

Chapter 6

Matrici e trasformazioni: prodotto di matrici e algebra matriciale

In questo capitolo vedremo come le matrici, oltre a rappresentare sistemi di equazioni, possono rappresentare 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.

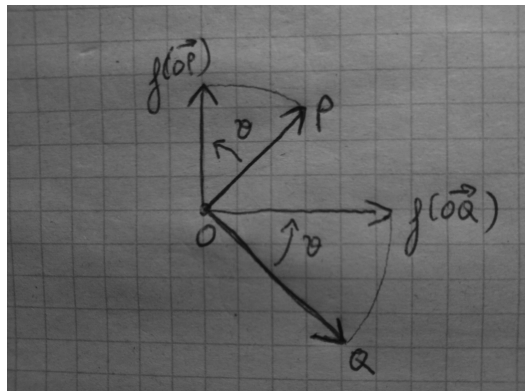
6.1 Matrici e trasformazioni

Mostriamo in questo paragrafo come alcune importanti trasformazioni geometriche possano essere rappresentate mediante una matrice.

Più precisamente, supponiamo di aver fissato nel piano un sistema di coordinate cartesiano ortogonale (ovvero un'origine O e una base $\{\vec{OP}_1, \vec{OP}_2\}$ fatta di vettori ortogonali e di lunghezza 1). 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 sui punti del piano la trasformazione che consiste nel ruotarli 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 vedere la nostra trasformazione come una funzione che opera sui vettori anziché sui punti, ovvero considerare equivalentemente 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 .

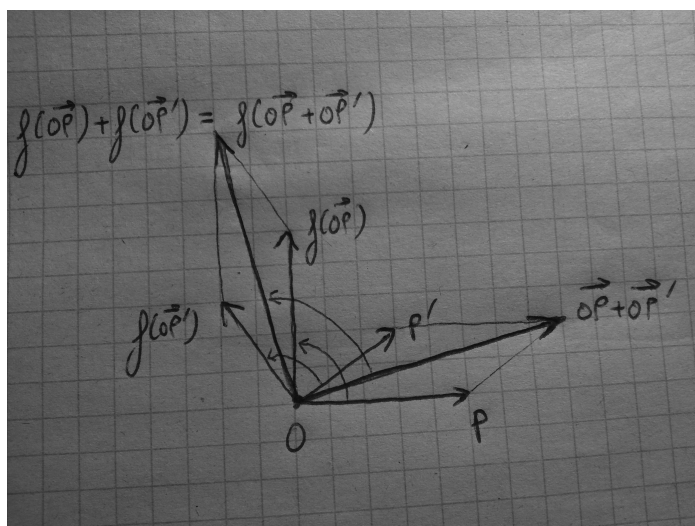


Ora, ci poniamo l'obiettivo di rappresentare tale trasformazione in coordinate, ovvero se (x_1, x_2) sono le coordinate di $\vec{OP}$ vogliamo determinare le coordinate di $f(\vec{OP})$.

A questo scopo, osserviamo preliminarmente che la rotazione ha le seguenti due importanti proprietà:

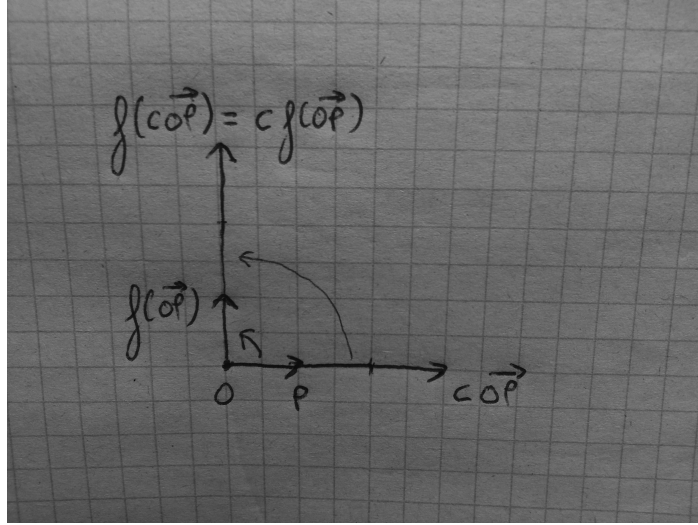
- (1) Dati due vettori $\vec{OP}$ e $\vec{OP}'$, sommarli e poi ruotare il vettore risultante oppure prima ruotarli e poi sommare i vettori ruotati è equivalente, ovvero

