \Rightarrow AB \in M_{m,p}(\mathbb{K}).$$

Il prodotto di matrici dato dalla Definizione 6.3 si chiama anche **prodotto righe per colonne**. Il motivo è che nell'espressione (6.13) della generica entrata di posto $i\,j$ appaiono tutte e sole le entrate con primo indice i di A e tutte e sole le entrate con secondo indice j di B. Dal momento che il primo è l'indice di riga e il secondo quello di colonna, l'entrata di posto $i\,j$ di AB è la somma dei prodotti entrata per entrata della

i-esima riga di A per la j-esima colonna di B.

$$\begin{pmatrix} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \\ \end{pmatrix} \begin{pmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{pmatrix} = \begin{pmatrix} & \cdots & \\ \cdots & a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} & \cdots \\ & \cdots & \end{pmatrix}$$

Esempio 6.4. Consideriamo le due matrici

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{pmatrix} \in M_{3,2}(\mathbb{R}), \quad B = \begin{pmatrix} 5 & 2 & 0 & 1 \\ 3 & 1 & -1 & 2 \end{pmatrix} \in M_{2,4}(\mathbb{R}).$$

In base a quello che abbiamo detto, il prodotto di A per B è ben definito e il risultato AB sarà una matrice 3 per 4.
Per trovare l'entrata di AB che sta *nella prima riga* e *prima colonna* si considerano *la prima riga di* A $\begin{pmatrix} 1 & 2 \end{pmatrix}$, la *prima colonna* di B $\left(\begin{smallmatrix} 5 \\ 3 \end{smallmatrix}\right)$, se ne moltiplicano gli elementi corrispondenti (il primo con il primo, il secondo con il secondo) e si sommano i risultati: $1 \cdot 5 + 2 \cdot 3 = 5 + 6 = 11$. Quindi l'entrata 1 1 di AB sarà 11.
Analogamente, per trovare l'entrata di AB che sta *nella prima riga* e *seconda colonna* si considerano *la prima riga di* A $\begin{pmatrix} 1 & 2 \end{pmatrix}$, la *seconda colonna* di B $\left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$, se ne moltiplicano gli elementi corrispondenti e si sommano i risultati: $1 \cdot 2 + 2 \cdot 1 = 2 + 2 = 4$. Quindi l'entrata 1 2 di AB sarà 4.
Svolgendo questo calcolo per tutte le entrate, si vede che

$$AB = \begin{pmatrix} 11 & 4 & -2 & 5 \\ 35 & 13 & -5 & 14 \\ 33 & 12 & -6 & 15 \end{pmatrix}.$$

Vediamo ora le proprietà fondamentali del prodotto di matrici.
In primo luogo, è importante osservare che tale prodotto *non gode della proprietà commutativa*. Date due matrici A e B e ammesso che si possano eseguire entrambi i prodotti AB e BA, in generale si avrà $AB \neq BA$.[2]

Esempio 6.5. Siano $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e $B = \left(\begin{smallmatrix} -1 & 0 \\ 5 & 2 \end{smallmatrix}\right)$. Allora, per definizione di prodotto righe per colonne vediamo che

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 9 & 4 \\ 17 & 8 \end{pmatrix}$$

[2] Questo fatto può essere spiegato ricordando che il prodotto tra matrici rappresenta la composizione delle funzioni da loro determinate, e come sappiamo in generale la composizione non gode della proprietà commutativa (si veda l'Esempio 2.37).

$$BA = \begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 11 & 18 \end{pmatrix}$$

ovvero $AB \neq BA$.

Si può dimostrare invece che il prodotto di matrici soddisfa la proprietà associativa, ovvero $(AB)C = A(BC)$. Questa è una conseguenza della stessa proprietà per la composizione di funzioni.

Esempio 6.6. Siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ e $C = \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix}$. Allora

$$(AB)C = \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \right) \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 16 & 10 \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \left(\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 10 \\ 1 & -5 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 16 & 10 \end{pmatrix}$$

Ci chiediamo ora se il prodotto di matrici ammetta un elemento neutro come nel caso del prodotto tra numeri reali, in cui esiste il numero 1 per cui $a \cdot 1 = 1 \cdot a = a$ per ogni a. La risposta è affermativa. Più precisamente, per ogni n consideriamo la seguente matrice con n righe e n colonne

$$Id_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{6.14}$$

ovvero la matrice che le cui entrate con lo stesso indice di riga e di colonna (a_{11}, a_{22} etc.) sono 1 e tutte le altre entrate sono 0.

Tale matrice si chiama **matrice identica** (o **matrice identità**) **di ordine** n e si denota con Id_n. Ad esempio,

$$Id_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Id_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Osservazione 6.7. Data una matrice quadrata A di ordine n, le sue entrate $a_{11}, a_{22}, \cdots, a_{nn}$ che hanno stesso indice di riga e di colonna formano la cosiddetta **diagonale della matrice**. La matrice identica Id_n può essere quindi descritta come la matrice che ha 1 sulla diagonale e 0 nelle altre entrate. Le entrate di Id_n si denotano solitamente con il simbolo

δ_{ij}, detto **delta di Kronecker**, definito da

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

Ora, si può verificare che, per ogni $A \in M_{m,n}(\mathbb{K})$ si ha

$$AId_n = A, \quad Id_m A = A$$

(l'ordine della matrice identica cambia perché deve essere tale che si possa svolgere il prodotto). Quindi la matrice identica è l'elemento neutro per il prodotto righe per colonne.

Esempio 6.8.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Osservazione 6.9. Si noti che la matrice identica Id_n rappresenta esattamente la funzione identica $id_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$ che manda ogni elemento in se stesso. Infatti

$$id_{\mathbb{K}^n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1x_1 + 0x_2 + \cdots + 0x_n \\ 0x_1 + 1x_2 + \cdots + 0x_n \\ \vdots \\ 0x_1 + 0x_2 + \cdots + 1x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{6.15}$$

Questo spiega perché tale matrice sia l'elemento neutro per il prodotto. Infatti il prodotto tra matrici rappresenta la composizione delle applicazioni corrispondenti e la funzione identica è l'elemento neutro per la composizione, come visto a pagina 50.

Ora, ispirandoci a quello che succede in un campo numerico, dove ogni elemento a diverso da zero ha un inverso rispetto al prodotto (ovvero un elemento b tale che $ab = ba = 1$) e dalla composizione di funzioni diamo la seguente

Definizione 6.10. Una matrice A è detta **invertibile** se esiste una matrice B tale che $AB = Id$ e $BA = Id$, dove Id è una matrice identica di ordine opportuno, dipendente da A. In tal caso la matrice B sarà detta **inversa di** A e verrà denotata con A^{-1}.

Si ha il seguente risultato che dimostreremo in seguito.

Teorema 6.11. *Una matrice $A \in M_{m,n}(\mathbb{K})$ è invertibile se e solo se $m = n$ e il rango di A è uguale a n.*

In altre parole, sono invertibili solo le matrici quadrate $A \in M_n(\mathbb{K})$ e di rango massimo, ovvero in cui nessuna riga è dipendente dalle altre. Sotto tale condizione, affermiamo che esiste una matrice $B \in M_n(\mathbb{K})$ (cioè quadrata dello stesso ordine) per cui $AB = Id_n$ e $BA = Id_n$.

Ora, per capire invece quale sia il ruolo giocato dal rango, facciamo alcune considerazioni che ci suggeriranno anche un modo per calcolare l'inversa quando essa esiste.

Concentriamoci sulla condizione $AB = Id_n$, che equivale a

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \cdots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Tenuto conto della definizione di prodotto righe per colonne, vediamo che moltiplicando le righe di A per la prima colonna di B deve valere

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} = 1 \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} = 0 \\ \cdots \\ a_{n1}b_{11} + a_{n2}b_{21} + \cdots + a_{nn}b_{n1} = 0 \end{cases}$$

ovvero la prima colonna di B soddisfa il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{6.16}$$

con matrice dei coefficienti uguale a A e termini noti uguali alla prima colonna della matrice identica.

