

CAPITOLO II: MATRICI.¹

Vincenzo Di Gennaro

1. Richiami e notazioni.

Nel capitolo precedente abbiamo dato la definizione di matrice, e, fissati interi $m, n \geq 1$, abbiamo visto che l'insieme di tutte le matrici $\mathcal{M}(m, n)$ con m righe ed n colonne e' il sostegno di uno spazio vettoriale $(\mathcal{M}(m, n), +, \cdot)$ in cui l'addizione interna e la moltiplicazione esterna sono cosi' definite. Per ogni coppia di matrici $A = (a_{ij})$ e $B = (b_{ij})$ in $\mathcal{M}(m, n)$, ed ogni scalare $c \in \mathbf{R}$, si pone:

$$A + B := (a_{ij} + b_{ij}), \quad c \cdot A := (ca_{ij}).$$

Data una matrice $A = (a_{ij}) \in \mathcal{M}(m, n)$, denoteremo con A_i la sua riga di posto i , mentre con A^j la sua colonna di posto j . Si osservi che A_i e' un vettore numerico di $\mathbf{R}^n = \mathcal{M}(1, n)$, mentre A^j e' una matrice con una colonna ed m righe, cioe' $A^j \in \mathcal{M}(m, 1)$. Per esempio, se

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathcal{M}(2, 3)$$

allora

$$A_1 = (1, 2, 3), \quad A_2 = (4, 5, 6) \in \mathbf{R}^3 = \mathcal{M}(1, 3),$$

mentre

$$A^1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \in \mathcal{M}(2, 1).$$

Data una matrice $A = (a_{ij}) \in \mathcal{M}(m, n)$, denoteremo con A^T la *matrice trasposta di* A , cioe' quella matrice $A^T \in \mathcal{M}(n, m)$ che si ottiene da A scambiando le righe con le colonne. Con riferimento all'esempio precedente, si ha:

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \in \mathcal{M}(3, 2).$$

In modo formale, possiamo dire che se $A = (a_{ij}) \in \mathcal{M}(m, n)$ allora $A^T = (b_{ij}) \in \mathcal{M}(n, m)$, con $b_{ij} := a_{ji}$. In particolare se $\mathbf{x} \in \mathbf{R}^n$ e' un vettore numerico, allora \mathbf{x}^T e' lo stesso vettore \mathbf{x} messo in colonna:

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \quad \mathbf{x}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

¹ Ultimo aggiornamento: 26 ottobre 2019

Sia ora A una matrice $n \times n$, cioè una matrice che ha lo stesso numero di colonne e di righe. Una tale matrice si dice *quadrata di ordine n* . In una matrice quadrata di ordine n , gli elementi del tipo a_{ii} compongono un vettore numerico (a_{11}, \dots, a_{nn}) di ordine n , che viene detto *la diagonale principale di A* . Una matrice quadrata A si dice *simmetrica* se coincide con la propria trasposta, cioè se $A = A^T$. Ciò equivale a dire che per ogni componente a_{ij} di A si ha $a_{ij} = a_{ji}$. Di seguito vediamo la generica matrice simmetrica di ordine 2 e di ordine 3.

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}.$$

Si osservi che, in una matrice simmetrica, le componenti della diagonale principale non sono sottoposte ad alcuna condizione. Una matrice quadrata $A = (a_{ij})$ di ordine n si dice *matrice diagonale* se tutte le componenti di A fuori della diagonale principale sono nulle, cioè se $a_{ij} = 0$ per ogni $i \neq j$. Ecco la generica matrice diagonale di ordine 2 e 3.

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}.$$

Una matrice quadrata $A = (a_{ij})$ si dice *triangolare superiore* se tutte le componenti di A al di sotto della diagonale principale sono nulle, cioè se $a_{ij} = 0$ ogniqualvolta $i > j$. Esempio di ordine 3:

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}.$$

Una matrice quadrata $A = (a_{ij})$ si dice *triangolare inferiore* se tutte le componenti di A al di sopra della diagonale principale sono nulle, cioè se $a_{ij} = 0$ ogniqualvolta $i < j$. Esempio di ordine 3:

$$\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}.$$

E' chiaro che A e' triangolare inferiore se e solo se A^T e' triangolare superiore, così come A e' diagonale se e solo se e' simultaneamente triangolare superiore ed inferiore.

2. Prodotto riga per colonna tra matrici.

Ci proponiamo di definire una moltiplicazione tra matrici, la cui motivazione sara' chiara quando studieremo le applicazioni lineari nel Capitolo 4. La moltiplicazione che definiremo agisce così. Si parte da una coppia di matrici $A \in \mathcal{M}(m, n)$ e $B \in \mathcal{M}(n, p)$ tali che la seconda matrice B abbia tante righe quante sono le colonne della prima matrice A . Il prodotto $A \cdot B$ e' una matrice con lo stesso numero di righe di A , e lo stesso numero di colonne di B , cioè $A \cdot B$ e' una matrice di tipo $m \times p$.

$$A \in \mathcal{M}(m, n), \quad B \in \mathcal{M}(n, p) \longrightarrow A \cdot B \in \mathcal{M}(m, p).$$

Si osservi che la moltiplicazione tra matrici non e' una operazione interna allo spazio delle matrici, tranne nel caso notevole $m = n = p$, di cui parleremo tra poco. Prima di definire il prodotto tra matrici, abbiamo bisogno di una definizione preliminare.

- *Il prodotto punto tra vettori numerici.*

Siano \mathbf{x}, \mathbf{y} vettori numerici di \mathbf{R}^n . Si definisce prodotto punto $\mathbf{x} \cdot \mathbf{y}$ tra \mathbf{x} e \mathbf{y} il seguente numero:

$$\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \in \mathbf{R}.$$

Per esempio, se ci mettiamo in \mathbf{R}^3 allora:

$$(1, 0, 0) \cdot (1, 0, 0) = 1, \quad (1, 0, 0) \cdot (0, 1, 0) = 0, \quad (1, -4, 2) \cdot \left(-3, 5, \frac{1}{2}\right) = -22.$$

Il prodotto punto puo' essere visto come un'applicazione

$$\mathbf{x}, \mathbf{y} \in \mathbf{R}^n \rightarrow \mathbf{x} \cdot \mathbf{y} \in \mathbf{R},$$

che nel caso $n = 1$ si riduce alla moltiplicazione interna tra numeri reali, mentre per $n > 1$ non e' una operazione interna ad \mathbf{R}^n . Applicazioni di questo tipo saranno l'oggetto di studio del Capitolo 7 sulle forme quadratiche. In qualche libro il prodotto punto viene chiamato prodotto scalare (canonico, o standard). Le proprieta' di calcolo del prodotto punto sono le seguenti.

- 1) Per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^n$ ed ogni scelta di scalari $a, b \in \mathbf{R}$ si ha

$$(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a\mathbf{x} \cdot \mathbf{z} + b\mathbf{y} \cdot \mathbf{z}.$$

Tale proprieta' viene detta *linearita' a sinistra* del prodotto punto.

- 2) Per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ si ha

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}.$$

Tale proprieta' viene detta *proprieta' di simmetria* del prodotto punto.

- 3) Per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^n$ ed ogni scelta di scalari $a, b \in \mathbf{R}$ si ha

$$\mathbf{x} \cdot (a\mathbf{y} + b\mathbf{z}) = a\mathbf{x} \cdot \mathbf{y} + b\mathbf{x} \cdot \mathbf{z}.$$

Tale proprieta' viene detta *linearita' a destra* del prodotto punto, e chiaramente segue dalle prime due proprieta'.

La dimostrazione di tali proprieta' si riconduce facilmente alla associativita', alla commutativita', ed alla distributivita' in \mathbf{R} . Siamo in condizione di definire il prodotto tra matrici.

- *La definizione di prodotto riga per colonna tra matrici.*

Si considerino matrici $A \in \mathcal{M}(m, n)$ e $B \in \mathcal{M}(n, p)$. Si definisce prodotto riga per colonna $C := A \cdot B$ della matrice A con la matrice B quella matrice $C = (c_{ij}) \in \mathcal{M}(m, p)$ con:

$$c_{ij} = A_i \cdot B^j,$$

cioe' quella matrice $C = (c_{ij}) \in \mathcal{M}(m, p)$ la cui componente c_{ij} di posto ij coincide con il prodotto punto della riga A_i di posto i di A con la colonna B^j di posto j di B . Si osservi che proprio perche' il numero di colonne di A coincide con il numero di righe di B , allora A_i e B^j sono vettori numerici della stessa lunghezza n , e percio' il prodotto punto $A_i \cdot B^j$ e' eseguibile. Naturalmente conveniamo nella definizione precedente di considerare il vettore colonna B^j come un vettore di \mathbf{R}^n . Inoltre osserviamo che se e' definito il prodotto $A \cdot B$ allora il prodotto $B \cdot A$ e' definito solo nel caso $m = p$.

Esempio. Calcoliamo il prodotto riga per colonna delle seguenti matrici:

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

Allora

$$A \cdot B = \begin{bmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 \\ A_2 \cdot B^1 & A_2 \cdot B^2 \end{bmatrix} = \begin{bmatrix} -7 & -4 \\ -16 & -10 \end{bmatrix}.$$

Osservazione. Siano \mathbf{x} ed \mathbf{y} vettori numerici in \mathbf{R}^n . Osserviamo che il prodotto punto tra \mathbf{x} ed \mathbf{y} puo' essere riguardato anche come il prodotto riga per colonna di \mathbf{x} con \mathbf{y}^T :

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y}^T.$$

Conseguenze dirette delle proprieta' del prodotto punto sono le proprieta' di calcolo del prodotto tra matrici.

• *Proprieta' di calcolo del prodotto tra matrici.*

1) Comunque si assegnino matrici $A \in \mathcal{M}(m, n)$, $B \in \mathcal{M}(n, p)$, $C \in \mathcal{M}(p, q)$ si ha:

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

Tale proprieta' si chiama *la proprieta' associativa del prodotto tra matrici*.

2) Comunque si assegnino matrici $A \in \mathcal{M}(m, n)$, $B, C \in \mathcal{M}(n, p)$ si ha:

$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

E comunque si assegnino matrici $A, B \in \mathcal{M}(m, n)$, $C \in \mathcal{M}(n, p)$ si ha:

$$(A + B) \cdot C = A \cdot C + B \cdot C.$$

Tale proprieta' si chiama *la proprieta' distributiva del prodotto tra matrici*.

3) Comunque si assegnino matrici $A \in \mathcal{M}(m, n)$, $B \in \mathcal{M}(n, p)$, ed uno scalare $c \in \mathbf{R}$, si ha:

$$c(A \cdot B) = (cA) \cdot B = A \cdot (cB).$$

Osservazione. Consideriamo due matrici $A \in \mathcal{M}(m, n)$, $B \in \mathcal{M}(n, p)$. Osserviamo che in generale $A \cdot B \neq B \cdot A$. Innanzitutto affinché $A \cdot B$ e $B \cdot A$ siano confrontabili occorre che $m = n = p$, cioè che A e B siano quadrate dello stesso ordine. Ed in tal caso può accadere che $A \cdot B \neq B \cdot A$. Per esempio:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Come già osservato, è notevole il caso in cui si esegua il prodotto riga per colonna tra matrici quadrate dello stesso ordine, diciamo n . Infatti in tal caso possiamo riguardare il prodotto tra matrici come un'operazione interna:

$$A, B \in \mathcal{M}(n, n) \rightarrow A \cdot B \in \mathcal{M}(n, n).$$

Ne consegue che sull'insieme $\mathcal{M}(n, n)$ abbiamo tre operazioni: un'operazione di addizione interna $+$, una operazione di moltiplicazione esterna \cdot (per cui $(\mathcal{M}(n, n), +, \cdot)$ è uno spazio vettoriale), ed una operazione di moltiplicazione interna \cdot , data dal prodotto riga per colonna. In tal caso si dice che la struttura algebrica

$$(\mathcal{M}(n, n), +, \cdot, \cdot)$$

è un'algebra, detta anche *l'algebra delle matrici quadrate di ordine n* . Nel caso $n = 1$ tale algebra si riduce all'algebra dei numeri reali $(\mathbf{R}, +, \cdot, \cdot)$. Tuttavia ci sono delle notevoli differenze tra il caso $n = 1$ ed il caso $n > 1$. Infatti nel caso $n = 1$, cioè nel caso dell'algebra dei numeri reali, sappiamo che la moltiplicazione interna soddisfa le seguenti proprietà:

- 1) Per ogni $a, b, c \in \mathbf{R}$ si ha $(ab)c = a(bc)$ (proprietà associativa della moltiplicazione tra numeri reali);
- 2) Esiste un elemento neutro 1, tale cioè che $1 \cdot a = a \cdot 1 = a$ per ogni $a \in \mathbf{R}$ (esistenza dell'elemento neutro);
- 3) Per ogni $a \in \mathbf{R} \setminus \{0\}$ esiste un elemento inverso b tale che $ab = ba = 1$ (esistenza dell'inverso);
- 4) Per ogni $a, b \in \mathbf{R}$ si ha $ab = ba$ (proprietà commutativa).²

Nel caso $n > 1$ le prime due proprietà continuano a valere. Infatti già sappiamo che vale la proprietà associativa. E c'è anche l'elemento neutro per la moltiplicazione. Infatti la matrice diagonale I di ordine n , che ha sulla diagonale principale tutti 1, funge da elemento neutro. Cioè:

$$A \cdot I = I \cdot A = A$$

per ogni $A \in \mathcal{M}(n, n)$. La matrice I si chiama anche *la matrice identica*, e può essere descritta dalle seguenti condizioni.