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



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

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



Le due proprietà appena osservate ci permettono di risolvere il nostro problema della rappresentazione della rotazione in coordinate: infatti, se P ha coordinate (x_1, x_2) , questo come abbiamo ricordato sopra significa che $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$: 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)

$$f(\vec{OP}) = x_1f(\vec{OP}_1) + x_2f(\vec{OP}_2) \quad (6.3)$$

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, ovvero si avranno uguaglianze del tipo

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

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

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

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

ovvero, svolgendo i conti tramite le proprietà (4.1)-(4.8) della somma di vettori e del prodotto di un vettore per un numero, viste nel Capitolo 4

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

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, per conoscere la rotazione in coordinate basta conoscere i coefficienti $a_{11}, a_{21}, a_{12}, a_{22}$: disponendo tali numeri nella matrice

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

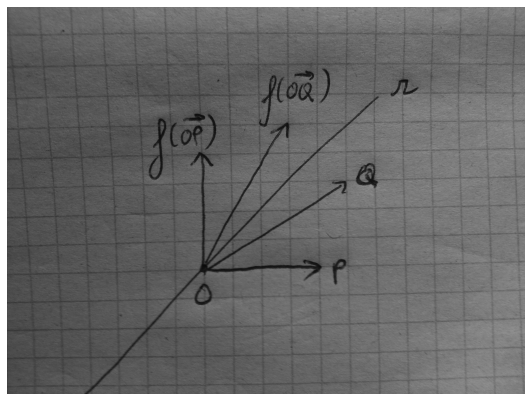
abbiamo che essa rappresenta la rotazione, nel senso che la funzione

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

che associa a ogni coppia¹ (x_1, x_2) la coppia $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$ ci dice chi sono le coordinate di un vettore ruotato a partire dalle coordinate del vettore di partenza.

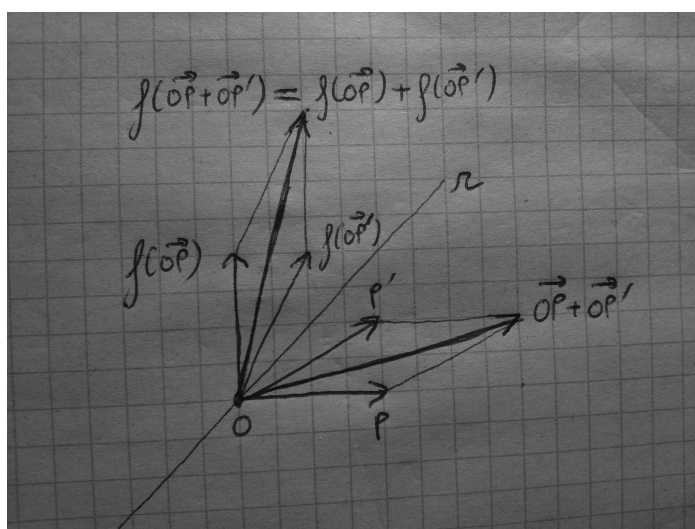
Lo stesso ragionamento fatto per arrivare alla matrice che rappresenta una rotazione nel piano può essere ripetuto anche per altre importanti trasformazioni geometriche: ad esempio la riflessione rispetto a una retta per O , che manda ogni vettore $\vec{OP}$ nel vettore simmetrico rispetto alla retta, come nel seguente disegno

¹Qui e nel seguito indicheremo spesso le n -uple in colonna anziché in riga quando ci sarà comodo per questioni di spazio.



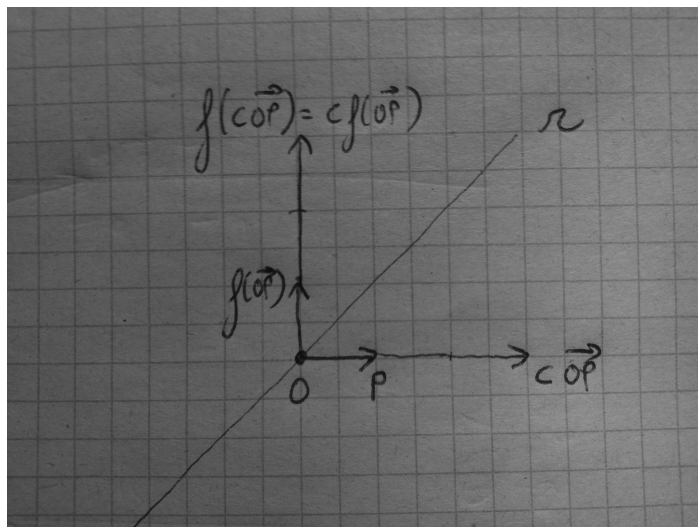
ha esattamente le stesse proprietà (6.1) e (6.2) viste nel caso della rotazione, ovvero