Analogamente, moltiplicando le righe di A per la seconda colonna di B si vede che devono essere soddisfatte le seguenti uguaglianze

$$\begin{cases} a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2} = 0 \\ a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2} = 1 \\ \cdots \\ a_{n1}b_{12} + a_{n2}b_{22} + \cdots + a_{nn}b_{n2} = 0 \end{cases}$$

ovvero la seconda colonna di B deve soddisfare il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 1 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{6.17}$$

la cui matrice dei coefficienti è ancora A ma con termini noti uguali alla seconda colonna della matrice identica. Possiamo ragionare allo stesso modo fino all'ultima colonna di B che dovrà soddisfare

$$\begin{cases} a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1n}b_{nn} = 0 \\ a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2n}b_{nn} = 0 \\ \dots \\ a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn} = 1 \end{cases}$$

ovvero essere soluzione del sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 1 \end{cases} \tag{6.18}$$

Riassumendo, si ha $AB = Id_n$ se e solo se le colonne

$$\begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix}, \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{pmatrix}, \dots, \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix}$$

sono soluzione rispettivamente degli n sistemi (6.16), (6.17), ..., (6.18).
Ora, come sappiamo, per verificare se un sistema ha soluzioni basta scrivere la sua matrice completa e ridurla a gradini col metodo di Gauss-Jordan.
Poiché i sistemi (6.16), (6.17), ..., (6.18) hanno tutti la stessa matrice dei coefficienti, cioè A, e differiscono solo per i termini noti, le operazioni elementari che eseguiremo per ridurli sono le stesse, quindi conviene svolgerle una sola volta. A questo scopo, basta scrivere la matrice

$$(A|Id_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ & & \cdots & & & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right) \tag{6.19}$$

ottenuta affiancando tutti i termini noti dei sistemi (6.16), (6.17), ..., (6.18) e risolverli contemporaneamente con una sola riduzione.

Se la matrice A avesse rango minore di n, in seguito alla riduzione si annullerebbe almeno una sua riga. In tal caso, affinché gli n sistemi rappresentati dalla (6.19) siano tutti compatibili, dovrebbe annullarsi almeno una riga del blocco Id_n sulla destra che rappresenta i termini noti (altrimenti in uno dei sistemi corrispondenti ci sarebbe un'uguaglianza falsa $0 = b \neq 0$).

Tuttavia, questo non può mai accadere. Se in seguito alle stesse operazioni elementari applicate per ridurre A si annullasse anche una riga del blocco di destra, allora le righe di Id_n sarebbero linearmente dipendenti. Ma questo è falso come si evince facilmente dal fatto che Id_n è una matrice a gradini senza righe nulle (ovvero di rango massimo).

Quindi, se la matrice di A avesse rango minore di n in seguito alla riduzione in contemporanea degli n sistemi (6.16), (6.17), ..., (6.18), comparirebbe necessariamente un'uguaglianza falsa $0 = b \neq 0$. Di conseguenza almeno uno di tali sistemi sarebbe incompatibile e l'uguaglianza $AB = Id_n$ non potrebbe essere verificata.

Questo spiega perché se A non ha rango massimo allora essa sicuramente non è invertibile.

Viceversa, se A ha rango massimo allora riducendo la matrice (6.19) nella parte dei coefficienti non si annulla nessuna riga e tutti i sistemi (6.16), (6.17), ..., (6.18) hanno soluzione. Esiste quindi una matrice B tale che $AB = Id_n$ le cui colonne sono date dalle soluzioni dei sistemi. Si può a questo punto mostrare (ma omettiamo i dettagli) che la matrice B così trovata soddisfa anche l'uguaglianza inversa $BA = Id_n$, ovvero A è invertibile con inversa uguale a B.

Ispirandoci a quanto appena detto, verifichiamo che la matrice $A = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ è invertibile e calcoliamone l'inversa. Come fatto sopra, affianchiamo a tale matrice la matrice identica dello stesso ordine

$$(A|Id_2) = \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right)$$

che rappresenta i due sistemi le cui soluzioni sono le colonne della matrice inversa. Applichiamo il procedimento di riduzione a gradini: in questo caso basta il singolo passaggio

$$(A|Id_n) = \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \underset{R_2 \to R_2 - R_1}{\longrightarrow} \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array}\right) \tag{6.20}$$

Notiamo che, dopo la riduzione a gradini, A non ha righe nulle e quindi, come abbiamo detto sopra, A è invertibile. La prima colonna dell'inversa B di A è data dalla soluzione del sistema ridotto

$$\begin{cases} x_1 - x_2 = 1 \\ 3x_2 = -1 \end{cases} \tag{6.21}$$

Risolvendo dal basso otteniamo la coppia $\left(\frac{2}{3}, -\frac{1}{3}\right)$, che è quindi la prima colonna della matrice inversa. Analogamente, la seconda colonna dell'inversa B di A è data dalla soluzione del sistema ridotto

$$\begin{cases} x_1 - x_2 = 0 \\ 3x_2 = 1 \end{cases} \tag{6.22}$$

Questa, come si vede risolvendo dal basso, è la coppia $\left(\frac{1}{3}, \frac{1}{3}\right)$, che è quindi la seconda colonna della matrice inversa.

In conclusione, l'inversa della matrice A è

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

Infatti, si verifica subito con un calcolo che

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

concordemente con la definizione di inversa.

Per evitare di scrivere e risolvere separatamente i sistemi (6.21) e (6.22) e trovare invece in modo più diretto la matrice inversa, si può procedere come segue. Dopo aver effettuato la riduzione a gradini in (6.20), si applicano ulteriori operazioni elementari fino a trasformare la matrice A del blocco di sinistra nella matrice identica. A questo punto nel blocco di destra si leggerà direttamente l'inversa. Per vederlo, riprendiamo da (6.20) e facciamo comparire prima uno zero in posizione 1 2 eseguendo

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array}\right) \underset{R_1 \to 3R_1 + R_2}{\longrightarrow} \left(\begin{array}{cc|cc} 3 & \mathbf{0} & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right)$$

Applichiamo poi a ogni riga l'operazione elementare del secondo tipo che consiste nel dividerla per l'elemento che si trova sulla diagonale:

$$\left(\begin{array}{cc|cc} 3 & 0 & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right) \underset{\substack{R_1 \to (1/3)R_1 \\ R_2 \to (1/3)R_2}}{\longrightarrow} \left(\begin{array}{cc|cc} 1 & 0 & 2/3 & 1/3 \\ 0 & 1 & -1/3 & 1/3 \end{array}\right). \tag{6.23}$$

Come si vede la matrice identica che avevamo affiancato ad A si è trasformata nella matrice inversa di A già trovata sopra.

Per capire perché, ricordiamo che le operazioni elementari che stiamo eseguendo servono a risolvere contemporaneamente i sistemi che poi ci danno come soluzione le colonne della matrice inversa. Ma allora, riducendo la matrice A (cioè la matrice dei coefficienti di tali sistemi) alla matrice identica come in (6.23) non stiamo facendo altro che ridurre i due sistemi alla forma

$$\begin{cases} x_1 = \frac{2}{3} \\ x_2 = -\frac{1}{3} \end{cases}, \quad \begin{cases} x_1 = \frac{1}{3} \\ x_2 = \frac{1}{3} \end{cases}$$

Abbiamo quindi fatto comparire direttamente nel blocco di destra le soluzioni cercate (che sono proprio le colonne della matrice inversa).

Esempio 6.12. Vogliamo trovare ora l'inversa della matrice

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

Iniziamo con il trasformare la matrice $(A|Id_n)$ in una matrice a gradini

$$(A|Id_n) = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 - 2R_1}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array}\right) \underset{R_3 \to 2R_3 + R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

Il fatto che non si sia annullata nessuna riga nel blocco di sinistra ci dice che la matrice A è invertibile.

Ora, come spiegato sopra, facciamo comparire zeri sopra la diagonale, effettuando una sorta di riduzione a gradini "inversa", dal basso verso l'altro e da destra verso sinistra.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{\substack{R_2 \to 2R_2 + R_3 \\ R_1 \to 2R_1 + R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 2 & \mathbf{0} & -1 & 1 & 2 \\ 0 & 4 & \mathbf{0} & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 2 & 2 & 0 & -1 & 1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{R_1 \to 2R_1 - R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 4 & \mathbf{0} & 0 & -1 & -1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

Infine, dividiamo ogni riga per l'elemento sulla diagonale applicando operazioni elementari del secondo tipo

$$\left(\begin{array}{ccc|ccc} 4 & 0 & 0 & -1 & -1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{\substack{R_1 \to (1/4)R_1 \\ R_1 \to (1/4)R_2 \\ R_3 \to -(1/4)R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/4 & -1/4 & 1/2 \\ 0 & 1 & 0 & -1/4 & 3/4 & 1/2 \\ 0 & 0 & 1 & 3/4 & -1/4 & -1/2 \end{array}\right)$$

Abbiamo quindi ottenuto

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

Come anticipato sopra, diamo una spiegazione del perché le matrici $A \in M_{m,n}(\mathbb{K})$ non quadrate (ovvero con numero di righe m diverso dal numero di colonne n) non possano essere invertibili. Più precisamente, mostriamo che se è verificata la condizione $AB = I$ allora necessariamente $m \leqslant n$ e, analogamente, se è verificata la condizione $BA = I$ allora necessariamente $m \geqslant n$.