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

²Inoltre ricordiamo che valgono anche le proprietà distributive.

Nel caso $n = 1, 2, 3$ abbiamo:

$$I = 1, \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Invece abbiamo appena visto che la proprieta' commutativa in generale e' falsa, cosi' come non e' detto che una matrice non nulla abbia un inverso. Per esempio, se consideriamo la matrice

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

allora non esiste nessuna matrice B tale che $A \cdot B = B \cdot A = I$. Infatti se

$$B = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$$

e' una qualsiasi matrice di ordine 2 allora:

$$A \cdot B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

L'esempio precedente motiva la seguente definizione.

- *La definizione di matrice invertibile.*

Una matrice quadrata A di ordine n si dice *invertibile* se esiste una matrice quadrata B di ordine n tale che $AB = BA = I$. In tal caso B dicesi la *matrice inversa* di A , e si denota col simbolo $A^{-1} := B$.

Abbiamo visto che non tutte le matrici sono invertibili. Si pone percio' naturale la questione su come riconoscere il fatto che una matrice sia invertibile. Questo e' un problema che riprenderemo e risolveremo dopo. Per il momento osserviamo le proprieta' generali delle matrici invertibili.

- *Le proprieta' generali delle matrici invertibili.*

- 1) *Se A e' invertibile, la sua inversa e' unica.*
- 2) *Se A e' invertibile allora lo e' anche la sua inversa A^{-1} , e $(A^{-1})^{-1} = A$.*
- 3) *La matrice identica I e' invertibile, e $I = I^{-1}$.*
- 4) *Se A e B sono invertibili, allora anche $A \cdot B$ e' invertibile e $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.*
- 5) *Se A_1, A_2, \dots, A_h sono matrici invertibili, lo e' anche la matrice prodotto $A_1 \cdot A_2 \cdot \dots \cdot A_h$, e $(A_1 \cdot A_2 \cdot \dots \cdot A_h)^{-1} = (A_h)^{-1} \cdot (A_{h-1})^{-1} \cdot \dots \cdot (A_2)^{-1} \cdot (A_1)^{-1}$.*

Dimostrazione. Le proprieta' 2) e 3) sono ovvie, e la 5) segue dalla 4) per induzione. Quindi proveremo solo 1) e 4).

1) Siano B e C matrici inverse per A . Allora $AB = BA = I$ e $AC = CA = I$. Utilizzando tali formule e la proprieta' associativa, abbiamo:

$$B = BI = B(AC) = (BA)C = IC = C.$$

4) Semplicemente andiamo a verificare che $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$. Infatti:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

E similmente si prova che $(B^{-1}A^{-1})(AB) = I$. ■

Osservazione. Nella definizione di matrice inversa si impone la condizione $AA^{-1} = A^{-1}A$, che, a priori, non e' scontata, visto che sappiamo che in generale la proprieta' commutativa e' falsa. Tuttavia proveremo piu' in la' che in realta' la condizione e' sovrabbondante. Infatti dimostreremo che, data una matrice A , se esiste una matrice B tale che $AB = I$ allora A e' invertibile e B e' l'inversa di A (e percio' necessariamente $BA = I$). Cio' vale anche se esiste una matrice B tale che $BA = I$.

Esempio. Per le matrici di ordine 2 e' molto facile scrivere l'inversa di una matrice. Infatti se

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{e} \quad ad - bc \neq 0$$

allora A e' invertibile e la sua inversa e' la matrice:

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

Come vedremo dopo, per le matrici di ordine $n > 2$ esistono formule esplicite analoghe, ma piu' complicate.

Concludiamo il paragrafo enunciando le proprieta' della matrice trasposta.

• *Le proprieta' generali della matrice trasposta.*

1) $I^T = I$.

2) Se $A, B \in \mathcal{M}(m, n)$, allora $(A + B)^T = A^T + B^T$.

3) Se $A \in \mathcal{M}(m, n)$ e $c \in \mathbf{R}$, allora $(c \cdot A)^T = c \cdot (A^T)$.

4) Se $A \in \mathcal{M}(m, n)$, allora $(A^T)^T = A$.

5) Se $A \in \mathcal{M}(m, n)$ e $B \in \mathcal{M}(n, p)$, allora $(A \cdot B)^T = B^T \cdot A^T$.

6) Se $A \in \mathcal{M}(n, n)$ ed e' invertibile, allora anche A^T e' invertibile, e l'inversa di A^T coincide con la trasposta dell'inversa di A . Cioe': $(A^T)^{-1} = (A^{-1})^T$.

Dimostrazione. 1) E' evidente. Piu' in generale, ogni matrice diagonale e' simmetrica.

2) e 3) Poniamo $A = (a_{ij})$, e $B = (b_{ij})$, e sia $c \in \mathbf{R}$ uno scalare. Allora:

$$(A + B)^T = ((a_{ij}) + (b_{ij}))^T = (a_{ij} + b_{ij})^T = (a_{ji} + b_{ji}) = (a_{ji}) + (b_{ji}) = A^T + B^T.$$

Inoltre:

$$(c \cdot A)^T = (ca_{ij})^T = (ca_{ji}) = c \cdot (A^T).$$

4) Sia $A = (a_{ij})$. Allora:

$$(A^T)^T = (a_{ji})^T = (a_{ij}) = A.$$

5) Poniamo $A = (a_{ij})$, e $B = (b_{hk})$. Cominciamo con l'osservare che $B^T \in \mathcal{M}(p, n)$, e che $A^T \in \mathcal{M}(n, m)$. Perciò possiamo eseguire la moltiplicazione $B^T \cdot A^T$, che è una matrice in $\mathcal{M}(p, m)$. Anche $(A \cdot B)^T \in \mathcal{M}(p, m)$. Quindi possiamo confrontare $B^T \cdot A^T$ con $(A \cdot B)^T$. Ora denotiamo con $((A \cdot B)^T)_{ki}$ la componente di posto ki di $(A \cdot B)^T$, e con $(B^T \cdot A^T)_{ki}$ quella di $B^T \cdot A^T$. Con conseguente significato per i simboli, si ha:

$$((A \cdot B)^T)_{ki} = (A \cdot B)_{ik} = A_i \cdot B^k.$$

D'altra parte

$$(B^T \cdot A^T)_{ki} = (B^T)_k \cdot (A^T)_i = B^k \cdot A_i = A_i \cdot B^k$$

(l'ultimo passaggio segue dalla simmetria del prodotto punto). Deduciamo che le due matrici $(A \cdot B)^T$ e $B^T \cdot A^T$ hanno, ordinatamente, le stesse componenti. Quindi sono uguali.

6) Andiamo a verificare che la matrice $(A^{-1})^T$ soddisfa le condizioni richieste per l'inversa di A^T . Infatti, tenuto conto della proprietà 4) appena dimostrata, abbiamo:

$$(A^{-1})^T \cdot A^T = (A \cdot A^{-1})^T = I^T = I.$$

E similmente:

$$A^T \cdot (A^{-1})^T = (A^{-1} \cdot A)^T = I^T = I. \quad \blacksquare$$

3. Il rango di una matrice.

Sia A una matrice, con m righe $A_1, \dots, A_m \in \mathbf{R}^n$ ed n colonne $A^1, \dots, A^n \in \mathbf{R}^m$. Sia \mathcal{R}_A il sottospazio di \mathbf{R}^n generato dalle righe di A :

$$\mathcal{R}_A := \text{Span}(A_1, \dots, A_m) \subseteq \mathbf{R}^n.$$

Diremo che \mathcal{R}_A è lo spazio delle righe di A . Accanto allo spazio delle righe possiamo considerare anche lo spazio delle colonne \mathcal{C}_A di A , che per definizione è il sottospazio di \mathbf{R}^m generato dalle colonne di A :

$$\mathcal{C}_A := \text{Span}(A^1, \dots, A^n) \subseteq \mathbf{R}^m.$$

Questi due spazi vettoriali, nonostante risiedano in spazi diversi, cioè \mathbf{R}^n ed \mathbf{R}^m , hanno la stessa dimensione. Ciò è il contenuto del cosiddetto:

- *Teorema del rango:* $\dim \mathcal{R}_A = \dim \mathcal{C}_A$.

Dimostrazione del Teorema del rango. Sia $A = (a_{ij})$ una matrice $m \times n$. Sia μ la dimensione dello spazio delle righe di A , e ν la dimensione dello spazio delle colonne. Vogliamo provare che $\mu = \nu$. A tale proposito, fissiamo una base per lo spazio delle colonne di A , che sia costituita da colonne di A :

$$A^{j_1}, \dots, A^{j_\nu}.$$

Ogni colonna di A è una combinazione lineare dei vettori di tale base, e perciò per ogni $j = 1, \dots, n$ ed $l = 1, \dots, \nu$ esistono scalari k_{jl} tali che

$$A^j = k_{j1}A^{j_1} + k_{j2}A^{j_2} + \dots + k_{j\nu}A^{j_\nu}.$$

Esplicitando sulle componenti abbiamo:

$$a_{ij} = k_{j1}a_{ij_1} + k_{j2}a_{ij_2} + \dots + k_{j\nu}a_{ij_\nu}$$

per ogni $i = 1, \dots, m$ e $j = 1, \dots, n$. Da queste formule deduciamo

$$A_i = a_{ij_1}\mathbf{k}^1 + a_{ij_2}\mathbf{k}^2 + \dots + a_{ij_\nu}\mathbf{k}^\nu,$$

dove con \mathbf{k}^l denotiamo il vettore numerico di \mathbf{R}^n con componenti $\mathbf{k}^l = (k_{1l}, k_{2l}, \dots, k_{nl})$. Ciò implica che lo spazio delle righe è contenuto nel sottospazio di \mathbf{R}^n generato dai ν vettori $\mathbf{k}^1, \mathbf{k}^2, \dots, \mathbf{k}^\nu$. Quindi la dimensione dello spazio delle righe di A , cioè μ , è al massimo ν :

$$\mu \leq \nu.$$

Lo stesso argomento applicato alla matrice A^T proverà che la dimensione dello spazio delle righe di A^T , cioè ν , è minore o uguale della dimensione dello spazio delle colonne di A^T , cioè μ . Quindi $\mu = \nu$. ■

- *La definizione di rango di una matrice.*

Il rango di una matrice A è, per definizione, la dimensione dello spazio delle righe di A , o, ciò che è lo stesso, la dimensione dello spazio delle colonne di A . Denoteremo il rango di A con il simbolo rkA . Perciò:

$$rkA = \dim \mathcal{R}_A = \dim \mathcal{C}_A.$$

Osservazione. (i) Si osservi che per una data matrice A si ha:

$$\mathcal{C}_A = \mathcal{R}_{A^T}.$$

Perciò lo studio dello spazio delle colonne può sempre essere ricondotto allo studio dello spazio delle righe.

(ii) Saper calcolare il rango di una matrice è equivalente a saper calcolare la dimensione di un sottospazio W di \mathbf{R}^n , dati i generatori di W . Infatti, se $\mathbf{b}_1, \dots, \mathbf{b}_h$ generano

W , allora la dimensione di W coincide con il rango della matrice che si ottiene mettendo in riga i vettori $\mathbf{b}_1, \dots, \mathbf{b}_h$.

(iii) Il problema del calcolo del rango di una matrice A ha grande importanza, anche nelle applicazioni in Fisica ed Ingegneria. Ci sono diversi metodi per effettuare tale calcolo. Uno lo abbiamo visto nel capitolo precedente, consiste nel calcolare le relazioni tra le righe (o le colonne), ed, in base a tali relazioni, scartare le eventuali righe sovrabbondanti, fino a pervenire ad un sistema di generatori libero. Il numero delle righe che così si ottengono è il rango della matrice. Ci sono metodi diversi, ed uno particolarmente rapido si basa sulle *operazioni elementari* che andremo a studiare nel prossimo paragrafo. Un altro metodo che si basa sul *determinante* lo vedremo dopo. Osserviamo infine che questi metodi per il calcolo del rango, cioè per il calcolo della dimensione di un sottospazio dello spazio numerico \mathbf{R}^n , si possono applicare per il calcolo della dimensione di un sottospazio di uno spazio vettoriale *qualunque*, grazie all'*applicazione delle coordinate* che consente di identificare uno spazio vettoriale qualsiasi ad uno spazio numerico. Su quest'ultimo punto, cioè su cosa è l'applicazione delle coordinate, e come si usa per il calcolo della dimensione, rinviando al Capitolo 4 sulle applicazioni lineari.