- (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



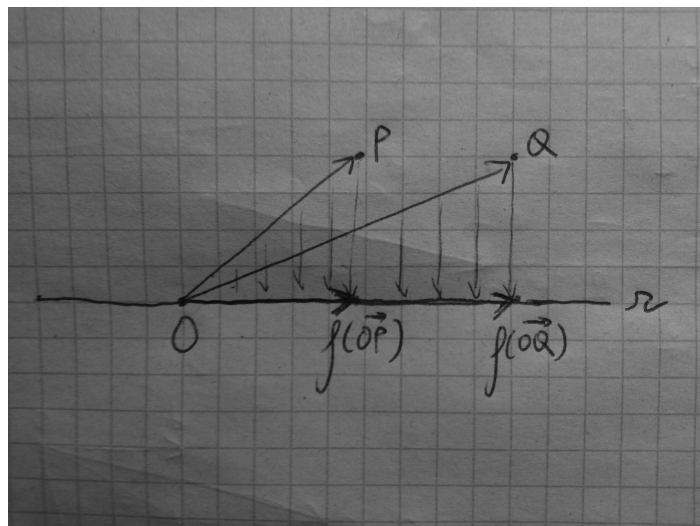
(cioè $f(\vec{OP} + \vec{OP}') = f(\vec{OP}) + f(\vec{OP}')$)

- (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



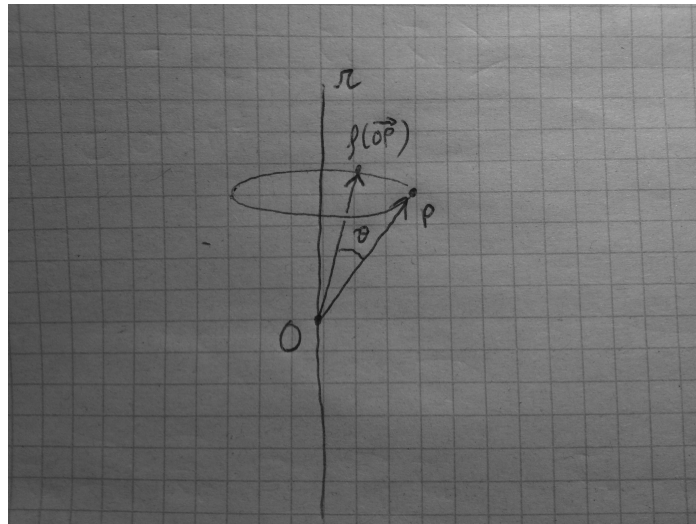
$$(\text{cioè } f(c\vec{OP}) = cf(\vec{OP}))$$

Quindi possiamo ripetere esattamente gli stessi passaggi visti sopra per la rotazione e arrivare a 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, che proietta ortogonalmente i vettori su una retta fissata passante per O .

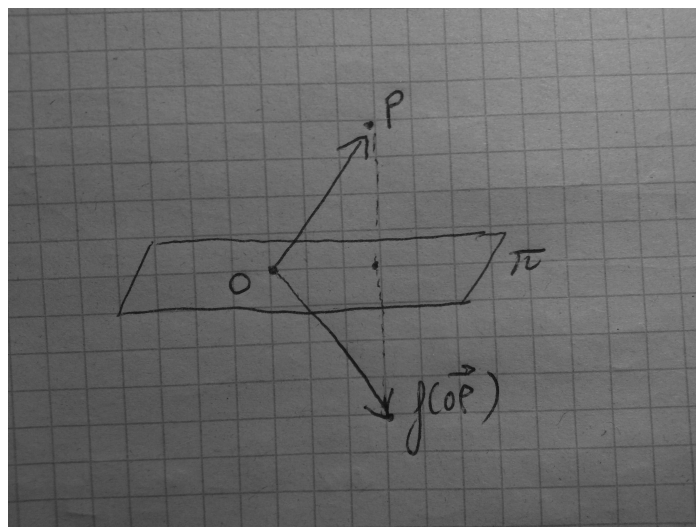


Quanto visto fino ad ora sul piano risulta vero anche per alcune importanti trasformazioni geometriche dello spazio tridimensionale: ad esempio, non è difficile vedere che le proprietà (6.1) e (6.2) sono soddisfatte anche dalla