Proposizione 6.13. *Sia A una matrice con m righe e n colonne. Allora L_A è iniettiva se e solo se il rango di A è uguale a n. In particolare, se L_A è iniettiva si deve avere $n \leqslant m$.*

Dimostrazione. Consideriamo la funzione

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \qquad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Per definizione, L_A è iniettiva se ogni m-upla $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$ nell'immagine di L_A è l'immagine di un'*unica* n-upla $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, ovvero se e solo se il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{cases}$$

ammette un'unica soluzione per una certa m-upla di arrivo fissata. Tuttavia per la Proposizione 5.17 un sistema compatibile ha soluzione unica se e solo se il rango della matrice dei coefficienti è uguale al numero di incognite. In altre parole il numero di righe non nulle della matrice A dopo la sua riduzione a gradini dovrebbe essere n. Questo equivale a dire che il rango di A è esattamente n.

Per quanto riguarda l'ultima affermazione, si ricordi che il rango per righe è uguale al rango della matrice, si veda l'Osservazione 5.16. Per cui non si può avere $m < n$ in quanto, in tal caso, il rango di A sarebbe necessariamente minore o uguale al numero di righe m e conseguentemente ad n. □

Analogamente possiamo dimostrare la seguente

Proposizione 6.14. *Sia A una matrice con m righe e n colonne: allora L_A è suriettiva se e solo se il rango di A è uguale a m. In particolare, se L_A è suriettiva si deve avere $m \leqslant n$.*

Dimostrazione. Per definizione, L_A è suriettiva se e solo se per ogni m-upla $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{K}^m$ esiste una n-upla $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ tale che $L_A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$, ovvero se e solo se il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{cases}$$

è compatibile per qualunque scelta dei termini noti $y_1, y_2, \cdots, y_m$. Questo accade se e solo se nessuna riga di A si annulla quando viene ridotta a gradini (ovvero se e solo se il rango di A è uguale al numero di righe m). Infatti se nessuna riga si annulla, allora il sistema ammette soluzione qualunque siano i termini noti. Viceversa, se una riga è nulla allora possiamo scegliere i termini noti in modo che il corrispondente termine non si annulli ed ottenere così un sistema incompatibile.
Per quanto riguarda l'ultima affermazione, si ricordi che il rango per colonne è uguale al rango della matrice, si veda l'Osservazione 5.16. Per cui non si può avere $n < m$ in quanto, in tal caso, il rango di A sarebbe necessariamente minore o uguale al numero di colonne n e conseguentemente ad m. □

Siamo quindi pronti a dimostrare il Teorema 6.11.

Dimostrazione del Teorema 6.11. Una matrice $A \in M_{m,n}(\mathbb{K})$ ammette inversa se e solo se L_A è invertibile. Ovvero se e solo se L_A è suriettiva ed iniettiva. Combinando la Proposizione 6.14 e la Proposizione 6.13, otteniamo che questo accade se e solo se il rango di A è uguale sia ad m che ad n. Questo conclude la dimostrazione.

□

Concludiamo questo paragrafo osservando che l'insieme

$$\{A \in M_n(\mathbb{K}) \mid A \text{ è invertibile}\}$$

delle matrici invertibili di ordine n forma un gruppo (Definizione 2.56) rispetto all'operazione di prodotto di matrici.
Infatti osserviamo che il prodotto AB di due matrici A, B invertibili è ancora invertibile perché ha come inversa $B^{-1}A^{-1}$:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AId_nA^{-1} = AA^{-1} = Id_n.$$

Analogamente notiamo che $(B^{-1}A^{-1})(AB) = Id_n$. Questo mostra che il prodotto tra due matrici invertibili è ancora invertibile.
Essendo il prodotto di matrici associativo, la proprietà 1 della Definizione 2.56 è verificata. Inoltre,l'elemento neutro per il prodotto (la matrice identica Id_n) è invertibile (essendo $Id_n \cdot Id_n = Id_n$, essa ha come inversa se stessa). Quindi anche la proprietà 2 della Definizione 2.56 è verificata. Per quello che riguarda la proprietà 3, ci basta mostrare che l'inversa A^{-1} di una matrice A è anch'essa invertibile. Ma questo è chiaro in quanto A stessa è l'inversa di A^{-1}.

Definizione 6.15. Il gruppo $GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid A \text{ è invertibile}\}$ la cui operazione è la moltiplicazione tra matrici è detto **gruppo generale lineare** di ordine n.

Il fatto che $GL_n(\mathbb{K})$ sia un gruppo ci consente di usare i risultati che valgono in generale per qualunque gruppo senza doverli ridimostrare in questo caso specifico. Ad esempio, usando il Lemma 2.63, possiamo affermare che l'inversa del prodotto $A_1A_2 \cdots A_k$ di k matrici invertibili $A_1, A_2, \cdots, A_k$ è data da $A_k^{-1} \cdots A_2^{-1}A_1^{-1}$, senza bisogno di dimostrarlo.

6.3 Invertibilità e determinante di una matrice

In questa sezione introdurremo uno strumento di calcolo alternativo alla riduzione a gradini per determinare se una matrice quadrata A è invertibile e, in tal caso, determinarne l'inversa. Questo è il *determinante*, definito per matrici quadrate.
Come vedremo, il determinante è una funzione che associa un elemento di $\mathbb{K}$, denotato con $\det(A)$, ad ogni matrice A quadrata di ordine n con entrate appartenenti un campo $\mathbb{K}$. Il determinante rileva l'invertibilità di una matrice nel senso che A è invertibile se e solo se $\det(A) \neq 0$. Dal momento che, come visto nel paragrafo precedente, una matrice è invertibile esattamente quando le sue righe sono indipendenti, questo equivale a dire che $\det(A) = 0$ se e solo se le righe della matrice A sono dipendenti.

Prima di dare la definizione generale, mostriamo che nel caso di matrici quadrate di ordine 2 e di ordine 3 non è difficile trovare una funzione delle entrate che sia nulla se e solo se le righe della matrice sono dipendenti.

Infatti, data una generica matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

di ordine 2, essendoci solo due righe queste saranno dipendenti se e solo se sono proporzionali. Questo, supponendo per il momento che a_{21} e a_{22} siano diversi da zero, significa che

$$\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} \tag{6.24}$$

ovvero, moltiplicando l'equazione per a_{21} (così da semplificarlo al primo membro) e per a_{22} (così da semplificarlo al secondo),

$$a_{11}a_{22} = a_{12}a_{21}.$$

Portando tutto al primo membro, si ottiene

$$a_{11}a_{22} - a_{12}a_{21} = 0 \tag{6.25}$$

che ci dà, come volevamo, una funzione delle entrate della matrice che si annulla esattamente quando le righe sono dipendenti.

Si noti che la (6.25) rileva la dipendenza delle righe anche senza l'ipotesi, che era necessaria per poter scrivere la (6.24), che a_{21} e a_{22} fossero diversi da zero. Infatti se, ad esempio, $a_{21} = 0$ allora la (6.25) diventa $a_{11}a_{22} = 0$, che equivale a dire che o $a_{11} = 0$ o $a_{22} = 0$. Nel primo caso la matrice è della forma $A = \left(\begin{smallmatrix} 0 & a_{12} \\ 0 & a_{22} \end{smallmatrix}\right)$, nel secondo $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ 0 & 0 \end{smallmatrix}\right)$, e in entrambi i casi le righe sono proporzionali.

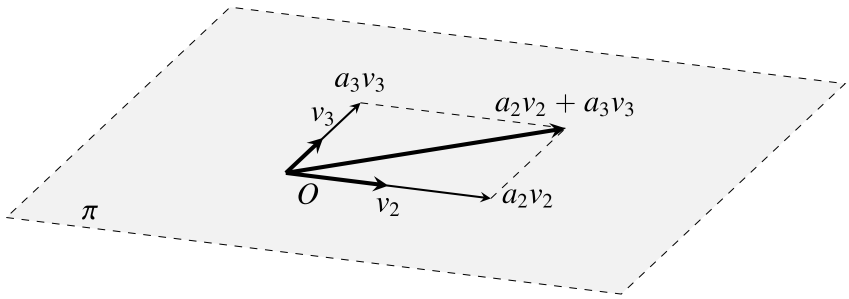
Anche per matrici di ordine 3 non è difficile ottenere, aiutandosi con un'interpretazione geometrica, una funzione delle entrate che si annulli se e solo se le righe sono dipendenti. Più precisamente, consideriamo

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

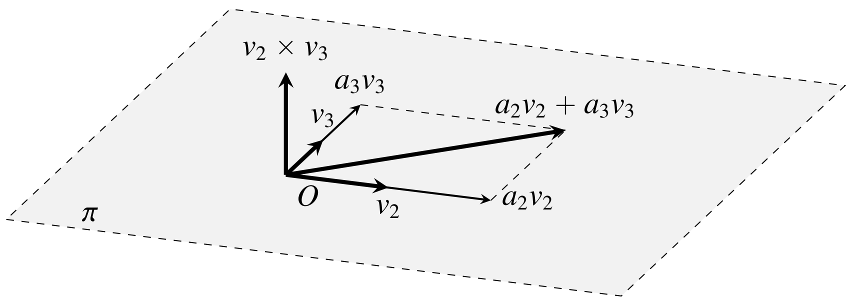
Possiamo pensare che le righe R_1, R_2, R_3 della matrice rappresentino le terne delle coordinate, rispetto a un sistema di riferimento ortonormale, di tre vettori geometrici v_1, v_2, v_3 applicati nello spazio.