• *La nozione di sottomatrice.*

Sia A una matrice. Una matrice M si dice *sottomatrice di A* se si ottiene da A cancellandone qualche riga, oppure qualche colonna, oppure qualche riga e qualche colonna. Per esempio, se

$$A = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 6 & 9 & 12 \\ -1 & 0 & 4 \end{bmatrix},$$

allora le seguenti matrici sono sottomatrici di A :

$$M = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 6 & 9 & 12 \end{bmatrix}, \quad N = \begin{bmatrix} 4 & 5 \\ 1 & 2 \\ 6 & 9 \\ -1 & 0 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 3 \\ 9 & 12 \\ 0 & 4 \end{bmatrix}.$$

Infatti M si ottiene da A cancellando l'ultima riga, N l'ultima colonna, P si ottiene da A cancellandone la prima riga e la prima colonna. Invece la matrice

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 & 9 & 12 \end{bmatrix}$$

non è una sottomatrice di A .

Osservazione. Se M è una sottomatrice di A allora $rkM \leq rkA$. Per provare ciò, supponiamo che M si ottenga da A cancellando certe righe e certe colonne, e sia N la sottomatrice di A ottenuta cancellando solo le righe che abbiamo soppresso per formare M . Poiché lo spazio delle righe di N è un sottospazio di quello di A , è chiaro che il rango di N non supera il rango di A . D'altra parte M è una sottomatrice di N che si

ottiene da N cancellandone alcune colonne. Quindi lo spazio delle colonne di M è un sottospazio delle colonne di N , e quindi il rango di M è minore o uguale a quello di N . In conclusione: $\text{rk}M \leq \text{rk}N \leq \text{rk}A$.

4. Il calcolo del rango tramite le operazioni elementari.

In questo paragrafo impareremo un algoritmo, detto *l'algoritmo di Gauss*, che ci consente di calcolare il rango di una matrice, in modo più veloce rispetto al calcolo delle relazioni che abbiamo studiato nel capitolo precedente. Per poter enunciare tale algoritmo abbiamo bisogno di una serie di definizioni preliminari.

- *Le operazioni elementari su una matrice.*

Sia A una matrice qualsiasi di tipo $m \times n$. Sulle righe di A si possono eseguire le cosiddette *operazioni elementari sulle righe*. Tali operazioni sono di tre tipi, e sono definite come segue.

- L'operazione p_{ij} , che significa scambiare la riga A_i di A con la riga A_j .
- L'operazione $e_{ij}(k)$, che significa sommare alla riga A_i di A la riga A_j moltiplicata per lo scalare k (in questa operazione si assume $i \neq j$). In altre parole, alla riga A_i si sostituisce la riga $A_i + kA_j$.
- L'operazione $e_i(k)$, che significa moltiplicare la riga A_i per lo scalare k (in questa operazione si assume $k \neq 0$).

Esempio. Se partiamo dalla matrice

$$A = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 6 & 9 & 12 \end{bmatrix}$$

ed eseguiamo, nell'ordine, le operazioni p_{12} , $e_{21}(-4)$, $e_{31}(-6)$, $e_{32}(-1)$, $e_2(-\frac{1}{3})$, arriveremo alla matrice

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

Infatti:

$$\begin{aligned} A = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 6 & 9 & 12 \end{bmatrix} &\xrightarrow{p_{12}} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 & 9 & 12 \end{bmatrix} \xrightarrow{e_{21}(-4)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 6 & 9 & 12 \end{bmatrix} \\ &\xrightarrow{e_{31}(-6)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -3 & -6 \end{bmatrix} \xrightarrow{e_{32}(-1)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{e_2(-\frac{1}{3})} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = S. \end{aligned}$$

Osservazione. (i) Nell'esempio precedente, se rappresentiamo la matrice A con le sue righe al seguente modo:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix},$$

allora possiamo rappresentare l'effetto delle operazioni elementari sulle righe di A così:

$$\begin{aligned} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} &\xrightarrow{p_{12}} \begin{bmatrix} A_2 \\ A_1 \\ A_3 \end{bmatrix} \xrightarrow{e_{21}(-4)} \begin{bmatrix} A_2 \\ A_1 - 4A_2 \\ A_3 \end{bmatrix} \xrightarrow{e_{31}(-6)} \begin{bmatrix} A_2 \\ A_1 - 4A_2 \\ A_3 - 6A_2 \end{bmatrix} \\ &\xrightarrow{e_{32}(-1)} \begin{bmatrix} A_2 \\ A_1 - 4A_2 \\ -A_1 - 2A_2 + A_3 \end{bmatrix} \xrightarrow{e_2(-\frac{1}{3})} \begin{bmatrix} A_2 \\ -\frac{1}{3}A_1 + \frac{4}{3}A_2 \\ -A_1 - 2A_2 + A_3 \end{bmatrix} = S. \end{aligned}$$

Poiché l'ultima riga di S è nulla allora deduciamo che

$$A_1 + 2A_2 - A_3 = \mathbf{0}.$$

Possiamo interpretare ciò al seguente modo. Il fatto che in S appaia una riga nulla implica la presenza di una relazione non banale per le righe di A . Come diremo in seguito, questo è un fatto di carattere generale. In più possiamo dire che, poiché la riga nulla di S appare al terzo posto, allora è sovrabbondante la terza riga di A . Anche questo è un fatto di carattere generale, purché si tenga conto che la terza riga di A non è stata sottoposta ad operazioni di scambio. In altre parole, per individuare la riga sovrabbondante in A , occorre tener conto degli scambi effettuati. Per esempio se sottoponiamo alla matrice

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

le operazioni $e_{21}(-2)$ e poi p_{23} otteniamo la matrice

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ma è chiaro che la riga sovrabbondante in A certo non è la terza, bensì la seconda.

(ii) Accanto alle operazioni elementari sulle righe si possono considerare, in modo analogo, anche le *operazioni elementari sulle colonne*. Sono di tre tipi, cioè p^{ij} , $e^{ij}(k)$ ed $e^i(k)$. Per le cose che diremo in questo capitolo ci sarà sufficiente considerare solo le operazioni sulle righe, ma, come sarà chiaro in seguito, si può sviluppare una teoria del tutto simile per le operazioni sulle colonne.

- *Le matrici elementari.*

Le operazioni elementari si possono interpretare in termini di prodotto riga per colonna, nel senso che ora andiamo a vedere. Premettiamo la definizione di *matrice elementare*. Una matrice elementare e', per definizione, una matrice quadrata che si ottiene dalla matrice I tramite una sola operazione sulle righe. Percio' avremo tre tipi di matrici elementari, che denoteremo, con ovvio significato per la notazione, con P_{ij} , $E_{ij}(k)$, $E_i(k)$. Per esempio, le seguenti matrici sono tutte e sole le matrici elementari di ordine $n = 2$.

$$P_{12} = P_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_{21}(k) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \quad E_{12}(k) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix},$$

$$E_1(k) = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}, \quad E_2(k) = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}.$$

Il nesso tra le operazioni elementari e le matrici elementari consiste nel fatto che *eseguire un'operazione elementare "e" sulle righe di una matrice A equivale a moltiplicare a sinistra di A per la matrice elementare E corrispondente ad "e", cioe'*

$$A \xrightarrow{e} EA.$$

Osservazione. (i) Osserviamo anche che le matrici elementari sono matrici invertibili, ed evidentemente si ha:

$$P_{ij}^{-1} = P_{ij}, \quad E_{ij}(k)^{-1} = E_{ij}(-k), \quad E_i(k)^{-1} = E_i\left(\frac{1}{k}\right).$$

In particolare la matrice inversa di una matrice elementare e' ancora una matrice elementare.

(ii) Anche la matrice trasposta di una matrice elementare e' elementare. Infatti:

$$P_{ij}^T = P_{ij}, \quad E_{ij}(k)^T = E_{ji}(k), \quad E_i(k)^T = E_i(k).$$

(iii) Accanto alle matrici elementari P_{ij} , $E_{ij}(k)$, $E_i(k)$ possiamo anche definire le matrici P^{ij} , $E^{ij}(k)$, $E^i(k)$ ottenute dalla matrice I tramite una sola operazione elementare sulle *colonne*. In tal caso pero' non otteniamo una nuova famiglia di matrici in quanto:

$$P^{ij} = P_{ij}^T = P_{ij}, \quad E^{ij}(k) = E_{ij}(k)^T = E_{ji}(k), \quad E^i(k) = E_i(k)^T = E_i(k).$$

E, in analogia con il caso delle operazioni elementari sulle righe, si ha che *eseguire un'operazione elementare "e" sulle colonne di una matrice A equivale a moltiplicare a destra di A per la matrice elementare P^{ij} , $E^{ij}(k)$, $E^i(k)$ corrispondente ad "e", cioe'*

$$A \xrightarrow{e} AE.$$

- *Matrici equivalenti per righe.*

Siano A e B due matrici $m \times n$. Diremo che A è *equivalente per righe* a B , e scriveremo $A \sim_r B$, se B si ottiene da A tramite un numero finito di operazioni elementari sulle righe di A . Per quanto detto in precedenza, $A \sim_r B$ se e solo se esistono un numero finito di matrici elementari E_1, E_2, \dots, E_h tali che $B = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot A$.

Osservazione. (i) La relazione di equivalenza per righe è una relazione di equivalenza nell'insieme delle matrici $\mathcal{M}(m, n)$. Infatti per ogni matrice A possiamo scrivere $A = E_1(1)A$, quindi $A \sim_r A$ e perciò la relazione è riflessiva. Se $A \sim_r B$ allora per opportune matrici elementari E_1, E_2, \dots, E_h abbiamo $B = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot A$. Da cui $A = E_h^{-1} \cdot E_{h-1}^{-1} \cdot \dots \cdot E_1^{-1} \cdot B$, e ciò implica che $B \sim_r A$ in quanto matrici inverse di matrici elementari sono ancora matrici elementari. Quindi la relazione è simmetrica. È anche transitiva, perché se $A \sim_r B$ e $B \sim_r C$, allora, per opportune matrici elementari $E_1, E_2, \dots, E_h, F_1, F_2, \dots, F_k$ avremo $B = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot A$ e $C = F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot B$. Da cui $C = F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot B = F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot A$. Ciò mostra che C si ottiene da A tramite operazioni elementari sulle righe, e perciò $A \sim_r C$.

(ii) Se A è equivalente per righe a B allora lo spazio delle righe di A coincide con lo spazio delle righe di B , cioè:

$$A \sim_r B \implies \mathcal{R}_A = \mathcal{R}_B.$$

Infatti, supponiamo che $A \sim_r B$. Allora esistono opportune operazioni elementari che portano A in B . Per la definizione stessa di operazione elementare, ciò implica che le righe di B sono combinazioni lineari delle righe di A , per cui $\mathcal{R}_B \subseteq \mathcal{R}_A$. D'altra parte, abbiamo appena visto che se $A \sim_r B$ allora $B \sim_r A$, e quindi per lo stesso motivo di prima deve valere anche l'altra inclusione $\mathcal{R}_A \subseteq \mathcal{R}_B$. Ne consegue $\mathcal{R}_B = \mathcal{R}_A$.

(iii) Per analogia possiamo anche dare la nozione di *equivalenza per colonne*. Diremo che A è *equivalente per colonne* a B , e scriveremo $A \sim_c B$, se B si ottiene da A tramite un numero finito di operazioni elementari sulle colonne di A . Ciò equivale a dire che esistono un numero finito di matrici elementari E_1, E_2, \dots, E_h tali che $B = A \cdot E_1 \cdot E_2 \cdot \dots \cdot E_h$. Anche l'equivalenza per colonne è una relazione di equivalenza, che in generale è diversa dall'equivalenza per righe. Per esempio, le matrici

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

sono equivalenti per righe, ma non per colonne.

(iv) Come nel caso delle operazioni per righe, le operazioni per colonne non alterano lo spazio delle colonne, cioè

$$A \sim_c B \implies \mathcal{C}_A = \mathcal{C}_B.$$

(v) Si osservi che $A \sim_c B$ se e solo se $A^T \sim_r B^T$. Infatti se $B = A \cdot E_1 \cdot E_2 \cdot \dots \cdot E_h$, allora $B^T = E_h^T \cdot E_{h-1}^T \cdot \dots \cdot E_1^T \cdot A^T$, da cui si evince che $A^T \sim_r B^T$ in quanto le trasposte di matrici elementari sono a loro volta elementari.

- *La nozione di matrice a scala.*

Sia A_i una riga non nulla di una data matrice A . Partendo dalla sinistra, la prima componente non nulla che figura in A_i si chiama il *pivot di A_i* . Similmente, se A^j è una colonna non nulla di A , partendo dall'alto, la prima componente non nulla che figura in A^j si chiama il *pivot di A^j* .