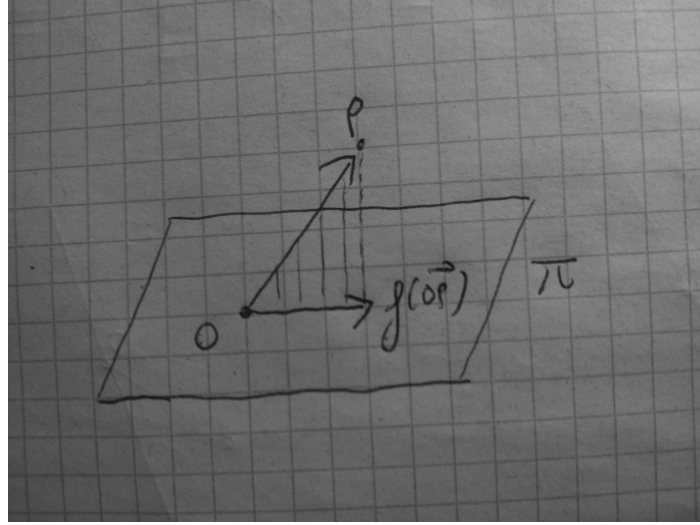
rotazione di un angolo fissato attorno a una retta fissata passante per O
(detta *asse della rotazione*)



dalla riflessione rispetto a un piano passante per O



o 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

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

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

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

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

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

Quindi, stavolta, per conoscere la trasformazione considerata in coordinate basta 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).

Osservazione 6.1. In generale, dato un generico spazio vettoriale V (si veda l'Osservazione 4.1), una funzione $f : V \rightarrow W$ (dove W è un altro spazio vettoriale) con le proprietà che $f(v + v') = f(v) + f(v')$ e $f(cv) = cf(v)$ per ogni $v, v' \in V$ e ogni scalare c si dice *funzione lineare* (o *applicazione lineare*).

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

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

(dove $\mathbb{K}$ può essere il campo $\mathbb{R}$ dei numeri reali o quello $\mathbb{C}$ dei numeri complessi) che risultano più in generale determinate dalla matrice con m righe e n colonne

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

Chiameremo l'applicazione (6.10) associata a tale matrice l'*applicazione lineare determinata da A* e la denoteremo L_A .

Nel paragrafo successivo concentreremo la nostra attenzione su tali funzioni: usando la nozione di composizione vista nel Capitolo 2 riusciremo tramite di esse a definire un'operazione di prodotto tra matrici e considerare quindi 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.

Introduciamo alcune notazioni preliminari: l'insieme delle matrici con m righe e n colonne e 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 $M_n(\mathbb{K})$ come notazione semplificata al posto di $M_{n,n}(\mathbb{K})$.

Una matrice con n righe e n colonne (ovvero un elemento di $M_n(\mathbb{K})$) si dice anche *matrice quadrata di ordine n* .

6.2 Composizione e prodotto di matrici

In questo paragrafo consideriamo il problema di calcolare la composizione di due funzioni del tipo (6.10) visto alla fine del paragrafo precedente.

Più precisamente, consideriamo $f = L_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$,

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

e $g = L_B : \mathbb{K}^p \rightarrow \mathbb{K}^n$:

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

In base a quanto visto nel Paragrafo 2.3, 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

$$\begin{aligned} (f \circ g) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} &= f \left(g \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \right) = f \begin{pmatrix} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2p}y_p \\ \vdots \\ b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p \end{pmatrix} = \\ &= \begin{pmatrix} a_{11}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{1n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ a_{21}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{2n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ \vdots \\ a_{m1}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{mn}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \end{pmatrix} = \end{pmatrix} \quad (6.13)$$

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

$$= \begin{pmatrix} (a_{11}b_{11} + \cdots + a_{1n}b_{n1})y_1 + (a_{11}b_{12} + \cdots + a_{1n}b_{n2})y_2 + \cdots + (a_{11}b_{1p} + \cdots + a_{1n}b_{np})y_p \\ (a_{21}b_{11} + \cdots + a_{2n}b_{n1})y_1 + (a_{21}b_{12} + \cdots + a_{2n}b_{n2})y_2 + \cdots + (a_{21}b_{1p} + \cdots + a_{2n}b_{np})y_p \\ \vdots \\ (a_{m1}b_{11} + \cdots + a_{mn}b_{n1})y_1 + (a_{m1}b_{12} + \cdots + a_{mn}b_{n2})y_2 + \cdots + (a_{m1}b_{1p} + \cdots + a_{mn}b_{np})y_p \end{pmatrix} \quad (6.14)$$

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

$$C = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1p} + \cdots + a_{mn}b_{np} \end{pmatrix} \quad (6.15)$$

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

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} \quad (6.16)$$

in cui il primo indice dell'entrata di A è fisso e sempre uguale a i , il secondo indice dell'entrata di B è fisso e sempre uguale a j mentre 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 $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$).