Per cui esiste una relazione di dipendenza tra R_1, R_2, R_3 se e solo se la stessa relazione di dipendenza vale tra i vettori corrispondenti (ad esempio, $R_3 = R_1 + R_2$ se e solo se

$v_3 = v_1 + v_2$, oppure $R_1 = 2R_2 - R_3$ se e solo se $v_1 = 2v_2 - v_3$ e così via).
Ora, se esistesse una relazione di dipendenza tra i vettori questo vorrebbe dire che essi stanno sullo stesso piano. Ad esempio, se avessimo $v_1 = a_2v_2 + a_3v_3$ allora v_1 sarebbe sullo stesso piano su cui stanno v_2 e v_3 in quanto una qualunque combinazione di v_2 e v_3 appartiene a tale piano



Ma appartenere al piano su cui stanno v_2 e v_3 significa essere perpendicolare al prodotto vettoriale $v_2 \wedge v_3$ che è infatti perpendicolare a v_2 e v_3 e quindi a qualunque altro vettore si trovi su quel piano.



Traduciamo allora tale condizione in coordinate. Assumendo, come abbiamo detto sopra, che le coordinate di v_1, v_2, v_3 siano date rispettivamente dalla prima, seconda e terza riga della matrice A, ovvero $v_1 \equiv (a_{11}, a_{12}, a_{13})$, $v_2 \equiv (a_{21}, a_{22}, a_{23})$, $v_3 \equiv (a_{31}, a_{32}, a_{33})$, abbiamo

$$v_2 \wedge v_3 \equiv (a_{22}a_{33} - a_{23}a_{32}, a_{23}a_{31} - a_{21}a_{33}, a_{21}a_{32} - a_{22}a_{31}).$$

Ricordando che due vettori sono perpendicolari se e solo se il loro prodotto vettoriale è nullo, si ha che v_1 è perpendicolare a $v_2 \wedge v_3$ (ovvero sta sullo stesso piano su cui stanno v_2 e v_3) se e solo se

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = 0 \qquad (6.26)$$

ovvero, svolgendo i calcoli,

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} = 0. \qquad (6.27)$$

Come volevamo, abbiamo ottenuto anche nel caso di matrici di ordine 3 una funzione delle entrate che si annulla se e solo se le sue righe sono dipendenti.
Ora, notiamo che i primi membri sia dell'espressione (6.25) che della (6.27) sono somma (con segno $+$ o $-$) di addendi del tipo

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \tag{6.28}$$

dove $\sigma(1), \sigma(2), \cdots, \sigma(n)$ è una permutazione degli indici di colonna. Ad esempio, nella (6.27) abbiamo che

> l'addendo $a_{11}a_{22}a_{33}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 1,\ \sigma(2) = 2,\ \sigma(3) = 3$
>
> l'addendo $a_{11}a_{23}a_{32}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 1,\ \sigma(2) = 3,\ \sigma(3) = 2$
>
> l'addendo $a_{12}a_{23}a_{31}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1$
>
> l'addendo $a_{12}a_{21}a_{33}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 2,\ \sigma(2) = 1,\ \sigma(3) = 3$
>
> l'addendo $a_{13}a_{21}a_{32}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 3,\ \sigma(2) = 1,\ \sigma(3) = 2$
>
> l'addendo $a_{13}a_{22}a_{31}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1) = 3,\ \sigma(2) = 2,\ \sigma(3) = 1$

Analogamente, nella (6.25) abbiamo che

> l'addendo $a_{11}a_{22}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}$ con σ data dalla permutazione $\sigma(1) = 1,\ \sigma(2) = 2$
>
> l'addendo $a_{12}a_{21}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}$ con σ data dalla permutazione $\sigma(1) = 2,\ \sigma(2) = 1$

In entrambi i casi, si vede che abbiamo un addendo per ognuna delle possibili permutazioni degli indici di colonna (nel caso di ordine 2 abbiamo due sole permutazioni dell'insieme $\{1, 2\}$, nel caso di ordine 3 abbiamo 6 permutazioni dell'insieme $\{1, 2, 3\}$). Per quanto riguarda il segno di tali addendi, possiamo notare che sia nella (6.25) che nella (6.27) gli addendi che corrispondono a permutazioni pari (si veda da Definizione 2.62) hanno segno positivo, mentre quelli dati da permutazioni dispari hanno segno negativo. Ad esempio, l'addendo $a_{11}a_{22}$ ha segno positivo perché corrisponde alla permutazione identica, che è pari; mentre l'addendo $a_{12}a_{21}$ ha segno - perché corrisponde alla trasposizione che scambia 1 e 2, che è una permutazione dispari. Analogamente, nella (6.27) l'addendo $a_{11}a_{22}a_{33}$ ha segno positivo perché corrisponde alla permutazione identica, l'addendo $a_{11}a_{23}a_{32}$ ha segno negativo perché corrisponde alla trasposizione che scambia 2 e 3, l'addendo $a_{12}a_{23}a_{31}$ ha segno positivo perché corrisponde al ciclo di lunghezza tre $(1\ 2\ 3)$, che si può ottenere tramite la composizione $(1\ 3) \circ (1\ 2)$ di due trasposizioni, e quindi è pari, e così via.

Definizione 6.16. Il **segno** $s(\sigma)$ di una permutazione $\sigma \in S_n$ è definito da

$$s(\sigma) = \begin{cases} 1 & \text{se } \sigma \text{ è pari} \\ -1 & \text{se } \sigma \text{ è dispari} \end{cases}$$

Con questa definizione sia la (6.25) che la (6.27) possono essere scritte come

$$\sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

rispettivamente per $n = 2$ e $n = 3$.

Quello che si può dimostrare è che, anche per matrici di ordine $n > 3$ si ha che tale espressione si annulla se e solo se le righe della matrice sono dipendenti. Possiamo allora dare la seguente

Definizione 6.17. Data una matrice quadrata $A \in M_n(\mathbb{K})$, il suo determinante $\det(A)$ è l'elemento di $\mathbb{K}$ dato da

$$\sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)} \tag{6.29}$$

Notiamo che la (6.29) ha tanti addendi quante sono le possibili permutazioni degli n indici di colonna che, come sappiamo dalla Sezione 2.9, sono $n!$. Per le matrici di ordine 4, si ha $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, per quelle di ordine 5 si ha $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$, quindi vediamo che il numero di addendi nell'espressione del determinante cresce molto rapidamente, il che rende l'uso di tale definizione non molto pratico per il calcolo.

Per questo motivo, introduciamo ora una formula che ci consente di calcolare il determinante senza dover utilizzare la (6.29).

Definizione 6.18. Data una matrice quadrata $A \in M_n(\mathbb{K})$ di ordine $n \geqslant 2$ si dice **cofattore** (o **complemento algebrico**) di una sua entrata a_{ij}, l'elemento $C_{ij} \in \mathbb{K}$ dato dal determinante della matrice che si ottiene da A cancellando la i-esima riga e la j-esima colonna, moltiplicato per $(-1)^{i+j}$.

Esempio 6.19. Calcoliamo i cofattori delle entrate della matrice $A \in M_3(\mathbb{R})$ data da

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

I cofattori sono

$$C_{11} = (-1)^{1+1} \det \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} = -2, \;\; C_{12} = (-1)^{1+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = -1$$

$$C_{13} = (-1)^{1+3} \det \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = 1, \;\; C_{21} = (-1)^{2+1} \det \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = -4$$

$$C_{22} = (-1)^{2+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1, \;\; C_{23} = (-1)^{2+3} \det \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = 2$$

$$C_{31} = (-1)^{3+1} \det \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = 3, \;\; C_{32} = (-1)^{3+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0$$

$$C_{33} = (-1)^{3+3} \det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = -3.$$

Definizione 6.20. Sia A una matrice quadrata di ordine $n \geqslant 2$ e siano $\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ la sua i-esima riga e $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ la sua j-esima colonna.

Si dice **sviluppo di Laplace rispetto alla *i*-esima riga** l'elemento di $\mathbb{K}$ definito da

$$a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}. \tag{6.30}$$

Analogamente si dice **sviluppo di Laplace rispetto alla *j*-esima colonna** l'elemento di $\mathbb{K}$ definito da

$$a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}. \tag{6.31}$$

Di seguito enunciamo un fondamentale risultato senza dimostrarlo.