Una matrice A si dice che è *a scala per righe* se soddisfa le seguenti due condizioni:

- (1) Le eventuali righe nulle di A si trovano in fondo alla matrice.
- (2) Per ogni riga non nulla A_i di A , detto a_{ij} il suo pivot, sono nulle tutte le componenti di A che si trovano al di sotto di a_{ij} , nella stessa colonna e nelle colonne precedenti.

Esempio. La matrice S dell'esempio precedente è una matrice a scala per righe:

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

Mentre la matrice A da cui parte l'esempio citato non è a scala per righe:

$$A = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 6 & 9 & 12 \end{bmatrix}.$$

Pero' abbiamo visto che con opportune operazioni elementari sulle righe, possiamo ricondurre A ad S , cioè ad una matrice a scala per righe. Questo è un fatto di carattere generale che andremo a discutere tra poco.

Osservazione. Si può anche dare la nozione di *matrice a scala per colonne*. Una matrice A si dice a scala per colonne se le eventuali colonne nulle di A si trovano in fondo a destra della matrice, e se, per ogni colonna non nulla A^j di A , detto a_{ij} il suo pivot, sono nulle tutte le componenti di A che si trovano a destra di a_{ij} , nella stessa riga e nelle righe precedenti. In altre parole A è a scala per colonne se e solo se A^T lo è per righe.

L'interesse delle matrici a scala per righe consiste nei seguenti due fatti:

1) *Ogni matrice A è equivalente per righe ad una matrice a scala S per righe. Anzi, esiste un algoritmo, detto l'algoritmo di Gauss, che fornisce una sequenza automatica di operazioni elementari che portano A in S .*

2) *Le righe non nulle di una matrice a scala per righe sono linearmente indipendenti. Perciò il rango di una matrice a scala per righe si calcola "a vista", contando semplicemente le sue righe non nulle.*

Una immediata conseguenza è l'annunciato metodo per il calcolo del rango di una matrice qualsiasi tramite le operazioni elementari.

- *Il calcolo del rango tramite le operazioni elementari.*

Sia A una matrice qualsiasi. Sia S una matrice a scala per righe equivalente per righe ad A , a cui si perviene per esempio tramite l'algoritmo di Gauss. Poiche' A e' equivalente per righe ad S , allora lo spazio delle righe di A coincide con lo spazio delle righe di S . D'altra parte le righe non nulle di S sono linearmente indipendenti, e percio' formano una base per $\mathcal{R}_A (= \mathcal{R}_S)$, ed il rango di A coincide con il numero di righe non nulle di S . Inoltre le righe nulle di S corrispondono alle righe sovrabbondanti nel sistema delle righe di A , tenuto conto degli scambi effettuati.

Per dimostrare la proprieta' 1) esibiamo direttamente l'algoritmo di Gauss. Tale algoritmo prevede in ingresso una qualunque matrice A , e produce in uscita come risultato una matrice a scala per righe S , equivalente per righe alla matrice di partenza A :

$$A \xrightarrow{\text{algoritmo di Gauss}} S.$$

L'algoritmo funziona cosi'.

Passo 1.

1.1 Se la matrice A e' formata da una sola riga, l'algoritmo termina; altrimenti:

1.2 individuare la colonna non nulla con indice piu' basso, ed il suo pivot, cioe' la sua prima componente non nulla; se non esistono colonne non nulle, la matrice e' nulla, quindi e' gia' a scala: in questo caso l'algoritmo termina qui.

1.3 Se il pivot e' nella riga di posto i , scambiare la prima riga con quella di posto i (qui stiamo usando operazioni tipo p_{ij}).

1.4 Rendere nulle tutte le altre componenti della colonna che contiene il pivot della prima riga, sommando alle varie righe opportuni multipli della prima (qui stiamo usando operazioni tipo $e_{ij}(k)$).

Passo 2.

Ripetere il Passo 1 sulla matrice ottenuta dal passo precedente, schermandone pero' la prima riga.

Passo 3.

Ripetere il Passo 2 sulla matrice schermata, fino ad esaurimento delle righe.

Cio' e' quanto basta per provare la proprieta' 1).

Per la proprieta' 2) ci limitiamo ad un breve commento. Se S e' a scala per righe con h righe non nulle, ed andiamo alla ricerca delle relazioni (a_1, \dots, a_h) fra le sue righe non nulle, cosi' come abbiamo visto negli esempi nel Capitolo sugli spazi vettoriali, si perviene ad un sistema di equazioni di primo grado con incognite date dalle componenti della generica relazione (a_1, \dots, a_h) . Ora, proprio perche' S e' a scala per righe, con un calcolo immediato si vede che l'unica possibilita' e' $a_1 = \dots = a_h = 0$. Fra poco faremo un esempio.

Osservazione. (i) Nell'algoritmo di Gauss non intervengono operazioni elementari del tipo $e_i(k)$. Queste operazioni interverranno in un altro algoritmo, detto *algoritmo di Gauss-Jordan*, che studieremo dopo, per il calcolo della matrice inversa.

(ii) Naturalmente possiamo enunciare anche un algoritmo di Gauss per lo spazio delle colonne. In particolare ogni matrice è equivalente per colonne ad una matrice a scala per colonne, ed il rango di una matrice a scala per colonne coincide con il numero delle sue colonne non nulle. Ricordiamo, come già osservato, che lo studio dello spazio delle colonne di una matrice A si riconduce a quello delle righe grazie al passaggio alla matrice trasposta A^T .

(iii) Durante il processo di riduzione a scala di una matrice, può accadere che sia possibile fare qualche operazione elementare che renda più rapido il calcolo rispetto all'algoritmo di Gauss. Per esempio, può accadere che una colonna abbia una componente uguale ad 1, che non sia il pivot. In tal caso conviene portare in alto questa componente uguale ad 1, invece del pivot, se si ha intenzione di evitare l'uso delle frazioni.

Esempio. Consideriamo la seguente matrice

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 3 & 6 & 0 & 9 \\ -1 & 0 & 1 & -1 \end{bmatrix}.$$

Andiamo ad applicare l'algoritmo di Gauss ad A . Innanzitutto partendo dalla sinistra individuiamo la prima colonna non nulla, in questo caso la prima colonna di A , ed il suo pivot, che è $a_{21} = 3$. Tale pivot si trova nella seconda riga, perciò occorre scambiare la prima riga con la seconda in modo da portare il pivot in alto, cioè nella prima riga.

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 3 & 6 & 0 & 9 \\ -1 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{p_{12}} \begin{bmatrix} 3 & 6 & 0 & 9 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & -1 \end{bmatrix}.$$

Ora con l'operazione $e_{31}(\frac{1}{3})$ rendiamo nulla l'unica componente diversa da 0 posta sotto il pivot della prima colonna.

$$\begin{bmatrix} 3 & 6 & 0 & 9 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{e_{31}(\frac{1}{3})} \begin{bmatrix} 3 & 6 & 0 & 9 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 1 & 2 \end{bmatrix}.$$

Ora che abbiamo annullato tutte le componenti della prima colonna poste al di sotto del pivot, ci dimentichiamo della prima riga, e ripetiamo la procedura sulla matrice formata solo dalle ultime due righe. Per tale matrice, la prima colonna da sinistra non nulla è la seconda, il cui pivot è 1 ed è posto in alto. Perciò possiamo operare con $e_{32}(-2)$ ed ottenere

$$\begin{bmatrix} 3 & 6 & 0 & 9 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 1 & 2 \end{bmatrix} \xrightarrow{e_{32}(-2)} \begin{bmatrix} 3 & 6 & 0 & 9 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix} =: S.$$

L'algoritmo finisce qui, perché la matrice S a cui siamo pervenuti è a scala. Andiamo a verificare che le tre righe non nulle di S sono libere. Infatti se (x, y, z) è una relazione tra le righe di S allora:

$$x(3, 6, 0, 9) + y(0, 1, 1, 0) + z(0, 0, -1, 2) = (0, 0, 0, 0).$$

Cio' equivale a dire che

$$\begin{cases} 3x = 0 \\ 6x + y = 0 \\ y - z = 0 \\ 9x + 2z = 0 \end{cases}$$

da cui deduciamo subito, come avevamo detto, che $x = y = z = 0$.

Esercizio. Sia W il sottospazio di \mathbf{R}^4 generato dai vettori $(1, 1, 1, 1)$, $(1, 1, 1, 2)$, $(0, 0, 0, 1)$, $(3, 3, 3, 4)$. Determinare una base di W .

Svolgimento. Lo spazio W coincide con lo spazio delle righe della matrice

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 3 & 3 & 3 & 4 \end{bmatrix}$$

ottenuta disponendo i generatori di W in riga. Riducendo a scala per righe la matrice A otteniamo:

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Percio' una base per W e' data dai vettori $(1, 1, 1, 1)$, $(0, 0, 0, 1)$. ■

Esercizio. Provare che il vettore $(2, 2, -4)$ sta in $\text{Span}((1, -1, 0), (0, 1, -1))$.

Svolgimento. Disponendo i vettori in riga formiamo una matrice

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

Riducendo a scala tale matrice si riesce, senza operazioni di scambio, a pervenire alla matrice

$$S = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Cio' implica che la terza riga della matrice A e' sovrabbondante, cioe' che $(2, 2, -4) \in \text{Span}((1, -1, 0), (0, 1, -1))$. ■

Esercizio. Sia U il sottospazio di \mathbf{R}^4 generato dai vettori $(1, 1, 1, 1)$, $(2, 0, 1, 3)$, $(3, 1, 2, 4)$, e sia V il sottospazio generato dai vettori $(1, 0, 0, 0)$, $(0, 1, 0, -2)$. Determinare una base per $U + V$.

Svolgimento. Disponendo in riga i vettori assegnati otteniamo una matrice

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

il cui spazio delle righe coincide con $U + V$. Riducendo a scala A otteniamo la matrice

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -1 & 1 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Perciò una base per $U + V$ è formata dai vettori $(1, 1, 1, 1)$, $(0, -2, -1, 1)$, $(0, 0, -\frac{1}{2}, -\frac{3}{2})$. ■

Esercizio. Sia U il sottospazio di \mathbf{R}^4 generato dai vettori $(1, 1, -1, 1)$, $(1, 2, 0, 2)$, $(1, 2, 0, 3)$. Determinare un sottospazio V di \mathbf{R}^4 tale che $\mathbf{R}^4 = U \oplus V$.

Svolgimento. Mettiamo in riga i generatori di U , formando la matrice

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

Sappiamo che $U = \mathcal{R}_A$. Riducendo A a scala per righe, otteniamo la matrice

$$S = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

e $U = \mathcal{R}_S$. Adesso inseriamo tra le righe di S un vettore riga canonico (in questo caso \mathbf{e}_3) in modo da formare una matrice quadrata a scala T di rango massimo 4:

$$T = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Poiché T ha rango 4, le sue righe formano una base di \mathbf{R}^4 . Perciò possiamo porre come spazio cercato $V := \text{Span}(\mathbf{e}_3)$. ■

Esercizio. Al variare del parametro $h \in \mathbf{R}$, sia W_h il sottospazio di \mathbf{R}^4 generato dai vettori $(h, -1, 1, 1)$, $(-1, -1, 1, 3)$, $(-1, -2, 2, 4)$. Sia U il sottospazio di \mathbf{R}^4 generato dai vettori $(1, -2, 2, 0)$, $(-1, 0, 0, 2)$. Per quali valori di h si ha $W_h = U$?

Svolgimento. Poiche' W_h ha dimensione almeno 2, ed U ha dimensione 2, allora $W_h = U$ se e solo se $W_h + U$ ha dimensione 2. Percio' si tratta di determinare per quali valori del parametro h la matrice

$$A_h = \begin{bmatrix} 1 & -2 & 2 & 0 \\ -1 & 0 & 0 & 2 \\ -1 & -1 & 1 & 3 \\ -1 & -2 & 2 & 4 \\ h & -1 & 1 & 1 \end{bmatrix}$$

ha rango 2. D'altra parte riducendo a scala la matrice A_h si ottiene la matrice

$$S_h = \begin{bmatrix} 1 & -2 & 2 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 2h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Percio' $W_h = U$ se e solo se $h = 0$. ■

Esercizio. Al variare del parametro $h \in \mathbf{R}$, sia A_h la matrice

$$A_h = \begin{bmatrix} h & h & h \\ h & h & h \\ h & h & h \end{bmatrix}.$$

Quanto vale il rango di A_h ?

Svolgimento. Il rango di A_h e' 1 se $h \neq 0$, ed e' 0 se $h = 0$. ■

Esercizio. Al variare del parametro $h \in \mathbf{R}$, sia A_h la matrice

$$A_h = \begin{bmatrix} 1 & 1 & 1 \\ 1 & h & 2 \\ 1 & 1 & h-1 \end{bmatrix}.$$

Quanto vale il rango di A_h ?