Diamo allora la seguente importante

Definizione 6.2. La matrice C le cui entrate sono date dalla (6.16) si chiama *prodotto di A per B* , e si scrive $C = AB$.

In pratica, abbiamo definito il prodotto di due matrici A e B in modo che la matrice $C = AB$ ottenuta sia la matrice che determina la composizione $L_A \circ L_B$ delle funzioni determinate da A e B .

Ricordando 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 $L_B : \mathbb{K}^p \rightarrow \mathbb{K}^n$ deve essere uguale al dominio di $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$) si osserva di conseguenza che anche il prodotto di due matrici può essere fatto solo sotto opportune condizioni: più precisamente, dal momento che la matrice di L_A ha m righe e n colonne, mentre la matrice B di L_B ha n righe e p colonne, vediamo che 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 6.2 si chiama anche *prodotto righe per colonne*: il motivo è che nell'espressione (6.16) 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 indice ci dice in che riga siamo, e il secondo in quale colonna, questo significa che per calcolare l'entrata di posto i, j di AB dobbiamo prendere la riga i -esima di A , la j -esima colonna di B , moltiplicare tra di loro le entrate corrispondenti e sommare il tutto

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

Esempio 6.3. 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, le due matrici possono essere moltiplicate 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 prendono *la prima riga di A* $\begin{pmatrix} 1 & 2 \end{pmatrix}$, *la prima colonna di B* $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$, se ne moltiplicano gli elementi corrispondenti (il primo con il primo, il secondo con il secondo) e si somma il risultato: $1 \cdot 5 + 2 \cdot 3 = 5 + 6 = 11$. Quindi nella prima riga, prima colonna di AB dobbiamo scrivere 11.

Analogamente, per trovare l'entrata di AB che sta *nella prima riga e seconda colonna* si prendono *la prima riga di A* $\begin{pmatrix} 1 & 2 \end{pmatrix}$, *la seconda colonna di B* $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, se ne moltiplicano gli elementi corrispondenti e si somma il risultato: $1 \cdot 2 + 2 \cdot 1 = 2 + 2 = 4$. Quindi nella prima riga, seconda colonna di AB c'è 4.

Facendo questo calcolo per tutte le entrate, si vede che

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

Vediamo ora le proprietà fondamentali del prodotto di matrici.

In primo luogo, è importante osservare che tale prodotto *non gode della proprietà commutativa*, ovvero, date due matrici A e B e ammesso che si

possano eseguire entrambi i prodotti AB e BA , questi non saranno in generale uguali².

Ad esempio, siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix}$. Allora, per definizione di prodotto righe per colonne vediamo che

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

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

ovvero $AB \neq BA$.

Per quello che riguarda invece la proprietà associativa, si può dimostrare che questa vale, ovvero $(AB)C = A(BC)$.

Non diamo una dimostrazione di questo fatto, ma limitiamoci a illustrarlo con un esempio: 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, 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 matrice con n righe e n colonne seguente

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & 1 \end{pmatrix} \quad (6.17)$$

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

²Questo 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 a pagina 45).

Tale matrice si chiama *matrice identica di ordine n* e si denota I_n . Ad esempio,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Osservazione 6.4. In generale, data una matrice quadrata di ordine n , le entrate $a_{11}, a_{22}, \dots, a_{nn}$ che hanno stesso indice di riga e di colonna formano la cosiddetta *diagonale della matrice*. La matrice identica I_n può essere quindi descritta come la matrice che ha 1 sulla diagonale e 0 nelle altre entrate. Le entrate di I_n si denotano solitamente con il simbolo δ_{ij} , detto *delta di Kronecker*, che vale quindi 1 se $i = j$, e 0 se $i \neq j$.

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

$$AI_n = A, \quad I_m A = A$$

(l'ordine della matrice identica cambia perché deve essere tale che si possa svolgere il prodotto) e quindi la matrice identica svolge esattamente il ruolo di elemento neutro per il prodotto righe per colonne.

Ad esempio,

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

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

Osservazione 6.5. Si noti che la matrice identica I_n non è nient'altro che la matrice che determina la funzione identica $id_{\mathbb{K}^n} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ che manda ogni elemento in se stesso: infatti,

$$id_{\mathbb{K}^n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1x_1 + 0x_2 + \dots + 0x_n \\ 0x_1 + 1x_2 + \dots + 0x_n \\ \vdots \\ 0x_1 + 0x_2 + \dots + 1x_n \end{pmatrix} \quad (6.18)$$

Questo spiega perché tale matrice sia l'elemento neutro per il prodotto, in quanto il prodotto tra matrici rappresenta la composizione delle applicazioni corrispondenti e la funzione identica è esattamente l'elemento neutro per la composizione, come visto a pagina 47.