Proposizione 6.21. *Sia A una matrice quadrata di ordine $n \geqslant 2$. Allora lo sviluppo di Laplace rispetto ad una qualunque sua riga o colonna è uguale al determinante di A. In formule*

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

indipendentemente dalla riga o dalla colonna scelta.

Ad esempio, nel caso di una matrice di ordine 3, scegliendo la prima riga di

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

si ha

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = \\ &= a_{11}(-1)^{1+1}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} + \\ &+ a_{12}(-1)^{1+2}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + \\ &+ a_{13}(-1)^{1+3}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

che coincide proprio con la (6.27).

Osserviamo che lo sviluppo di Laplace è una formula *ricorsiva*. Infatti per calcolare il determinante di una matrice A di ordine n prevede che si conoscano i cofattori, che sono determinanti di matrici di ordine $n-1$ (quelle che si ottengono cancellando una riga e una colonna da A). A loro volta, ciascuno di questi determinanti va calcolato con la stessa formula, riducendoli al calcolo di determinanti di matrici di ordine $n-2$. Si continua così fino ad arrivare a determinanti di matrici di ordine 2, per i quali si ha la semplice formula $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

Esempio 6.22. Si consideri la matrice

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

Scegliamo di sviluppare il determinante rispetto alla terza riga, usando cioè la formula (6.30) nel caso $i = 3$:

$$\det(A) = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} + a_{34}C_{34}.$$

Tenendo conto che $a_{32} = a_{33} = 0$ (quindi senza calcolare i corrispondenti cofattori) si ha

$$\det(A) = 1 \cdot (-1)^{3+1} \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{pmatrix} + 1 \cdot (-1)^{3+4} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix}. \tag{6.32}$$

Ora, ciascuno dei due determinanti di ordine 3 che compaiono in questa uguaglianza va di nuovo calcolato usando la formula data. Ad esempio, se per il primo determinante scegliamo la terza colonna (e usiamo quindi la formula $\det(A) = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33}$) tenendo conto che $a_{13} = a_{33} = 0$, si ha

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{pmatrix} = 1 \cdot (-1)^{2+3} \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = -1$$

dove il determinante di ordine 2 è stato calcolato usando la formula $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. Analogamente, se per l'altro determinante scegliamo la prima riga (quindi usiamo $\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$) tenendo conto che $a_{13} = 0$, si ha

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix} = 1 \cdot (-1)^{1+1} \det \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} + 1 \cdot (-1)^{1+2} \det \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = 3 - 1 = 2.$$

Sostituendo questi risultati nella (6.32), si trova allora

$$\det(A) = 1 \cdot (-1)^{3+1}(-1) + 1 \cdot (-1)^{3+4}\, 2 = -3.$$

Ora, le nozioni di determinante e cofattore appena introdotte, oltre a dirci se una matrice è invertibile, ci danno un modo alternativo al procedimento descritto nel paragrafo precedente per calcolarne l'inversa. Infatti, si ha il seguente risultato, che non dimostriamo.

Proposizione 6.23. *Sia $A \in M_n(\mathbb{K})$ una matrice invertibile (ovvero con $\det(A) \neq 0$). Allora la sua inversa A^{-1} è data da*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix} \tag{6.33}$$

dove C_{ij} indica, come sopra, il cofattore di a_{ij}, e $\frac{1}{\det(A)}$ davanti alla matrice dei cofattori significa che ogni entrata di tale matrice deve essere moltiplicata per $\frac{1}{\det(A)}$.

Osservazione 6.24. A proposito della disposizione dei cofattori nella (6.33), si noti che i cofattori delle entrate della prima *riga* di A sono nella prima *colonna* della (6.33), i cofattori delle entrate della seconda *riga* di A sono nella seconda *colonna* della (6.33), e così via. In altri termi le righe sono scambiate con le colonne e viceversa.

Esempio 6.25. Calcoliamo l'inversa della matrice

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

di cui abbiamo calcolato i cofattori nell'Esempio 6.19. Da quel calcolo risulta

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{2n} & C_{33} \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -1 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{4}{3} & -1 \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{2}{3} & 1 \end{pmatrix}.$$

Il determinate di A è stato calcolato sviluppandolo secondo Laplace rispetto alla terza riga, usando i cofattori già calcolati:

$$\det(A) = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} = 1 \cdot 3 + 0 \cdot 0 + 2 \cdot (-3) = -3.$$

Osservazione 6.26. Si noti che la (6.33), nel caso $n = 2$, diventa la semplice formula

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \tag{6.34}$$

in quanto i cofattori sono semplicemente $C_{11} = (-1)^{1+1}a_{22}$, $C_{12} = (-1)^{1+2}a_{21}$, $C_{21} = (-1)^{2+1}a_{12}$, $C_{22} = (-1)^{2+2}a_{11}$. In pratica, a parte dividere per il determinante, la matrice inversa si ottiene scambiando tra loro i due elementi sulla diagonale e cambiando di segno le restanti entrate.

Essendo il determinante uno strumento fondamentale per la determinazione dell'invertibilità di una matrice, il calcolo della sua inversa e per molte altre apllicazioni, è fondamentale conoscerne le proprietà, anche allo scopo di agevolarne il calcolo.
Iniziamo con il chiederci come si comporta il determinante di una matrice rispetto alle operazioni elementari.

(1) Se A' si ottiene da una matrice $A \in M_n(\mathbb{K})$ moltiplicando una riga di A per $c \in \mathbb{K}$, allora $\det(A') = c\det(A)$.

Dimostrazione. In base alla Definizione 6.17 il determinante della matrice A' è

$$\det(A') = \sum s(\sigma)a_{1\sigma(1)} \cdots (ca_{i\sigma(i)}) \cdots a_{n\sigma(n)}.$$

Il fattore c è comune a tutti gli addendi della sommatoria, possiamo quindi metterlo in evidenza davanti alla somma e scrivere

$$\det(A') = c \sum s(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)},$$

ovvero $\det(A') = c \det(A)$, come volevamo. $\square$

Esempio 6.27. Sia $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e sia $A' = \left(\begin{smallmatrix} 5 & 10 \\ 3 & 4 \end{smallmatrix}\right)$ è la matrice ottenuta da A moltiplicando la prima riga per $c = 5$. Allora si ha

$$\begin{aligned} \det(A) &= 1 \cdot 4 - 3 \cdot 2 = -2, \\ \det(A') &= 5 \cdot 4 - 10 \cdot 3 = -10. \end{aligned}$$

(2) Se una riga R_i della matrice $A \in M_n(\mathbb{K})$ si scrive come somma $R_i = R_i' + R_i''$, si ha

$$\det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} = \det \begin{pmatrix} R_1' \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ R_i'' \\ \vdots \\ R_n \end{pmatrix}. \tag{6.35}$$

Dimostrazione. Siano

$$\begin{aligned} R_i' &= \begin{pmatrix} a_{i1}' & a_{i2}' & \cdots & a_{in}' \end{pmatrix} \\ R_i'' &= \begin{pmatrix} a_{i1}'' & a_{i2}'' & \cdots & a_{in}'' \end{pmatrix} \end{aligned}$$

e di conseguenza

$$R_i = R_i' + R_i'' = \begin{pmatrix} a_{i1}' + a_{i1}'' & a_{i2}' + a_{i2}'' & \cdots & a_{in}' + a_{in}'' \end{pmatrix}.$$

Allora, in base alla Definizione 6.17, si ha

$$\det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} = \sum s(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i-1\sigma(i-1)} \left(a_{i\sigma(i)}' + a_{i\sigma(i)}'' \right) a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} =$$

$$= \sum s(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)}' \cdots a_{n\sigma(n)} + \sum s(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)}'' \cdots a_{n\sigma(n)} =$$

$$= \det \begin{pmatrix} R_1 \\ \vdots \\ R_i' \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ R_i'' \\ \vdots \\ R_n \end{pmatrix}$$

dove abbiamo usato la proprietà distributiva in ogni termine della sommatoria. □

Esempio 6.28. Se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, la seconda riga si scrive come somma $\begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 3 \end{pmatrix}$ e, in accordo con la (6.35), si vede che

$$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

come può essere verificato col calcolo esplicito (si ottiene $-2 = -3 + 1$).

(3) Se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ scambiando tra loro due righe, o due colonne, allora $\det(A') = -\det(A)$.

Dimostrazione. Sia A' la matrice ottenuta da A scambiando la riga R_i con la riga R_j: per ogni k si ha $a'_{ik} = a_{jk}$ e $a'_{jk} = a_{ik}$, mentre tutte le altre entrate sono uguali a quelle di A.