Svolgimento. Tramite le operazioni elementari $e_{21}(-1)$ ed $e_{31}(-1)$ perveniamo alla matrice

$$B_h = \begin{bmatrix} 1 & 1 & 1 \\ 0 & h-1 & 1 \\ 0 & 0 & h-2 \end{bmatrix}.$$

Se $h \notin \{1, 2\}$ la matrice B_h e' a scala con tre righe non nulle, percio' in tal caso A_h ha rango 3. Anche la matrice B_2 e' a scala, ma ha rango 2. Invece B_1 non e' a scala, ma e' equivalente alla matrice a scala

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Perciò anche B_1 ha rango 2. In conclusione:

$$\operatorname{rk} A_h = \begin{cases} 3 & \text{se } h \notin \{1, 2\} \\ 2 & \text{se } h \in \{1, 2\}. \end{cases} \blacksquare$$

Osservazione. Come abbiamo già detto, le operazioni elementari possono essere utilizzate anche per il calcolo della dimensione di un sottospazio di uno spazio vettoriale qualsiasi, non necessariamente lo spazio numerico \mathbf{R}^n . Ciò si può fare grazie all'applicazione delle coordinate, per la quale rinviamo al Capitolo 4.

Esercizio. Sia A una matrice $m \times n$ e B una matrice $n \times p$. Provare che

$$\operatorname{rk}(A \cdot B) \leq \operatorname{rk} A.$$

Svolgimento. Sia S una matrice a scala per righe equivalente per righe ad A . Quindi possiamo scrivere $A = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot S$, con opportune matrici elementari. Allora si ha $AB = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot S \cdot B$. Quindi $\operatorname{rk}(A \cdot B) = \operatorname{rk}(S \cdot B)$, e poiché $\operatorname{rk} S = \operatorname{rk} A$, sarà sufficiente provare che

$$\operatorname{rk}(S \cdot B) \leq \operatorname{rk} S.$$

Ora la conclusione segue dal fatto che le (eventuali) righe nulle di S producono righe nulle nel prodotto SB . Perciò SB ha almeno tante righe nulle quante ne ha S . Avendo SB ed S lo stesso numero di righe, ne consegue che il numero di righe non nulle di SB (e quindi il rango di SB) è minore o uguale del numero di righe non nulle di S , che è il rango di S . In altri termini:

$$\begin{aligned} \operatorname{rk}(A \cdot B) &= \operatorname{rk}(S \cdot B) \leq \# \text{ righe non nulle di } S \cdot B \\ &= m - (\# \text{ righe nulle di } S \cdot B) \leq m - (\# \text{ righe nulle di } S) = \operatorname{rk} S = \operatorname{rk} A. \end{aligned} \blacksquare$$

Esercizio. Sia A una matrice $m \times n$ e B una matrice $n \times p$. Provare che

$$\operatorname{rk}(A \cdot B) \leq \operatorname{rk} B.$$

Svolgimento. Per il Teorema del rango, sappiamo che

$$\operatorname{rk}(A \cdot B) = \operatorname{rk}(A \cdot B)^T = \operatorname{rk}(B^T \cdot A^T).$$

Per l'esercizio precedente abbiamo anche

$$\operatorname{rk}(B^T \cdot A^T) \leq \operatorname{rk} B^T = \operatorname{rk} B. \blacksquare$$

Esercizio. Sia A una matrice $m \times n$, P una matrice $m \times m$ invertibile, Q una matrice $n \times n$ invertibile. Provare che

$$\operatorname{rk} A = \operatorname{rk}(Q \cdot A) = \operatorname{rk}(A \cdot P).$$

In altre parole, quando moltiplichiamo una matrice per una matrice invertibile, il rango non cambia.

Svolgimento. Per l'esercizio precedente sappiamo che:

$$\text{rk}(Q \cdot A) \leq \text{rk } A.$$

Per lo stesso motivo, abbiamo:

$$\text{rk } A = \text{rk}(Q^{-1} \cdot (Q \cdot A)) \leq \text{rk}(Q \cdot A).$$

L'altra uguaglianza $\text{rk } A = \text{rk}(A \cdot P)$ si dimostra in modo simile, utilizzando gli esercizi precedenti. ■

Esercizio. Sia A una matrice quadrata di ordine n , avente rango massimo $\text{rk } A = n$. Provare che A e' prodotto di matrici elementari.

Svolgimento. Sia S una matrice a scala per righe equivalente per righe ad A . Quindi possiamo scrivere $A = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot S$, con opportune matrici elementari. Poiche' il rango di A e' massimo, ed S e' a scala, tutte le componenti s_{ii} di S sulla diagonale principale sono non nulle. Percio' ciascuna componente s_{ii} e' il pivot della riga a cui appartiene. Quindi, partendo dall'ultima riga, con opportune operazioni elementari sulle righe di S , possiamo annullare tutte le componenti di S poste al di sopra della diagonale principale, e rendere S diagonale. Quindi esistono opportune matrici elementari F_1, F_2, \dots, F_k , tali che $S = F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot D$, dove D e' quella matrice diagonale che ha la stessa diagonale principale di S . Allora possiamo scrivere

$$A = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot D.$$

Infine, con opportune operazioni di tipo $e_i(k)$, possiamo rendere uguali ad 1 tutte le componenti di D sulla diagonale principale, portando D nella matrice unitaria I . Percio', con altre matrici elementari G_j , possiamo far si' che

$$D = G_1 \cdot G_2 \cdot \dots \cdot G_l \cdot I = G_1 \cdot G_2 \cdot \dots \cdot G_l.$$

Sostituendo, riusciamo ad esprimere A come prodotto di matrici elementari:

$$A = E_1 \cdot E_2 \cdot \dots \cdot E_h \cdot F_1 \cdot F_2 \cdot \dots \cdot F_k \cdot G_1 \cdot G_2 \cdot \dots \cdot G_l. \quad \blacksquare$$

Esercizio. Sia A una matrice quadrata di ordine n . Provare che A ha rango massimo $\text{rk } A = n$ se e solo se A e' una matrice invertibile.

Svolgimento. Se A ha rango massimo, abbiamo visto che e' prodotto di matrici elementari, e quindi A e' invertibile perche' prodotto di matrici invertibili.

Viceversa, se A e' invertibile, esiste una matrice B tale che $A \cdot B = I$. Poiche'

$$n = \text{rk } I = \text{rk}(A \cdot B) \leq \text{rk } A \leq n,$$

allora A ha rango massimo n . ■

5. Il determinante di una matrice quadrata.

Ad ogni matrice quadrata A si può associare un numero, $\det A \in \mathbf{R}$, detto *il determinante di A* , anche indicato con il simbolo $|A|$, che ha una grande importanza nello studio dell'algebra lineare. Per esempio, come vedremo in seguito, il determinante induce un nuovo algoritmo per il calcolo del rango, più lento rispetto all'algoritmo di Gauss, ma più interessante per certi aspetti di cui parleremo più in là.

- *La definizione di determinante di una matrice quadrata.*

Sia $A = (a_{ij})$ una matrice quadrata di ordine n .

Se $n = 1$ la matrice è un numero $A = [a]$, ed in tal caso si pone

$$\det A := a.$$

Supponiamo ora $n > 1$. In ogni riga di A scegliamo un elemento, in modo tale che la scelta ricada in colonne tutte diverse. Otteniamo n elementi che possiamo denotare così:

$$a_{1\sigma(1)}, a_{2\sigma(2)}, \dots, a_{n\sigma(n)}$$

dove $a_{i\sigma(i)}$ denota l'elemento che abbiamo scelto nella riga di posto i , e $\sigma(i)$ l'indice della colonna a cui tale elemento appartiene. Ne consegue che l'applicazione

$$\sigma : i \in \{1, 2, \dots, n\} \rightarrow \sigma(i) \in \{1, 2, \dots, n\}$$

è una permutazione sull'insieme $\{1, 2, \dots, n\}$ che identifica la scelta che abbiamo fatto.³ Denoteremo con

$$a_\sigma := a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$$

il prodotto degli elementi che abbiamo scelto. Al variare di tutte le possibili scelte, cioè al variare di σ tra tutte le possibili permutazioni di S_n , avremo calcolato $n!$ prodotti a_σ . Si definisce *il determinante della matrice A* come la somma di tutti questi elementi a_σ così calcolati, ciascuno di essi moltiplicato per la segnatura di σ . Cioè:

$$\det A := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_\sigma.$$

³ Una permutazione sull'insieme $\{1, 2, \dots, n\}$ è un'applicazione biiettiva di $\{1, 2, \dots, n\}$ in se'. L'insieme di tutte le permutazioni di $\{1, 2, \dots, n\}$ si denota S_n , e si chiama il gruppo simmetrico su n elementi. Il gruppo simmetrico S_n ha $n!$ elementi. Se σ è una permutazione sull'insieme $\{1, 2, \dots, n\}$, la coppia (i, j) si dice che è una inversione per σ se $i < j$ e $\sigma(i) > \sigma(j)$. Se σ ha un numero pari di inversioni, si dice che σ ha segnatura pari, ed in tal caso si pone $\operatorname{sgn}(\sigma) := 1$. Se σ ha un numero dispari di inversioni, si dice che σ ha segnatura dispari, ed in tal caso si pone $\operatorname{sgn}(\sigma) := -1$. Per esempio, l'applicazione identica su $\{1, 2, \dots, n\}$ non ha inversioni, perciò ha segnatura pari. Invece la permutazione che scambia due indici diversi, e fissa tutti gli altri, è una permutazione dispari. La segnatura della composizione di due permutazioni coincide con il prodotto delle singole segnature.

Nel caso $n = 2$ la precedente definizione si riduce alla formula:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Ma nel caso $n > 2$ il calcolo del determinante, così come è stato definito, è intrattabile. Tuttavia è possibile rendere il calcolo più veloce grazie alle operazioni elementari. Ciò sarà l'argomento del prossimo paragrafo.

6. Il calcolo del determinante tramite le operazioni elementari.

Occorre premettere alcune proprietà elementari dei determinanti, che si deducono facilmente dalla definizione di determinante.

- *Le proprietà elementari del determinante.*

1) $\det A = \det A^T$.

2) Se la matrice A possiede una riga nulla, oppure una colonna nulla, allora $\det A = 0$.

3) Se A' è la matrice che si ottiene da A moltiplicando per k una riga, oppure una colonna, allora $\det A' = k \det A$.

4) Siano A_1, A_2, \dots, A_n le righe di A , in modo da poter rappresentare A nel seguente modo:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Supponiamo che la prima riga di A sia la somma di h vettori numerici di \mathbf{R}^n :

$$A_1 = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_h.$$

Allora

$$\det \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} \mathbf{u}_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} + \det \begin{bmatrix} \mathbf{u}_2 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} + \dots + \det \begin{bmatrix} \mathbf{u}_h \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Tale proprietà vale anche se formulata rispetto ad una riga qualsiasi, ed anche se formulata rispetto ad una colonna.

5) Se la matrice A' si ottiene da A scambiando due righe, oppure due colonne, allora $\det A' = -\det A$.

6) Se la matrice A possiede due righe uguali, oppure due colonne uguali, allora $\det A = 0$.

7) Se le righe di A , oppure le colonne di A , sono linearmente dipendenti, allora $\det A = 0$.

8) Se ad una riga di A si somma una combinazione lineare delle altre righe, il determinante di A non cambia. Il determinante di A non cambia anche se ad una colonna di A si somma una combinazione lineare delle altre colonne.

9) Se la matrice A è triangolare, superiore o inferiore, allora il determinante di A è semplicemente il prodotto degli elementi della diagonale principale di A . In particolare $\det I = 1$.

Dimostrazione. 1) Poniamo $A = (a_{ij})$, ed $A^T = (a'_{ij})$. Ricordiamo che $a'_{ij} = a_{ji}$. Per ogni permutazione σ abbiamo:

$$\begin{aligned} a'_\sigma &= a'_{1\sigma(1)} \cdot a'_{2\sigma(2)} \cdot \dots \cdot a'_{n\sigma(n)} = a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \dots \cdot a_{\sigma(n)n} \\ &= a_{1\sigma^{-1}(1)} \cdot a_{2\sigma^{-1}(2)} \cdot \dots \cdot a_{n\sigma^{-1}(n)} = a_{\sigma^{-1}}. \end{aligned}$$

Tenuto conto che σ e σ^{-1} hanno la stessa segnatura, e che al variare di σ in S_n , la permutazione inversa σ^{-1} descrive tutto S_n una sola volta,⁴ deduciamo:

$$\det A^T = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a'_\sigma = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \cdot a_{\sigma^{-1}} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_\sigma = \det A.$$

2) Se A ha una riga nulla, o anche una colonna, allora in ogni prodotto a_σ appare uno 0 come fattore, perciò $\det A = 0$.

3) Se A' si ottiene da A moltiplicando per k una riga di A , o anche una colonna, allora per ogni σ si avrà $a'_\sigma = k a_\sigma$, e perciò $\det A' = k \det A$.