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 $ab = ba = 1$), ci chiediamo per quali matrici A esista una matrice B tale che $AB = I$ e $BA = I$ (dove I è una matrice identica di ordine opportuno, dipendente dagli ordini di A e B). Chiameremo tali matrici *invertibili* e B sarà detta *l'inversa di A* .

Si ha il seguente

Teorema 6.6. 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 $AB = I_n$ e $BA = I_n$.

Vedremo più sotto alla fine del paragrafo (Osservazione 6.7) di dare una spiegazione al perché possano essere invertibili solo le matrici quadrate. 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.

Delle due condizioni $AB = I_n$ e $BA = I_n$ concentriamoci sulla prima, che equivale a scrivere

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

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

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

ovvero la prima colonna di B soddisfa il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \quad (6.19)$$

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

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

ovvero la seconda colonna di B deve soddisfare il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 1 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \quad (6.20)$$

sempre con matrice dei coefficienti uguale a A ma stavolta con termini noti uguali alla seconda colonna della matrice identica, e così via possiamo ragionare allo stesso modo fino all'ultima colonna di B che dovrà soddisfare

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

ovvero il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 1 \end{cases} \quad (6.21)$$

Riassumendo, si ha $AB = I_n$ se e solo se le colonne $\begin{pmatrix} b_{11} \\ b_{21} \\ \cdots \\ b_{n1} \end{pmatrix}, \begin{pmatrix} b_{12} \\ b_{22} \\ \cdots \\ b_{n2} \end{pmatrix}, \dots,$

$\begin{pmatrix} b_{1n} \\ b_{2n} \\ \cdots \\ b_{nn} \end{pmatrix}$ sono soluzione rispettivamente degli n sistemi (6.19), (6.20), $\dots$, (6.21).

Ora, come sappiamo, per verificare se un sistema ha soluzioni basta scrivere la sua matrice completa e ridurla a gradini mediante operazioni elementari sulle righe.

Poiché i sistemi (6.19), (6.20), ..., (6.21) hanno tutti la stessa matrice dei coefficienti, cioè A , e differiscono solo per i termini noti, le operazioni elementari che eseguiamo per ridurli sono le stesse, quindi conviene svolgerle una sola volta: a questo scopo, basta scrivere la matrice

$$(A|I_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ & & \dots & & & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array} \right) \quad (6.22)$$

ottenuta affiancando tutti i termini noti dei sistemi (6.19), (6.20), ..., (6.21) e risolverli contemporaneamente con una sola riduzione.

Ora, 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 (6.22) siano tutti compatibili, dovrebbe annullarsi anche tutta la riga del blocco I_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 di I_n , questo vorrebbe dire che le righe di I_n sono linearmente dipendenti, il che è falso: ad esempio non è possibile ottenere la prima riga $(1, 0, \dots, 0)$ come combinazione delle altre perché queste ultime hanno tutte prima componente nulla, e qualunque loro combinazione avrebbe ancora prima componente nulla (e quindi non potrebbe essere uguale a $(1, 0, \dots, 0)$ che ha prima componente uguale a 1), e per le altre righe si può ragionare in maniera analoga.

Quindi se la matrice di A avesse rango minore di n in seguito alla riduzione in contemporanea degli n sistemi (6.19), (6.20), ..., (6.21) comparirebbe necessariamente un'uguaglianza falsa $0 = b \neq 0$, ovvero almeno uno di tali sistemi sarebbe incompatibile e l'uguaglianza $AB = I_n$ 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 (6.22) nella parte dei coefficienti non si annulla nessuna riga, tutti i sistemi (6.19), (6.20), ..., (6.21) hanno soluzione e quindi esiste una matrice B tale che $AB = I_n$. Si può a questo punto mostrare (ma omettiamo i dettagli) che la matrice B così trovata soddisfa anche l'uguaglianza inversa $BA = I_n$, ovvero A è invertibile con inversa uguale a B .

Ispirandoci a quanto appena detto, verifichiamo se la matrice $A = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$ è invertibile e, in caso affermativo, calcoliamone l'inversa. Come fatto sopra, affianchiamo a tale matrice la matrice identica dello stesso ordine

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

che rappresenta i due sistemi le cui soluzioni sono le colonne della matrice inversa, e iniziamo con l'applicare il procedimento di riduzione a gradini: a questo scopo basta il singolo passaggio

$$(A|I_n) = \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - R_1} \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array} \right) \quad (6.23)$$

da cui vediamo che, dopo la riduzione a gradini, ad A non si annulla nessuna riga e quindi, come abbiamo detto sopra, A è invertibile.