In base alla Definizione 6.17, si ha

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} s(\sigma)\, a'_{1\sigma(1)} \cdots a'_{i\sigma(i)} \cdots a'_{j\sigma(j)} \cdots a'_{n\sigma(n)} = \\ &= \sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(i)} \cdots a_{i\sigma(j)} \cdots a_{n\sigma(n)}. \end{aligned} \tag{6.36}$$

Confrontando la (6.36) con l'espressione

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} \tag{6.37}$$

vediamo che in ogni addendo le entrate delle colonne $\sigma(i)$ e $\sigma(j)$ hanno gli indici di riga i e j scambiati tra loro.

Ora, allo scopo di dimostrare il risultato, riscriviamo la (6.36) come segue. Consideriamo la trasposizione $\tau \in S_n$ che scambia i con j e lascia invariati tutti gli altri elementi, e definiamo la funzione $f : S_n \to S_n$ che associa a ogni permutazione σ la composizione $\sigma \circ \tau$.

Tale funzione è invertibile e ha come inversa se stessa: $f(f(\sigma)) = f(\sigma \circ \tau) = \sigma \circ \tau \circ \tau = \sigma$ (dove abbiamo usato il fatto che, per ogni trasposizione, $\tau \circ \tau = id$).

Quindi f è biiettiva; il che significa che, al variare di $\sigma \in S_n$, $\sigma \circ \tau$ ci dà tutte le permutazioni di S_n.

Allora possiamo sostituire il σ che compare in ognuno degli addendi di (6.36) con $\sigma \circ \tau$ e ottenere comunque, al variare di $\sigma \in S_n$, tutte le permutazioni di S_n in modo che la somma sia invariata. Si ha quindi

$$\det(A') = \sum_{\sigma \in S_n} s(\sigma \circ \tau)\, a_{1(\sigma\circ\tau)(1)} \cdots a_{j(\sigma\circ\tau)(i)} \cdots a_{i(\sigma\circ\tau)(j)} \cdots a_{n(\sigma\circ\tau)(n)}.$$

Dal momento che τ scambia i e j lasciando invariati tutti gli altri elementi di $\{1, 2, \cdots, n\}$, si ha

$$(\sigma \circ \tau)(k) = \begin{cases} k & \text{se } k \neq i, j \\ j & \text{se } k = i \\ i & \text{se } k = j \end{cases}$$

e quindi otteniamo

$$\det(A') = \sum_{\sigma \in S_n} s(\sigma \circ \tau)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)}. \tag{6.38}$$

Notiamo che la (6.38) è uguale a quella (6.37) del determinante di A a parte che in essa compare $s(\sigma \circ \tau)$ invece che $s(\sigma)$. Ora, in base alla definizione di segno, si ha $s(\sigma \circ \tau) = -s(\sigma)$. Infatti, se $s(\sigma) = +1$ allora questo significa che σ si scrive come prodotto di un numero pari di trasposizioni, e quindi $\sigma \circ \tau$, contenendo in più la trasposizione τ, risulterà prodotto di un numero dispari di trasposizioni, da cui $s(\sigma \circ \tau) = -1$. Analogamente, se $s(\sigma) = -1$, allora σ è una permutazione dispari e, conseguentemente, $\sigma \circ \tau$ è una permutazione pari, da cui $s(\sigma \circ \tau) = +1$. Quindi la (6.38) diventa

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} -s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = \\ &= -\sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = -\det(A) \end{aligned}$$

come volevamo.

□

Esempio 6.29. Sia

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

e sia

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

la matrice ottenuta da A scambiando tra loro la prima e la seconda riga. Allora si ha

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2,$$
$$\det(A') = 3 \cdot 2 - 4 \cdot 1 = +2.$$

Le proprietà (1) e (2) appena viste ci consentono di capire come si comporta il determinante quando eseguiamo operazioni elementari del terzo tipo. Più precisamente, supponiamo che la matrice A' sia ottenuta da $A \in M_n(\mathbb{K})$ sommando alla sua i-esima riga R_i la j-esima riga R_j moltiplicata per c, ovvero

$$A' = \begin{pmatrix} R_1 \\ \vdots \\ R_i + cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix}.$$

Ma allora, per la proprietà (2) abbiamo

$$\det(A') = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i + cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix}. \tag{6.39}$$

Mentre il primo addendo è il determinante della matrice A, il secondo addendo è nullo, in quanto determinante di una matrice con le righe dipendenti (compaiono R_j e cR_j che sono proporzionali).[3]

[3] Se non si vuole utilizzare questo risultato che non abbiamo dimostratato basta usare prima la proprietà (1) ed applicare poi la proprietà (3) alle righe R_j ed R_j che si trovano in i-esima e j-esima posizione.

Quindi la (6.39) si legge

$$\det(A') = \det(A) + 0 = \det(A).$$

Concludiamo che *quando eseguiamo un'operazione elementare del terzo tipo sulle righe di una matrice, il determinante non cambia*.
Questo risultato può essere usato per facilitare il calcolo del determinante di una matrice, in quanto mediante operazioni elementari del terzo tipo si può modificare la matrice trasformando alcune entrate in 0 senza cambiare il determinante. Si può poi calcolare lo sviluppo di Laplace rispetto ad una riga o colonna che presenta il numero maggiore di questi zeri così comparsi (ricordiamo che se un'entrata è nulla non è necessario calcolare il corrispondente cofattore).
Mediante tali operazioni potremmo anche trasformare la matrice in una matrice a gradini

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \tag{6.40}$$

Il determinante di tale matrice è dato semplicemente dal prodotto $a_{11}a_{22}\cdots a_{nn}$ degli elementi della sua diagonale. Infatti sviluppando secondo Laplace rispetto alla prima colonna, essendo a_{11} l'unico elemento non nullo, si avrà

$$\det(A) = a_{11}(-1)^{1+1} \det \begin{pmatrix} a_{22} & \cdots & a_{2n-1} & a_{2n} \\ 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Se ora per lo sviluppo di Laplace del cofattore scegliamo nuovamente la prima colonna, che ha come unico elemento non nullo a_{22}, otteniamo

$$\det(A) = a_{11}a_{22} \det \begin{pmatrix} a_{33} & \cdots & a_{3n-1} & a_{3n} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

e continuando a sviluppare sempre rispetto alla prima colonna otteniamo proprio $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Osservazione 6.30. Tutte le proprietà del determinante che abbiamo visto valide per operazioni sulle righe sono vere anche rispetto alle colonne. Per cui valgono le seguenti proprietà che si dimostrano in modo del tutto analogo.

(1) Se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ moltiplicando una colonna di A per $c \in \mathbb{K}$, allora $\det(A') = c \det(A)$.

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e $A' = \left(\begin{smallmatrix} 5 & 2 \\ 15 & 4 \end{smallmatrix}\right)$ è la matrice ottenuta da A moltiplicando la prima colonna per $c = 5$, allora si ha

$$\det(A) = -2 \ \text{ e } \ \det(A') = 5 \cdot 4 - 10 \cdot 3 = -10.$$

(2) Se una colonna C_i della matrice $A \in M_n(\mathbb{K})$ si scrive come somma $C_i = C_i' + C_i''$, si ha

$$\begin{aligned} &\det \begin{pmatrix} C_1 & \cdots & C_i' + C_i'' & \cdots & C_n \end{pmatrix} = \\ &= \det \begin{pmatrix} C_1 & \cdots & C_i' & \cdots & C_n \end{pmatrix} + \det \begin{pmatrix} C_1 & \cdots & C_i'' & \cdots & C_n \end{pmatrix}. \end{aligned}$$

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, la seconda colonna si scrive come somma

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

e si ha

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

come può essere verificato calcolando i determinanti (si ottiene $-2 = -5 + 3$).

(3) se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ scambiando tra loro due colonne, allora $\det(A') = -\det(A)$.

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e $A' = \left(\begin{smallmatrix} 2 & 1 \\ 4 & 3 \end{smallmatrix}\right)$ è la matrice ottenuta da A scambiando tra loro la prima e la seconda colonna, si ha

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2 \ \text{ e } \ \det(A') = 2 \cdot 3 - 1 \cdot 4 = +2.$$

Come ulteriore proprietà del determinante, citiamo senza dimostrarlo il seguente

Teorema 6.31 (Teorema di Binet)**.** *Siano $A, B \in M_n(\mathbb{K})$ due matrici quadrate. Allora*

$$\det(AB) = \det(A)\det(B).$$

Ad esempio, sia $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$. Allora si ha $AB = \begin{pmatrix} 4 & 7 \\ 10 & 15 \end{pmatrix}$, e si vede che $\det(A) = -2$, $\det(B) = 5$ e $\det(AB) = -10 = (-2) \cdot 5$, come afferma il teorema di Binet.