4) Per ogni $l = 1, \dots, h$ poniamo $\mathbf{u}_l = (u_{l1}, u_{l2}, \dots, u_{ln})$, e

$$M^l = (m^l_{ij}) = \begin{bmatrix} \mathbf{u}_l \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Per il determinante di M^l useremo la notazione:

$$\det M^l = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot m^l_\sigma.$$

Poiché $A_1 = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_h$, allora

$$a_{1j} = u_{1j} + u_{2j} + \dots + u_{hj}$$

⁴ Cioè l'applicazione $\sigma \in S_n \rightarrow \sigma^{-1} \in S_n$ è biettiva.

per ogni j . Perciò

$$\begin{aligned} a_\sigma &= a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)} \\ &= (u_{1\sigma(1)} + u_{2\sigma(1)} + \dots + u_{h\sigma(1)}) \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)} = \\ &u_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)} + u_{2\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)} + \dots + u_{h\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}. \end{aligned}$$

Questo vuol dire che, per ogni permutazione σ , l'elemento a_σ relativo alla matrice A coincide con la somma dei corrispondenti elementi nelle h matrici M^l a destra dell'uguaglianza. Cioè che:

$$a_\sigma = m_\sigma^1 + m_\sigma^2 + \dots + m_\sigma^h.$$

Passando alla somma degli a_σ , si ottiene l'uguaglianza cercata.

E' chiaro che questo argomento funziona per ogni riga, non necessariamente la prima, così come può essere utilizzato per una colonna che si decomponga nella somma di certi vettori numerici colonna.

5) Per semplicità di notazione, supponiamo che A' si ottenga da A scambiando le prime due righe (l'argomento che segue si applica anche al caso delle colonne). Allora per ogni permutazione σ avremo:

$$\begin{aligned} a'_\sigma &= a'_{1\sigma(1)} \cdot a'_{2\sigma(2)} \cdot a'_{3\sigma(3)} \cdot \dots \cdot a'_{n\sigma(n)} = a_{2\sigma(1)} \cdot a_{1\sigma(2)} \cdot a_{3\sigma(3)} \cdot \dots \cdot a_{n\sigma(n)} \\ &= a_{1\sigma(2)} \cdot a_{2\sigma(1)} \cdot a_{3\sigma(3)} \cdot \dots \cdot a_{n\sigma(n)}. \end{aligned}$$

Cioè a'_σ coincide con $a_{\sigma \circ \tau}$, dove con τ denotiamo la permutazione che scambia 1 con 2. Tenuto conto che al variare di σ in S_n la permutazione $\sigma \circ \tau$ descrive tutto S_n una sola volta, e che τ ha segnatura dispari, ne consegue:

$$\begin{aligned} \det A' &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a'_\sigma = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma \circ \tau} = - \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma \circ \tau) \cdot a_{\sigma \circ \tau} \\ &= - \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_\sigma = - \det A. \end{aligned}$$

6) Se A possiede due righe uguali (o anche due colonne uguali), allora scambiando tra loro tali righe la matrice non cambia, ma, per la proprietà precedente, cambia il segno del determinante. Quindi $\det A = -\det A$, da cui $\det A = 0$.

7) Per semplificare la notazione, supponiamo che la prima riga dipenda linearmente dalle rimanenti, quindi si abbia un'espressione del tipo:

$$A_1 = k_2 A_2 + k_3 A_3 + \dots + k_n A_n.$$

Combinando la proprietà 4) con la proprietà 3), allora possiamo esprimere il determinante di A come una combinazione lineare, con pesi i numeri k_j , di determinanti di matrici che presentano due righe uguali. Quindi $\det A = 0$.

8) Sia A' la matrice che si ottiene da A sostituendo alla riga A_i la riga:

$$A_i + k_1 A_1 + \cdots + k_{i-1} A_{i-1} + k_{i+1} A_{i+1} + \cdots + k_n A_n$$

che si ottiene sommando ad A_i una combinazione lineare delle rimanenti righe di A . Per la proprietà 4) il determinante di A' coincide con la somma del determinante di A con il determinante di una certa matrice B che ha la riga di posto i :

$$k_1 A_1 + \cdots + k_{i-1} A_{i-1} + k_{i+1} A_{i+1} + \cdots + k_n A_n$$

combinazione lineare delle sue rimanenti righe. Per la proprietà precedente $\det B = 0$, e perciò $\det A' = \det A$.

9) Se la matrice A è triangolare, c'è solo un termine a_σ che può essere non nullo, ed è proprio quello con σ pari all'identità di $\{1, 2, \dots, n\}$. Cioè

$$\det A = a_{\text{id}_{\{1,2,\dots,n\}}} = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}. \quad \blacksquare$$

Osservazione. Se pensiamo ad una matrice quadrata A di ordine n come ad una n -pla (A_1, A_2, \dots, A_n) di vettori numerici di \mathbf{R}^n , cioè la n -pla delle sue righe, possiamo pensare il determinante come una funzione

$$\det : \mathbf{R}^n \times \mathbf{R}^n \times \cdots \times \mathbf{R}^n \rightarrow \mathbf{R}$$

che ad ogni matrice $A = (A_1, A_2, \dots, A_n)$ associa il suo determinante $\det A \in \mathbf{R}$. In tal caso le proprietà 3), 4) e 5) che abbiamo appena dimostrato, si esprimono dicendo che la funzione determinante è *una funzione multilineare alternante*. Si può dimostrare che la funzione determinante è *l'unica* funzione multilineare alternante che soddisfi la condizione $\det I = 1$.

Siamo in condizione di effettuare il calcolo del determinante tramite le operazioni elementari.

- *Il calcolo del determinante tramite le operazioni elementari.*

Dalle proprietà del determinante che abbiamo appena provato deduciamo, in particolare, che:

1. Se sulla matrice A si effettua un'operazione di scambio tipo p_{ij} , si ottiene una matrice A' tale che $\det A' = -\det A$.
2. Se sulla matrice A si effettua un'operazione del tipo $e_{ij}(k)$, si ottiene una matrice A' che ha lo stesso determinante di A .
3. Se sulla matrice A si effettua un'operazione del tipo $e_i(k)$ si ottiene una matrice A' tale che $\det A' = k \det A$.

Ora supponiamo di voler calcolare il determinante di una data matrice A di ordine n . Se riduciamo a scala la matrice A soltanto con operazioni tipo p_{ij} ed $e_{ij}(k)$, per esempio

con l'algoritmo di Gauss, otterremo una matrice a scala $S = (s_{ij})$ il cui determinante coincide con quello di A , a parte il segno che dipenderà dal numero di scambi effettuati:

$$\det A = (-1)^{\#\text{scambi}} \cdot \det S.$$

D'altra parte, essendo S una matrice a scala, dunque triangolare, il suo determinante è semplicemente il prodotto delle sue componenti sulla diagonale principale, e quindi arriviamo alla formula:

$$\det A = (-1)^{\#\text{scambi}} \cdot s_{11} \cdot s_{22} \cdot \dots \cdot s_{nn}.$$

In conclusione: *data una matrice quadrata A , ed S una matrice a scala per righe equivalente per righe ad A , che si ottenga da A soltanto con operazioni del tipo p_{ij} ed $e_{ij}(k)$, allora il determinante di A coincide con il prodotto degli elementi sulla diagonale principale di S , preso con lo stesso segno o con segno opposto a seconda che il numero di scambi effettuati sia pari o dispari.*

Esempio. Calcoliamo il determinante della seguente matrice:

$$A = \begin{bmatrix} 3 & -2 & 12 & 0 & 7 \\ 1 & 0 & 6 & 8 & 6 \\ 1 & -1 & 3 & -4 & 2 \\ 1 & -1 & 3 & -3 & 3 \\ 1 & -1 & 2 & 3 & 2 \end{bmatrix}.$$

Eseguendo le operazioni p_{15} , $e_{21}(-1)$, $e_{31}(-1)$, $e_{41}(-1)$, $e_{51}(-3)$, $e_{52}(-1)$, $e_{43}(-1)$, $e_{53}(-2)$, si perviene alla seguente matrice a scala:

$$S = \begin{bmatrix} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 4 & 5 & 4 \\ 0 & 0 & 1 & -7 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -3 \end{bmatrix}.$$

Quindi

$$\det A = (-1) \cdot (-3) = 3.$$

Osservazione. (i) Nelle proprietà elementari abbiamo visto che se le righe (o le colonne) di una matrice quadrata di ordine n sono legate (il che equivale a dire che $\text{rk} A < n$), allora il determinante di tale matrice è nullo. Vale anche il viceversa. Cioè che se A è una matrice quadrata di ordine n e $\det A = 0$ allora $\text{rk} A < n$. Infatti se $\det A = 0$, allora sarà nullo anche il determinante di una matrice a scala per righe S equivalente ad A . Poiché S è a scala e $\det S = 0$, allora S deve avere una riga nulla. Quindi $\text{rk} S < n$, e perciò anche $\text{rk} A < n$ in quanto A ed S hanno lo stesso rango. Tenuto conto che, in un esercizio precedente, abbiamo dimostrato che $\text{rk} A = n$ se e solo se A è invertibile, si deduce il seguente importante:

• *Corollario.* Sia A una matrice quadrata di ordine n . Allora sono equivalenti le seguenti proprietà'.

- (i) Il rango di A è massimo, cioè $\text{rk}A = n$.
- (ii) A è invertibile.
- (iii) $\det A \neq 0$.
- (iv) Le righe di A formano una base per \mathbf{R}^n .
- (v) Le colonne di A formano una base per \mathbf{R}^n .

Naturalmente il Corollario precedente possiamo anche enunciarlo così':

• *Corollario.* Sia A una matrice quadrata di ordine n . Allora sono equivalenti le seguenti proprietà'.

- (i) Il rango di A non è massimo, cioè $\text{rk}A < n$.
- (ii) A non è invertibile.
- (iii) $\det A = 0$.
- (iv) Le righe di A formano un sistema di vettori di \mathbf{R}^n legato.
- (v) Le colonne di A formano un sistema di vettori di \mathbf{R}^n legato.⁵

(ii) Sia A una matrice quadrata di ordine n , e sia E una matrice elementare. Dalle proprietà' elementari dei determinanti deduciamo in particolare che il determinante del prodotto $E \cdot A$ è uguale al prodotto dei determinanti $\det E \cdot \det A$, cioè':

$$\det(E \cdot A) = \det E \cdot \det A = \det(A \cdot E).$$

Questo è un caso particolare del Teorema di Binet che ora andiamo a provare.

• *Il Teorema di Binet.*

Siano A e B due matrici quadrate dello stesso ordine n . Allora

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

Dimostrazione del Teorema di Binet. Distinguiamo due casi.

Primo caso, $\det A = 0$ oppure $\det B = 0$. Se $\det A = 0$ allora $\text{rk}A < n$. In un esercizio precedente abbiamo visto che in generale $\text{rk}(A \cdot B) \leq \text{rk}A$. Quindi $\text{rk}(A \cdot B) < n$ e cioè' $\det(A \cdot B) = 0$. Ma allora è chiaro che $\det(A \cdot B) = \det A \cdot \det B$, essendo entrambi i numeri nulli. Poiché si ha anche $\text{rk}(B \cdot A) \leq \text{rk}A$, come prima $\det(B \cdot A) = \det A \cdot \det B = 0$. Similmente si prova il teorema quando $\det B = 0$.

Secondo caso, $\det A \neq 0$ e $\det B \neq 0$. Allora entrambe le matrici hanno rango massimo, ed abbiamo visto che in tal caso ciascuna matrice è prodotto di matrici elementari. Ma per i prodotti con una matrice elementare il Teorema di Binet segue dalle proprietà' elementari dei determinanti. ■

⁵ Questi due corollari dovrebbero essere imparati a memoria.

7. Il calcolo del determinante tramite lo sviluppo di Laplace.

C'è un altro modo per calcolare il determinante, e' un metodo induttivo sull'ordine della matrice, e si basa sul cosiddetto *sviluppo di Laplace*, per definire il quale abbiamo bisogno di una definizione preliminare.

- *La definizione di complemento algebrico.*

Sia A una matrice quadrata di ordine n . Fissati degli indici $i, j = 1, \dots, n$, sia A_{ij} la sottomatrice di A ottenuta cancellando da A la riga di posto i e la colonna di posto j . Si osservi che anche A_{ij} e' quadrata, ma di ordine $n - 1$. Il numero

$$(-1)^{i+j} \det A_{ij}$$

si chiama *complemento algebrico di A di posto ij* . In qualche libro il complemento algebrico si chiama anche *cofattore di A di posto ij* .

Possiamo enunciare lo sviluppo di Laplace.

- *Lo sviluppo di Laplace del determinante.*

Sia $A = (a_{ij})$ una matrice quadrata di ordine $n > 1$.