La prima colonna dell'inversa B di A è data dalla soluzione del sistema ridotto

$$\begin{cases} x_1 - x_2 = 1 \\ 3x_2 = -1 \end{cases} \quad (6.24)$$

ovvero, come si vede risolvendo dal basso, la coppia $(\frac{2}{3}, -\frac{1}{3})$, 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

$$\begin{cases} x_1 - x_2 = 0 \\ 3x_2 = 1 \end{cases} \quad (6.25)$$

ovvero, come si vede risolvendo dal basso, la coppia $(\frac{1}{3}, \frac{1}{3})$, 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 (6.24) e (6.25) e trovare invece in modo più diretto la matrice inversa, si può procedere come segue. Dopo aver effettuato la riduzione a gradini in (6.23), 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 legge direttamente l'inversa. Per vederlo, riprendiamo da (6.23) e facciamo comparire prima uno zero in posizione 1 2 eseguendo

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow 3R_1 + R_2} \left(\begin{array}{cc|cc} 3 & 0 & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array} \right)$$

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

$$\left(\begin{array}{cc|cc} 3 & 0 & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array} \right) \xrightarrow[\substack{R_1 \rightarrow (1/3)R_1 \\ R_2 \rightarrow (1/3)R_2}]{\substack{R_1 \rightarrow (1/3)R_1 \\ R_2 \rightarrow (1/3)R_2}} \left(\begin{array}{cc|cc} 1 & 0 & 2/3 & 1/3 \\ 0 & 1 & -1/3 & 1/3 \end{array} \right). \quad (6.26)$$

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 (6.26) non stiamo facendo altro che ridurre i due sistemi alla forma

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

e cioè far comparire direttamente le soluzioni cercate (che sono proprio le colonne della matrice inversa).

Vediamo un altro esempio:

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

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

$$(A|I_n) = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow[\substack{R_2 \rightarrow R_2 + R_1 \\ R_3 \rightarrow R_3 - 2R_1}]{\substack{R_2 \rightarrow R_2 + R_1 \\ R_3 \rightarrow R_3 - 2R_1}} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array} \right)$$

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

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 “all’incontrario”, dal basso verso l’altro e da destra verso sinistra (in ogni passaggio, mettiamo in evidenza in grassetto i nuovi zeri che facciamo comparire)

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array} \right) \xrightarrow[R_1 \rightarrow 2R_1 + R_3]{R_2 \rightarrow 2R_2 + R_3} \left(\begin{array}{ccc|ccc} 2 & 2 & \mathbf{0} & -1 & 1 & 2 \\ 0 & 4 & \mathbf{0} & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array} \right)$$

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

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

$$\left(\begin{array}{ccc|ccc} 4 & 0 & 0 & -1 & -1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array} \right) \xrightarrow[R_3 \rightarrow -(1/4)R_3]{R_1 \rightarrow (1/4)R_1, R_2 \rightarrow (1/4)R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/4 & -1/4 & 1/2 \\ 0 & 1 & 0 & -1/4 & 3/4 & 1/2 \\ 0 & 0 & 1 & 3/4 & -1/4 & -1/2 \end{array} \right)$$

Quindi

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

Osservazione 6.7. 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 delle due condizioni $AB = I$ e $BA = I$ che devono essere entrambe verificate per poter dire che A è invertibile, se $m < n$ allora non è mai verificata la seconda, mentre se $m > n$ allora non è mai verificata la prima.

Infatti, sia $m < n$ e supponiamo per assurdo che esista una matrice B con n righe e m colonne³ tale che $BA = I_n$. Dal momento che il prodotto di due matrici A e B corrisponde alla composizione delle funzioni L_A e L_B da loro determinate, $BA = I_n$ equivale a dire che la composizione $L_B \circ L_A$ è uguale all'applicazione identica (che è la funzione che corrisponde alla matrice identica): come abbiamo detto nel Capitolo 2 (pagina 47), questo significa che L_A ha un'inversa sinistra. Ma, come abbiamo visto nella Proposizione 2.24, una funzione ha un'inversa sinistra se e solo se è iniettiva.

L'assurdo si otterrà mostrando che se $m < n$ allora la funzione

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

non può essere iniettiva.