Dal teorema di Binet discende immediatamente il seguente

Corollario 6.32. *Se $A \in M_n(\mathbb{K})$ è una matrice invertibile, allora*

$$\det(A^{-1}) = \frac{1}{\det(A)}. \tag{6.41}$$

Dimostrazione. Basta applicare il teorema di Binet al caso $B = A^{-1}$. In questo caso infatti abbiamo

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

ovvero, tenuto conto che $AA^{-1} = Id_n$,

$$\det(Id_n) = \det(A)\det(A^{-1}).$$

Ma la matrice identica Id_n è una matrice del tipo (6.40), con $a_{11}, a_{22}, \cdots, a_{nn}$ tutti uguali a 1, quindi il suo determinante è 1. In definitiva, si ottiene

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

che implica subito la (6.41). □

Vediamo ora un'applicazione del determinante e più precisamente della formula (6.33) per l'inversa che ci fornisce un modo alternativo alla riduzione a gradini per risolvere i sistemi di equazioni lineari sotto certe ipotesi.

A tale scopo, osserviamo prima che un generico sistema lineare di m equazioni in n incognite

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{6.42}$$

può essere riscritto in forma molto più concisa grazie al prodotto di matrici. Più precisamente, denotiamo con

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

la matrice dei coefficienti del sistema, con

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

la matrice (costituita da una sola colonna) che ha come entrate le incognite, e con

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

la matrice (sempre costituita da una sola colonna) che ha come entrate i termini noti del sistema. Allora svolgendo il prodotto righe per colonne, notiamo che

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Quindi il sistema (6.42) può essere riscritto come prodotto di matrici nella forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

o nella forma più compatta

$$Ax = b. \tag{6.43}$$

Ora, supponiamo che la matrice A dei coefficienti del sistema sia quadrata di ordine n (quindi il sistema ha n equazioni e n incognite) e che abbia determinante diverso da zero. Allora A è invertibile e possiamo moltiplicare a sinistra entrambi i membri dell'uguaglianza (6.43) per l'inversa A^{-1} ottenendo

$$A^{-1}(Ax) = A^{-1}b.$$

Tenendo conto che il prodotto di matrici gode della proprietà associativa, questo può essere riscritto come

$$(A^{-1}A)x = A^{-1}b$$

ovvero, visto che per definizione di inversa $A^{-1}A = Id_n$,

$$Id_n x = A^{-1}b.$$

Per cui, essendo Id_n elemento neutro per il prodotto, otteniamo

$$x = A^{-1}b. \tag{6.44}$$

Quindi la formula (6.44) fornisce la soluzione x di un sistema con lo stesso numero di equazioni e di incognite e matrice dei coefficienti con determinante diverso da zero.

Osservazione 6.33. Si noti che non abbiamo fatto altro che applicare gli stessi passaggi, normalmente sottointesi, che si applicano quando si vuole risolvere una semplice equazione di primo grado in una sola incognita $ax = b$. Se, ad esempio, vogliamo risolvere l'equazione $2x = 3$, dividiamo entrambi i membri per 2 o, equivalentemente, moltiplichiamo per l'inverso $\frac{1}{2}$ di 2 ottenendo $\frac{1}{2}(2x) = \frac{1}{2}3 = \frac{3}{2}$. Per la proprietà associativa al primo membro si ha $(\frac{1}{2}2)x = \frac{3}{2}$ ovvero, essendo $\frac{1}{2}2 = 1$, si ha $1x = \frac{3}{2}$. Per cui, essendo 1 elemento neutro per la moltiplicazione tra reali, $x = \frac{3}{2}$, che è la soluzione dell'equazione.
L'unica differenza con il caso dei sistemi $Ax = b$ è che non sempre A è invertibile, mentre, a meno che a non sia zero, nell'equazione $ax = b$ tale ipotesi è sempre garantita.

Se combiniamo la (6.44) con la formula per l'inversa (6.33), vediamo che la soluzione di un sistema con n equazioni e n incognite e matrice dei coefficienti A invertibile è data da

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{6.45}$$

Quindi le componenti x_i della soluzione del sistema si ottengono moltiplicando le righe della matrice dei cofattori (divisa per il determinante di A) per la colonna dei termini noti:

$$x_1 = \frac{b_1C_{11} + b_2C_{21} + \cdots + b_nC_{n1}}{\det(A)}$$

$$x_2 = \frac{b_1C_{12} + b_2C_{22} + \cdots + b_nC_{n2}}{\det(A)}$$

$$\vdots$$

In generale

$$x_i = \frac{b_1C_{1i} + b_2C_{2i} + \cdots + b_nC_{ni}}{\det(A)} \tag{6.46}$$

Confrontando il numeratore della (6.46) con l'espressione del determinante di A calcolato con lo sviluppo di Laplace rispetto alla i-esima colonna

$$\det(A) = a_{1i}C_{1i} + a_{2i}C_{2i} + \cdots + a_{ni}C_{ni}$$

si nota che i termini noti $b_1, b_2, \cdots, b_n$ compaiono al posto degli elementi $a_{1i}, a_{2i}, \cdots, a_{ni}$ della colonna i-esima di A. In altre parole, tale numeratore coincide con lo sviluppo di Laplace del determinante della matrice che si ottiene da A ponendo i termini noti al posto della i-esima colonna. Riassumendo, abbiamo dimostrato il

Teorema 6.34 (Teorema di Cramer)**.** *Sia $Ax = b$ un sistema di n equazioni lineari in n incognite, con* $\det(A) \neq 0$*. Allora tale sistema ha un'unica soluzione $(x_1, x_2, \cdots, x_n)$ le cui componenti sono date da*

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

dove B_i è la matrice che si ottiene da A ponendo la matrice b dei termini noti al posto della i-esima colonna di A.

Esempio 6.35. Consideriamo il sistema

$$\begin{cases} 2x_1 + x_2 = 5 \\ x_1 - x_2 = 3 \end{cases}$$

La matrice dei coefficienti del sistema

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

ha determinante $\det(A) = -3$, quindi è invertibile e possiamo applicare il metodo di Cramer. Le componenti dell'unica soluzione sono date da

$$x_1 = \frac{\det(B_1)}{\det(A)} = \frac{\det\begin{pmatrix} 5 & 1 \\ 3 & -1 \end{pmatrix}}{\det(A)} = \frac{-8}{-3} = \frac{8}{3}$$

$$x_2 = \frac{\det(B_2)}{\det(A)} = \frac{\det\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}}{\det(A)} = \frac{1}{-3} = -\frac{1}{3}$$

Osservazione 6.36. Quanto visto finora ci permette di tracciare un parallelo tra il comportamento delle soluzioni di un sistema $Ax = b$ di n equazioni lineari in n incognite e

il comportamento delle soluzioni di un'equazione congruenziale $ax \equiv_n b$, descritto dal Teorema 3.29 e dal Teorema 3.31.
Infatti, in tali risultati abbiamo visto che $ax \equiv_n b$ ha soluzioni se e solo se il massimo comune divisore d di a e n divide b. In tal caso le soluzioni *distinte modulo* n sono esattamente d. Da ciò si evince che nel caso in cui il massimo comune divisore d di a e n sia 1, la soluzione esiste ed è unica qualunque sia b (1 divide qualunque b). Ma $d = 1$ significa che a e n sono primi tra loro, il che a sua volta equivale a dire, come visto a pagina 113, che a è invertibile modulo n. Possiamo allora riformulare quanto appena detto come segue: se a è invertibile modulo n, l'equazione $ax \equiv_n b$ ha una sola soluzione qualunque sia b, mentre se a non è invertibile (ovvero il massimo comune divisore d di a e n è maggiore di 1) allora la soluzione non esiste per certi valori di b (quelli per cui d non divide b), e per i valori per cui esiste sicuramente non è unica (come abbiamo detto, ci sono d soluzioni).
Un sistema lineare $Ax = b$ di n equazioni in n incognite si comporta esattamente allo stesso modo. Se la matrice A è invertibile, allora la soluzione esiste qualunque sia b ed è unica, data come abbiamo visto sopra da $x = A^{-1}b$. Alternativamente, se A è invertibile il suo rango r è uguale a n e quindi dopo la riduzione a gradini non si annulla nessuna riga; quindi non non ci sono incompatibilità del tipo $0 = c$ con $c \neq 0$ e la soluzione esiste qualunque sia la colonna dei termini noti b. Inoltre questa è unica, come afferma la Proposizione 5.17, in quanto il rango r è uguale al numero n di incognite.
Invece, se A non è invertibile, il suo rango r è minore di n e quindi dopo la sua riduzione a gradini si annulla almeno una riga. Questo comporta che per certe scelte della colonna dei termini noti b si otterrà in corrispondenza della riga annullata un'incompatibilità del tipo $0 = c$ con $c \neq 0$, ovvero per certi b non avremo soluzione. D'altra parte, per i b per cui abbiamo soluzione, questa non sarà certamente unica, in quanto come sappiamo le soluzioni di un sistema compatibile sono ∞^{n-r} (sempre per la Proposizione 5.17), ed essendo $r < n$ per ipotesi si avrà $n - r > 0$ e quindi le soluzioni saranno infinite.