Per ogni fissato indice di riga i si ha:

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

Cioe' il determinante di A e' uguale alla somma dei complementi algebrici $(-1)^{i+j} \det A_{ij}$ di una fissata riga A_i , ciscun complemento algebrico moltiplicato per la componente a_{ij} corrispondente. Tale formula e' nota sotto il nome di *sviluppo di Laplace del determinante di A rispetto alla riga di posto i* .

Inoltre, per ogni fissato indice di colonna j si ha:

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

Cioe' il determinante di A e' uguale alla somma dei complementi algebrici $(-1)^{i+j} \det A_{ij}$ di una fissata colonna A^j , ciscun complemento algebrico moltiplicato per la componente a_{ij} corrispondente. Tale formula e' nota sotto il nome di *sviluppo di Laplace del determinante di A rispetto alla colonna di posto j* .

Dimostrazione. La dimostrazione seguira' applicando le proprieta' elementari del determinante. Esaminiamo solo il caso dello sviluppo rispetto ad una riga, perche' nel caso delle colonne la dimostrazione e' simile.

Per semplificare la notazione fissiamo la prima riga, quindi poniamo $i = 1$. Allora dobbiamo provare che

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

Cominciamo con lo scrivere la prima riga di A nel seguente modo.

$$A_1 = a_{11}\mathbf{e}_1 + \cdots + a_{1n}\mathbf{e}_n$$

dove i vettori \mathbf{e}_j denotano i vettori canonici di \mathbf{R}^n . Per la multilinearita' del determinante allora sappiamo che

$$\det A = \det \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = a_{11} \det \begin{bmatrix} \mathbf{e}_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} + a_{12} \det \begin{bmatrix} \mathbf{e}_2 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} + \cdots + a_{1n} \det \begin{bmatrix} \mathbf{e}_n \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Poniamo

$$M^j = \begin{bmatrix} \mathbf{e}_j \\ A_2 \\ \vdots \\ A_n \end{bmatrix}.$$

Per concludere sara' sufficiente provare che

$$\det M^j = (-1)^{1+j} \det A_{1j}.$$

Ora dopo $j - 1$ scambi di colonna portiamo la matrice M^j nella matrice N^j che ha come prima colonna la colonna di posto j di M^j , e poi, nello stesso ordine, tutte le altre colonne di M^j . Quindi

$$\det M^j = (-1)^{j-1} \det N^j = (-1)^{1+j} \det N^j.$$

D'altra parte la matrice N^j ha come prima riga il vettore $\mathbf{e}_1 = (1, 0, \dots, 0)$, e tutte le componenti n_{hk} di N^j di posto $h \geq 2$ e $k \geq 2$ coincidono, ordinatamente, con le componenti di A_{1j} .

$$N^j = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{2j} & & & \\ \vdots & & A_{1j} & \\ \vdots & & & \\ a_{nj} & & & \end{bmatrix}$$

Percio', applicando l'algoritmo di Gauss per calcolare il determinante di N^j , troveremo il determinante di A_{1j} . ■

Esempio. Calcoliamo il determinante della matrice

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

con lo sviluppo rispetto alla prima riga. Allora:

$$\begin{aligned}\det A &= \sum_{j=1}^3 (-1)^{1+j} a_{1j} \det A_{1j} = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} \\ &= 1 \cdot \det \begin{bmatrix} 5 & 0 \\ 2 & 1 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 4 & 0 \\ -1 & 1 \end{bmatrix} + 3 \cdot \det \begin{bmatrix} 4 & 5 \\ -1 & 2 \end{bmatrix} = 5 - 8 + 39 = 36.\end{aligned}$$

Esempio. Naturalmente possiamo anche combinare le operazioni elementari con lo sviluppo di Laplace. Per esempio, andiamo a calcolare il determinante della seguente matrice:

$$A = \begin{bmatrix} 2-t & 1 & 1 \\ -11 & -5-t & -4 \\ 7 & 3 & 2-t \end{bmatrix}.$$

Con l'operazione $e^{32}(-1)$ otteniamo:

$$\det A = \det \begin{bmatrix} 2-t & 1 & 0 \\ -11 & -5-t & t+1 \\ 7 & 3 & -1-t \end{bmatrix}.$$

Possiamo portare fuori la componente $t+1$ dalla terza colonna, ottenendo:

$$\det A = (t+1) \det \begin{bmatrix} 2-t & 1 & 0 \\ -11 & -5-t & 1 \\ 7 & 3 & -1 \end{bmatrix}.$$

Ora applicando $e_{32}(1)$ abbiamo:

$$\det A = (t+1) \det \begin{bmatrix} 2-t & 1 & 0 \\ -11 & -5-t & 1 \\ -4 & -2-t & 0 \end{bmatrix}.$$

E sviluppando rispetto alla terza colonna si ha:

$$\det A = -(t+1) \det \begin{bmatrix} 2-t & 1 \\ -4 & -2-t \end{bmatrix} = -t^2(t+1).$$

Osservazione. (i) Come prima denotiamo con $A = (a_{ij})$ una matrice quadrata di ordine n . E' interessante osservare che se si fissa un indice di riga i , e si effettua la somma dei complementi algebrici $(-1)^{i+j} \det A_{ij}$ di A_i , ciascun complemento algebrico moltiplicato per la componente a_{hj} corrispondente, ma di un'altra riga A_h , $h \neq i$, si ottiene 0. Cioe' per ogni coppia di indici di riga $i, h = 1, \dots, n$ con $h \neq i$ si ha:

$$\sum_{j=1}^n (-1)^{i+j} a_{hj} \det A_{ij} = 0.$$

Tale formula vale anche per le colonne, cioè per ogni coppia di indici di colonna $j, h = 1, \dots, n$ con $h \neq j$ si ha:

$$\sum_{i=1}^n (-1)^{i+j} a_{ih} \det A_{ij} = 0.$$

La dimostrazione è molto semplice. Esaminiamo solo la prima formula, la seconda è simile. Sia A' la matrice che si ottiene da A sostituendo al posto della riga A_i la riga A_h . In questa nuova matrice appaiono due righe uguali, sappiamo quindi che $\det A' = 0$. Ma se sviluppiamo il determinante di A' rispetto alla riga di posto i otteniamo proprio $\det A' = \sum_{j=1}^n (-1)^{i+j} a_{hj} \det A_{ij}$.

(ii) Sulla base di quanto appena osservato, possiamo riassumere in un'unica formula lo sviluppo di Laplace del determinante e la formula di annullamento precedente. Cioè comunque si fissino due indici di riga $i, h = 1, \dots, n$ si ha:

$$\sum_{j=1}^n (-1)^{i+j} a_{hj} \det A_{ij} = \begin{cases} \det A & \text{se } i = h \\ 0 & \text{se } i \neq h. \end{cases}$$

Similmente, comunque si fissino due indici di colonna $j, h = 1, \dots, n$ si ha:

$$\sum_{i=1}^n (-1)^{i+j} a_{ih} \det A_{ij} = \begin{cases} \det A & \text{se } j = h \\ 0 & \text{se } j \neq h. \end{cases}$$

8. Il calcolo esplicito della matrice inversa.

Le formule che abbiamo appena dimostrato, ed il Teorema di Binet, ci mettono in condizione di capire quando una matrice è invertibile, e come è fatta la matrice inversa. A tale proposito, premettiamo una definizione.

- *La matrice aggiunta classica.*

Sia $A = (a_{ij})$ una matrice quadrata di ordine n . La *matrice aggiunta classica* di A , che denoteremo con $A^* = (a_{ij}^*)$, è quella matrice quadrata di ordine n la cui componente di posto ij è il complemento algebrico di A di posto ji . Cioè:

$$a_{ij}^* = (-1)^{i+j} \det A_{ji}.$$

Le formule che abbiamo dimostrato ci dicono proprio che:

$$A \cdot A^* = A^* \cdot A = (\det A) \cdot I.$$

Siamo in grado di calcolare la matrice inversa di una matrice invertibile.

- *Il calcolo esplicito della matrice inversa.*

Sia $A = (a_{ij})$ una matrice quadrata di ordine n . Allora A è invertibile se e solo se $\det A \neq 0$. In tal caso l'inversa di A è la matrice:

$$A^{-1} = \left(\frac{1}{\det A} \right) A^*.$$

Dimostrazione. Già sappiamo che A è invertibile se e solo se $\det A \neq 0$.

Se $\det A \neq 0$, allora nelle formule precedenti possiamo dividere per $\det A$ e dedurre:

$$A \cdot \left(\left(\frac{1}{\det A} \right) \cdot A^* \right) = \left(\left(\frac{1}{\det A} \right) \cdot A^* \right) \cdot A = I.$$

Cio' vuol dire che A è invertibile e la sua inversa è proprio $A^{-1} = \left(\frac{1}{\det A} \right) A^*$. ■

Esempio. Se A è una matrice di ordine 2:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

allora

$$A^* = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Perciò se $\det A \neq 0$, cioè se $ad - bc \neq 0$, allora A è invertibile e la sua inversa è:

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

Esercizio. Siano A e B matrici quadrate dello stesso ordine tali che $AB = I$. Provare che A è invertibile e che $A^{-1} = B$.

Svolgimento. Poiché $\det(AB) = \det A \cdot \det B = 1$ allora $\det A \neq 0$. Cio' prova che A è invertibile. Se A^{-1} è l'inversa di A allora:

$$B = IB = A^{-1}AB = A^{-1}I = A^{-1}. \quad \blacksquare$$

Esercizio. Siano A e B matrici quadrate dello stesso ordine tali che $BA = I$. Provare che A è invertibile e che $A^{-1} = B$.

Svolgimento. Simile al precedente. ■

Esercizio. Calcolare l'inversa della matrice

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

Svolgimento. Innanzitutto osserviamo che

$$\det A = -1$$

perciò possiamo dire che A è invertibile. Poi calcoliamo l'aggiunta classica di A :

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

Quindi

$$A^{-1} = \begin{bmatrix} -1 & 2 & 2 \\ 1 & -1 & -1 \\ -1 & 1 & 2 \end{bmatrix}. \quad \blacksquare$$

Osservazione. Utilizzando le operazioni elementari, si può rendere più veloce il calcolo dell'inversa di una matrice. Infatti si può usare un algoritmo, detto l'algoritmo di Gauss-Jordan, che ora andiamo a vedere.

- *L'algoritmo di Gauss-Jordan.*

Per calcolare l'inversa di una matrice quadrata A , $n \times n$, si può procedere come segue. Si affianca alla matrice A la matrice unitaria I di ordine n :

$$M_1 = [A | I].$$

Si ottiene così una matrice con n righe e $2n$ colonne. Ora si riduce a scala per righe la matrice M_1 tramite operazioni elementari sulle righe, per esempio, con l'algoritmo di Gauss.⁶ Queste stesse operazioni elementari porteranno la matrice A in una matrice a scala S , ed il prodotto complessivo delle matrici elementari corrispondenti apparirà al posto del blocco I a destra della matrice M_1 . Si perviene cioè ad una matrice intermedia M_2 del tipo:

$$M_2 := [S | E],$$

dove $S = EA$ è una matrice quadrata a scala di ordine n , ed E è il prodotto delle matrici elementari corrispondenti alle operazioni effettuate. Se e solo se A è invertibile, tutte le componenti di S sulla diagonale principale sono non nulle. Perciò con operazioni del tipo $e_i(k)$ su M_2 possiamo rendere le componenti della diagonale di S tutte uguali ad 1. Si perviene così ad una nuova matrice

$$M_3 = [T | F]$$

⁶ Se si esegue una data operazione elementare e sulle righe di M_1 , corrispondente alla matrice elementare E , la matrice M_1 si porta nella matrice $[EA | E]$. Cioè

$$[A | I] \xrightarrow{e} E \cdot [A | I] = [EA | E].$$

dove T e' una matrice quadrata a scala per righe di ordine n , le cui componenti sulla diagonale principale sono tutte uguali ad 1, F e' la matrice prodotto delle matrici elementari corrispondenti a tutte le operazioni eseguite, e $T = FA$. Infine, *partendo dall'ultima riga*, con operazioni di tipo $e_{ij}(k)$, si annullano tutte le componenti al di sopra delle componenti della diagonale principale di T (che sono i pivots delle righe di T). Così facendo la matrice T sarà sostituita dalla matrice I . In conclusione si arriva ad una matrice del tipo

$$M_4 = [I | G].$$

La matrice G e' il prodotto complessivo delle matrici elementari coinvolte nelle operazioni, e soddisfa la condizione $I = G \cdot A$. Perciò $G = A^{-1}$ e' la matrice inversa di A che stiamo cercando.

Questo procedimento e' noto in letteratura come *algoritmo di Gauss-Jordan*.

Riepilogando brevemente: l'idea e' di eseguire opportune operazioni elementari sulle righe di M_1 in modo da portare la matrice A in I . Così facendo la matrice I a destra in M_1 si porterà in una certa matrice G , prodotto di tutte le matrici elementari corrispondenti alle operazioni che hanno portato A in I . La matrice inversa di A e' proprio G , in quanto deve essere $I = GA$.