Per definizione, L_A è iniettiva se non ci sono due n -uple $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ diverse che vanno a finire mediante la L_A in una stessa m -upla di arrivo $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$, ovvero

non deve succedere che il sistema

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

abbia più di una soluzione per una certa m upla di arrivo fissata. Tuttavia, come sappiamo dalla teoria dei sistemi un sistema può avere soluzione unica solo se il rango della matrice dei coefficienti è uguale al numero di incognite, ovvero il numero di righe non nulle dopo la riduzione a gradini della matrice

³Da una parte, B deve avere necessariamente m colonne per poter essere moltiplicata a sinistra di A (che ha m righe), dall'altra deve anche avere necessariamente n righe (cioè tante quante le colonne di A) affinché il risultato sia la matrice identica che è una matrice quadrata.

A dovrebbe essere n . Tuttavia, essendo $m < n$, le righe sono già in partenza meno di n e quindi questa condizione non potrà mai essere verificata.

Riassumendo, se $m < n$ la funzione L_A non può essere iniettiva, quindi non può avere un'inversa sinistra e quindi l'uguaglianza $BA = I_n$ non può verificarsi.

Il ragionamento appena fatto mostra anche che

Proposizione 6.8. 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

Analogamente, sia $m > n$ e supponiamo per assurdo che esista una matrice B con n righe e m colonne⁴ tale che $AB = I_m$. Come sopra, questo equivale a dire che la composizione $L_A \circ L_B$ è uguale all'applicazione identica, il che significa che L_A ha un'inversa destra. Ma, come abbiamo visto nella Proposizione 2.23, una funzione ha un'inversa sinistra se e solo se è suriettiva.

Per definizione, L_A è suriettiva se per ogni m -upla di arrivo $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$ esiste

una n -upla $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ che ci va a finire mediante L_A , ovvero il sistema

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

deve essere compatibile per qualunque scelta dei termini noti $y_1, y_2, \dots, y_m$. Tuttavia, essendo $m > n$, quando andiamo a ridurre la matrice A a gradini per risolvere tale sistema necessariamente almeno una riga si deve annullare, altrimenti il rango sarebbe m , ovvero ci sarebbero m righe indipendenti e quindi (il rango per righe è uguale al rango per colonne) m colonne indipendenti: ma essendo le colonne in numero n minore di m , questo non è possibile. Ma se una riga si annulla, affinché il sistema sia compatibile deve annullarsi anche il termine noto corrispondente, e non è possibile che questo succeda per qualunque scelta dei termini noti.

⁴Come sopra, da una parte B deve avere necessariamente m colonne per poter essere moltiplicata a destra di A (che ha m righe), dall'altra deve avere necessariamente n righe affinché il risultato sia la matrice identica che è una matrice quadrata.

Riassumendo, se $m > n$ la funzione L_A non può essere suriettiva, quindi non può avere un'inversa destra e quindi l'uguaglianza $AB = I_m$ non può verificarsi.

Il ragionamento appena fatto mostra anche che

Proposizione 6.9. 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

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 2.30) rispetto all'operazione di prodotto di matrici che si denota $GL_n(\mathbb{K})$.

Infatti, prima di tutto osserviamo che il prodotto AB di due matrici A, B invertibili è ancora invertibile perché ha come inversa $B^{-1}A^{-1}$: infatti

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

e analogamente si vede che $(B^{-1}A^{-1})(AB) = I_n$. Questo mostra che il prodotto tra due elementi di $GL_n(\mathbb{K})$ rimane dentro $GL_n(\mathbb{K})$, ovvero il prodotto è ben definito dentro $GL_n(\mathbb{K})$.

Essendo il prodotto di matrici associativo, la proprietà (i) della Definizione 2.30 è verificata; inoltre, la matrice identica I_n , che funge da elemento neutro per il prodotto, è invertibile (essendo $I_n \cdot I_n = I_n$, essa ha come inversa se stessa), quindi sta dentro $GL_n(\mathbb{K})$: anche la proprietà (ii) della Definizione 2.30 è verificata. Per quello che riguarda la proprietà (iii), che chiede che ogni elemento di $GL_n(\mathbb{K})$ abbia un inverso contenuto anch'esso in $GL_n(\mathbb{K})$, essa è verificata in quanto per definizione ogni matrice A di $GL_n(\mathbb{K})$ è invertibile e ha un'inversa A^{-1} , e questa inversa sta chiaramente dentro $GL_n(\mathbb{K})$ perché è chiaramente a sua volta invertibile (ha come inversa proprio A).

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 2.34, che afferma che in qualunque gruppo G l'inverso del prodotto $g_1g_2 \cdots g_k$ è $g_k^{-1} \cdots g_2^{-1}g_1^{-1}$, possiamo affermare che il prodotto $A_1A_2 \cdots A_k$ di k matrici invertibili $A_1, A_2, \dots, A_k$ è dato da $A_k^{-1} \cdots A_2^{-1}A_1^{-1}$, senza bisogno di ridimostrarlo.