Oltre al prodotto, tra matrici può essere definita anche un'operazione di somma.
L'introduzione di tale operazione si rivela necessaria se vogliamo manipolare espressioni e uguaglianze tra matrici esattamente come manipoliamo espressioni e uguaglianze tra numeri. Ad esempio, come abbiamo già messo in evidenza in capitoli precedenti, per risolvere l'equazione tra numeri $ab = ac$ portiamo ac al primo membro cambiandolo di segno e ottenendo $ab - ac = 0$ e poi mettiamo in evidenza a ottenendo $a(b - c) = 0$. In questo caso stiamo sfruttando il fatto che tra numeri è definita un'operazione di somma per la quale valgono la proprietà associativa, l'esistenza di un elemento neutro, l'esistenza di un inverso e la proprietà distributiva. Infatti, per spostare ac al primo membro stiamo in realtà sommando entrambi i membri di $ab = ac$ con l'inverso additivo (ovvero l'opposto) di ac, cioè $-ac$, ottenendo $ab + (-ac) = ac + (-ac) = 0$, ovvero $ab - ac = 0$. Inoltre per mettere in evidenza a sfruttiamo la proprietà distributiva della somma rispetto al prodotto.

Allo scopo di poter eseguire manipolazioni analoghe anche nelle espressioni tra matricidiamo la seguente

Definizione 6.37. La **somma di due matrici** $A, B \in M_{m,n}(\mathbb{K})$, denotata con $A + B$, è la matrice di $M_{m,n}(\mathbb{K})$ la cui entrata di posto ij è la somma delle entrate di posto ij di A e B, ovvero

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

Esempio 6.38. Consideriamo le matrici

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & 1 \end{pmatrix}.$$

Allora

$$A + B = \begin{pmatrix} 1+1 & 2+0 & 3+(-2) \\ 4+3 & 5+1 & 6+1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 7 & 6 & 7 \end{pmatrix}.$$

La somma di matrici così definita ha, tra le altre, proprio le proprietà desiderate citate sopra. Date matrici $A, B, C \in M_{m,n}(\mathbb{K})$ e $D, E \in M_{n,p}(\mathbb{K})$ si ha

(1) Proprietà commutativa: $A + B = B + A$

(2) Proprietà associativa: $(A + B) + C = A + (B + C)$

(3) Esistenza dell'elemento neutro: la matrice 0 le cui entrate sono nulle soddisfa

$$A + 0 = 0 + A = A.$$

In ogni insieme $M_{m,n}(\mathbb{K})$, chiameremo 0 la *matrice nulla*.

(4) Esistenza dell'inverso additivo: ogni matrice A ha un'inversa additiva o opposta, data dalla matrice $-A$ che ha come entrata di posto ij l'opposto dell'entrata di posto ij di A:

$$(-A)_{ij} = -A_{ij}.$$

(5) Proprietà distributiva rispetto al prodotto: si ha $A(D + E) = AD + AE$ e, analogamente, $(A + B)D = AD + BD$.

Quindi, data un'uguaglianza tra matrici $AB = AC$, possiamo eseguire gli stessi passaggi che si eseguono tra numeri: $AB - AC = 0$ e $A(B - C) = 0$.

Osservazione 6.39. Si noti che da un'uguaglianza come l'ultima appena scritta, $A(B - C) = 0$, non possiamo però dedurre $A = 0$ o $B - C = 0$ come faremmo per un'uguaglianza numerica. Questo perché il prodotto di matrici *non* gode della cosiddetta "proprietà di annullamento del prodotto", ovvero può capitare che il prodotto di due matrici non nulle sia la matrice nulla. Ad esempio

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

Ricordiamo che abbiamo già visto altri contesti in cui tale proprietà non vale (ad esempio per il prodotto tra classi nell'aritmetica modulare).

Avendo visto le operazioni di prodotto e somma tra matrici, che sono operazioni *binarie* in quanto ci danno una matrice come risultato di un'operazione tra due matrici, concludiamo vedendo un'operazione *unaria* che ci dà una nuova matrice a partire da una sola data[4], l'operazione di *trasposta di una matrice*.

Definizione 6.40. Data una matrice $A \in M_{m,n}(\mathbb{K})$, definiamo la sua **trasposta** come la matrice $A^T \in M_{n,m}(\mathbb{K})$ che si ottiene da A scambiando le righe con le colonne:

$$A^T_{ij} = A_{ji}.$$

In pratica, in seguito allo scambio gli elementi di A che si trovavano nella riga i e colonna j si vengono a trovare sulla riga j e colonna i.

Esempio 6.41.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Longrightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

L'importanza della trasposta consiste nel fatto che il suo utilizzo consente una più agevole manipolazione di alcune espressioni. Ad esempio, il prodotto scalare $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$ tra due terne (x_1, x_2, x_3) e (y_1, y_2, y_3) può essere riscritto usando il prodotto di matrici e la trasposta. Più precisamente, se consideriamo x e y come matrici con una sola colonna

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

il prodotto scalare $x \cdot y$ coincide con il prodotto righe per colonne tra le matrici

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

[4] Ad esempio, in algebra booleana le operazioni di $\wedge$ e $\vee$ sono operazioni binarie in quanto ci danno una nuova proposizione (rispettivamente $P \wedge Q$ e $P \vee Q$) a partire da due proposizioni P e Q date, mentre $\neg$ è un'operazione unaria che ci dà una nuova proposizione $\neg P$ a partire da una sola proposizione P data.

Questo non è nient'altro che il prodotto righe per colonne $x^T y$ tra la trasposta x^T di x e y. Come vedremo nel prossimo capitolo, scrivere il prodotto scalare in questa forma ha vari vantaggi. Ad esempio, si possono così determinare quali matrici rappresentano trasformazioni rigide dello spazio (ovvero trasformazioni che non modificano lunghezze e angoli). Avremo allora bisogno di conoscere le proprietà della trasposta, tra cui le più importanti sono le seguenti (dimostriamo quelle meno immediate).

(1) $(A + B)^T = A^T + B^T$

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

Dimostrazione. Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,p}(\mathbb{K})$. L'identità discende dalla definizione di prodotto righe per colonne:

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk} B_{ki} = \sum_{k=1}^{n} B^T_{ik} A^T_{kj} = (B^T A^T)_{ij}.$$

□

(3) $(A^T)^T = A$

(4) $\det(A^T) = \det(A)$

Dimostrazione. Sia $A \in M_n(\mathbb{K})$ di entrate a_{ij} e denotiamo con a^T_{ij} le entrate di A^T. Allora, direttamente dalla Definizione 6.17, si ha

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Notiamo che ogni prodotto può essere riscritto come $a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$ semplicemente riordinando rispetto agli indici di colonna. Per cui abbiamo

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma^{-1}) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}.$$

Si noti che $s(\sigma) = s(\sigma^{-1})$. Poiché inoltre associare l'inversa, vista come funzione da S_n in S_n, è biiettiva possiamo concludere che

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\sigma \in S_n} s(\sigma) a^T_{1\sigma(1)} \cdots a^T_{n\sigma(n)} = \det(A^T).$$

□

Illustriamo con un esempio la (2) allo scopo di mettere in evidenza l'ordine invertito delle trasposte a secondo membro

Esempio 6.42. Siano
$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$
allora si ha
$$AB = \begin{pmatrix} -1 & 1 \\ -1 & 3 \end{pmatrix} \Longrightarrow (AB)^T = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}.$$
Si ha inoltre
$$A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \qquad B^T = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$
e quindi, come dimostrato nella (2),
$$B^T A^T = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}.$$
D'altra parte si vede subito che
$$A^T B^T = \begin{pmatrix} 4 & -1 \\ 6 & -2 \end{pmatrix} \neq (AB)^T.$$

Un'ultima operazione riguardante le matrici, spesso usata nel fare i conti, è il prodotto di un elemento del campo numerico $\mathbb{K}$ per una matrice. Più precisamente, si ha la seguente

Definizione 6.43. Dato $c \in \mathbb{K}$ e una matrice $A \in M_{m,n}(\mathbb{K})$, si definisce il **prodotto di** A **per lo scalare** c come la matrice $cA \in M_{m,n}(\mathbb{K})$ che si ottiene moltiplicando ogni entrata di A per c:
$$(cA)_{ij} = c(A_{ij}).$$

Osservazione 6.44. Abbiamo usato implicitamente tale operazione nella formula (6.33) per l'inversa quando abbiamo moltiplicato la matrice dei cofattori per il numero $\frac{1}{\det(A)}$.

Segnaliamo le seguenti proprietà del prodotto di una matrice per uno scalare le cui dimostrazioni sono una conseguenza immediata delle definizioni.

$$c(A + B) = cA + cB \tag{6.47}$$

$$A(cB) = c(AB) \tag{6.48}$$