Esempio. Andiamo a calcolare l'inversa della matrice

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

Si parte dalla matrice

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

Riducendo a scala con le operazioni $e_{21}(-1)$ ed $e_{32}(\frac{1}{2})$ si ottiene

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -2 & -1 & -1 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}.$$

Poi si rendono uguali ad 1 i pivots delle righe con le operazioni $e_2(-\frac{1}{2})$ ed $e_3(2)$:

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -1 & 1 & 2 \end{bmatrix}.$$

Infine si annullano tutte le componenti sopra i pivots partendo dall'ultima riga con le operazioni $e_{23}(-\frac{1}{2})$ ed $e_{12}(-2)$:

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

L'inversa di A e' la matrice

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

9. Il calcolo del rango con l'uso del determinante.

In questo paragrafo vedremo come il calcolo del rango di una matrice puo' essere ricondotto ad un calcolo di determinanti. Premettiamo alcune definizioni.

Sia A una matrice $m \times n$. Una sottomatrice quadrata M di A di ordine q si dice *minore di A di ordine q* . Poi diremo che un minore M e' *nonsingolare* se $\det M \neq 0$. Invece diremo che un minore e' *singolare* se $\det M = 0$. Alla matrice A possiamo associare il seguente insieme di numeri naturali:

$$\mathcal{M}(A) := \{q \in \mathbf{N} : \text{esiste un minore } M \text{ di } A \text{ nonsingolare di ordine } q\}.$$

Esempio. Consideriamo la matrice

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & 2 & 6 \\ 1 & 0 & -1 & 0 \end{bmatrix}.$$

Allora $[0]$ e' un minore di A di ordine 1 singolare. Mentre $[6]$ e' un minore di A di ordine 1 nonsingolare. Poi $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ e' un minore di A singolare di ordine 2. Mentre $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ e' un minore di A di ordine 2 nonsingolare. Osserviamo infine che tutti i minori di ordine 3 di A sono singolari, e che chiaramente non esistono minori di A di ordine > 3 . In questo esempio si ha

$$\mathcal{M}(A) := \{q \in \mathbf{N} : \text{esiste un minore } M \text{ di } A \text{ nonsingolare di ordine } q\} = \{1, 2\}.$$

Il primo fatto che vogliamo mettere in evidenza e' il seguente:

- *Il rango di una matrice e' l'ordine massimo dei suoi minori nonsingolari. Cioe', data una matrice A , $\text{rk}(A) = \max \mathcal{M}(A)$.*

Dimostrazione. Sappiamo che se M e' una sottomatrice di A allora $\text{rk}(M) \leq \text{rk}(A)$. Percio' se M e' un minore nonsingolare di A di ordine q , allora

$$q = \text{rk}(M) \leq \text{rk}(A).$$

Ne consegue che $\max \mathcal{M}(A) \leq \text{rk}(A)$. Per completare la dimostrazione sara' sufficiente provare che A possiede un minore nonsingolare di ordine $q_0 := \text{rk}(A)$. A tale proposito, con operazioni elementari sulle righe, riduciamo la matrice A ad una matrice a scala per righe S . La matrice S ha le prime q_0 righe che sono non nulle, e le restanti tutte nulle.

Sia M il minore di S costituito dalle componenti comuni alle prime q_0 righe, ed alle q_0 colonne passanti per i pivots delle righe. Tale minore M è una matrice triangolare superiore di ordine q_0 . Ed ha rango q_0 , perché i pivots delle sue righe sono esattamente quelli delle righe di S . Tale minore si ottiene da un minore N di A di ordine q_0 , con le stesse operazioni elementari che hanno portato A in S . Ciò prova l'esistenza di un minore N di A nonsingolare di ordine $q_0 = \text{rk}(A)$. ■

Esempio. Con riferimento all'esempio precedente, poiché $\mathcal{M}(A) = \{1, 2\}$, allora $\text{rk}(A) = \max \mathcal{M}(A) = 2$.

Osservazione. La proprietà precedente costituisce un nuovo modo per calcolare il rango di una matrice A . Questo metodo non altera la matrice, mentre con le operazioni elementari occorre cambiare la matrice di cui siamo interessati al calcolo del rango. Nella pratica il nuovo metodo prevede il calcolo del determinante di tutte le sottomatrici quadrate di A . Questo calcolo è molto lento rispetto alle operazioni elementari. Può essere sensibilmente migliorato, come andremo a vedere tra poco. Prima però vediamo una conseguenza della proprietà precedente.

- *La definizione di minore fondamentale di una matrice.*

Sia A una matrice di rango p . In base alla proprietà precedente *esiste un minore di A di ordine p nonsingolare, e tutti i minori di A aventi ordine maggiore di p sono singolari*. Un minore nonsingolare di A di ordine uguale al rango di A si dice *minore fondamentale di A* . La seguente proprietà è una semplice conseguenza della precedente, e rappresenta un metodo per riconoscere direttamente dalla matrice A una base per lo spazio delle righe ed una base per lo spazio delle colonne.

- *Sia M un minore fondamentale $p \times p$ di una matrice A . Supponiamo che M si ottenga da A considerando le componenti comuni alle righe A_{i_1}, \dots, A_{i_p} ed alle colonne A^{j_1}, \dots, A^{j_p} di A . Allora A_{i_1}, \dots, A_{i_p} formano una base per lo spazio delle righe di A , e A^{j_1}, \dots, A^{j_p} formano una base per lo spazio delle colonne di A .*

Dimostrazione. Poiché A ha rango p , è sufficiente provare che i sistemi A_{i_1}, \dots, A_{i_p} e A^{j_1}, \dots, A^{j_p} sono liberi. A tale scopo, sia B la sottomatrice di A formata dalle p righe A_{i_1}, \dots, A_{i_p} di A . La matrice B possiede M come minore nonsingolare di ordine p . Poiché B ha p righe, allora $\max \mathcal{M}(B) = p$. Perciò, per la proprietà precedente, p è il rango di B . In particolare le sue righe A_{i_1}, \dots, A_{i_p} sono linearmente indipendenti. Ciò implica che A_{i_1}, \dots, A_{i_p} formano una base per lo spazio delle righe di A . L'affermazione sulle colonne si prova in modo simile. ■

Esempio. Riprendiamo l'esempio

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & 2 & 6 \\ 1 & 0 & -1 & 0 \end{bmatrix}.$$

Abbiamo visto che A ha rango $p = 2$. Allora il minore $M = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ è un minore fondamentale di A . Poiché M è formato con le componenti di A comuni alla prima

ed alla terza riga, ed alle prime due colonne, allora $(1, 2, 1, 3), (1, 0, -1, 0)$ e' una base per lo spazio delle righe di A , mentre $(1, 2, 1)^T, (2, 4, 0)^T$ e' una base per lo spazio delle colonne di A .

- *La definizione di orlato.*

Sia N un minore di A di ordine $q + 1$, e sia M un minore di N di ordine q . Allora si dice che N e' un orlato di M in A . Cio' premesso siamo in grado di enunciare la seguente proprieta', nota come il Teorema degli orlati, attribuito a Kronecker:

- *Teorema degli orlati.* Sia M un minore nonsingolare di una matrice A . Se tutti gli orlati di M in A sono singolari allora M e' un minore fondamentale di A .

Dimostrazione. Supponiamo che M si ottenga da A considerando le componenti comuni alle righe A_{i_1}, \dots, A_{i_p} ed alle colonne A^{j_1}, \dots, A^{j_p} di A . Per la prima delle proprieta' provate in questo paragrafo, e' chiaro che le righe A_{i_1}, \dots, A_{i_p} di A sono libere. Per concludere la dimostrazione e' sufficiente provare che tali righe generano le righe di A . A tale proposito, sia A_i una qualunque riga di A , con indice $i \notin \{i_1, \dots, i_p\}$, ed A^j una qualunque colonna, e consideriamo la seguente matrice quadrata:

$$N := \begin{bmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_p} & a_{i_1 j} \\ \dots & \dots & \dots & \dots \\ a_{i_p j_1} & \dots & a_{i_p j_p} & a_{i_p j} \\ a_{i j_1} & \dots & a_{i j_p} & a_{i j} \end{bmatrix}.$$

Si osservi che N , a meno di scambi di righe o di colonne, e' un orlato di M in A , a meno che $j \in \{j_1, \dots, j_p\}$. In ogni caso $\det N = 0$, per ipotesi sugli orlati di M , oppure perche' N possiede due colonne uguali (quando $j \in \{j_1, \dots, j_p\}$). Sviluppando il determinante di N secondo Laplace rispetto all'ultima colonna, otteniamo:

$$\lambda_1 a_{i_1 j} + \dots + \lambda_p a_{i_p j} + \lambda a_{i j} = 0,$$

dove $\lambda_1, \dots, \lambda_p, \lambda$ sono i cofattori di N relativi ai posti $i_1 j, \dots, i_p j, i j$. In particolare $\lambda = \det M \neq 0$. Poiche' i cofattori coinvolti nello sviluppo non dipendono da j , la formula precedente implica che

$$\lambda_1 A_{i_1} + \dots + \lambda_p A_{i_p} + \lambda A_i = \mathbf{0}.$$

Essendo $\lambda \neq 0$, ne consegue che A_i dipende linearmente da A_{i_1}, \dots, A_{i_p} . ■

Osservazione. Il Teorema degli orlati rende piu' veloce il calcolo del rango con l'uso del determinante. Infatti una volta individuato un minore nonsingolare M di A di ordine p , in prima battuta ci si puo' limitare al calcolo degli orlati di M . Se sono tutti singolari, allora M e' un minore fondamentale e p e' il rango di A . Altrimenti troveremo un orlato N di M nonsingolare. N ha ordine $p + 1$, e ripetiamo la procedura per N . Cosi' continuando, al crescere dell'ordine, diminuiranno i possibili orlati, fino a che non troveremo un minore fondamentale.

Esempio. Si consideri la seguente matrice

$$A := \begin{bmatrix} 0 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 5 & -3 & 4 & -1 \end{bmatrix}.$$

Vogliamo calcolare il rango di A utilizzando il Teorema degli orlati. Cominciamo individuando un minore nonsingolare di A di ordine 1, candidato minore fondamentale. Per sempio $M := [2]$. Orlando M con la seconda riga e la terza colonna otteniamo il minore nonsingolare $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$. Chiaramente $M := [2]$ non puo' essere un minore

fondamentale, ed allora ripartiamo con $M := \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ come nuovo candidato minore fondamentale di A , studiandone gli orlati. Se orliamo tale minore con la terza riga di A e la quarta colonna otteniamo un altro minore nonsingolare $\begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$. Come

prima $M := \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ non puo' essere un minore fondamentale. Ripartiamo ponendo

$$M := \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix},$$

che e' un minore nonsingolare di A . Ci sono solo due orlati di M in A , e sono le matrici

$$\begin{bmatrix} 0 & 2 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \\ 1 & 5 & -3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 5 & -3 & 4 & -1 \end{bmatrix}.$$

La prima si ottiene orlando M con la prima colonna di A e l'ultima riga di A , e la seconda si ottiene orlando M con la quinta colonna e l'ultima riga. Poiche' questi due orlati sono singolari, allora per il Teorema degli orlati M e' un minore fondamentale di A . In particolare, poiche' M e' ottenuto considerando le componenti comuni alle righe A_1, A_2, A_3 ed alle colonne A^2, A^3, A^4 di A , allora le righe $(0, 2, -1, 1, 0)$, $(0, 0, 1, -1, 1)$, $(1, 1, 0, 1, 0)$ di A formano una base per lo spazio delle righe di A , e le colonne $(2, 0, 1, 5)^T$, $(-1, 1, 0, -3)^T$, $(1, -1, 1, 4)^T$ di A formano una base per lo spazio delle colonne di A .

Indice dei paragrafi.

1. *Richiami e notazioni.*
2. *Prodotto riga per colonna tra matrici.*
3. *Il rango di una matrice.*
4. *Il calcolo del rango tramite le operazioni elementari.*
5. *Il determinante di una matrice quadrata.*
6. *Il calcolo del determinante tramite le operazioni elementari.*
7. *Il calcolo del determinante tramite lo sviluppo di Laplace.*
8. *Il calcolo esplicito della matrice inversa.*
9. *Il calcolo del rango con l'uso del determinante.*

Bibliografia.

Nella stesura di questo capitolo ho consultato i seguenti libri.

Sistemi lineari e matrici, M. Barnabei - F. Bonetti, Pitagora Editrice Bologna, 1994.

Algebra Lineare e Geometria Analitica, A. Franchetta, Liguori Editore, 1979.

Algebra lineare, S. Lipschutz, McGraw-Hill, 1994.

Algebra, S. Mac Lane - G. Birkhoff, Mursia, 1978.

Lezioni di Geometria I, F. Orecchia, Aracne, 1993.