4. Operazioni elementari per righe e colonne

Sia K un campo, e sia A una matrice $m \times n$ a elementi in K. Una *operazione elementare per righe* sulla matrice A è una operazione di uno dei seguenti tre tipi:

1) scambio di due righe di A;
2) moltiplicazione di una riga di A per uno scalare non nullo;
3) sostituzione di una riga di A con la somma della riga stessa e di un multiplo di un'altra riga.

In modo analogo possiamo definire operazioni elementari per colonne. È utile osservare che l'operazione inversa di una operazione elementare per righe (o per colonne) è anch'essa un'operazione elementare per righe (o per colonne). In effetti, se B è ottenuta da A scambiando la riga i-esima con la j-esima, anche A si ottiene da B scambiando la riga i-esima con la j-esima. Se invece B è ottenuta da A moltiplicando per k la riga i-esima, A si ottiene da B moltiplicando la riga i-esima per $1/k$. Infine, se B è ottenuta da A sommando alla riga i-esima k volte la riga j-esima, A si ottiene da B sottraendo dalla riga i-esima k volte la riga j-esima.

Esaminiamo le colonne della matrice A partendo da sinistra. Ignoriamo eventuali colonne nulle e fissiamo la nostra attenzione sulla prima colonna non nulla (se $A \neq 0$). Supponiamo che si tratti della i_1-esima colonna, e che $a_{j,i_1} \neq 0$. Scambiando la prima e la j-esima riga, e dunque tramite una operazione elementare per righe di tipo 1), otteniamo da A una matrice

$$A' = \begin{pmatrix} 0 & \dots & 0 & a'_{1,i_1} & \dots & a'_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a'_{m,i_1} & \dots & a'_{m,n} \end{pmatrix}$$

con $a'_{1,i_1} \neq 0$. Effettuiamo operazioni elementari per righe di tipo 3), sostituendo la riga j-esima con la differenza tra la riga stessa e $a_{j,i_1}/a_{1,i_1}$ volte la prima riga, per ogni j compreso tra 2 e m. Il risultato è

$$B = \begin{pmatrix} 0 & \dots & 0 & b_{1,i_1} & b_{1,i_1+1} & \dots & b_{1,n} \\ 0 & & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & b_{m,i_1+1} & \dots & b_{m,n} \end{pmatrix},$$

dove $b_{1,i_1} = a'_{1,i_1} \neq 0$. Ripetiamo lo stesso procedimento con la sottomatrice

$$\begin{pmatrix} 0 & \dots & 0 & b_{2,i_1+1} & \dots & b_{2,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & b_{m,i_1+1} & \dots & b_{m,n} \end{pmatrix}.$$

Se questa non è nulla, una successione finita di operazioni elementari per righe dà, a partire da B, una matrice della forma

$$B' = \begin{pmatrix} 0 & \dots & 0 & b'_{1,i_1} & & \dots & & & & \\ 0 & & 0 & 0 & \dots & 0 & b'_{2,i_2} & \dots & & \\ \vdots & & & & & & 0 & b'_{3,i_2+1} & \dots & \\ & & & & & & \vdots & & & \\ 0 & & & \dots & & & 0 & b'_{m,i_2+1} & \dots & b'_{m,n} \end{pmatrix},$$

con $b'_{1,i_1} = b_{1,i_1} = a'_{1,i_1} \neq 0$ e $b'_{2,i_2} \neq 0$. Ripetiamo lo stesso procedimento sulla matrice ottenuta da B' cancellando le prime due righe, e così via. La conclusione è che, con una successione finita di operazioni elementare per righe, si ottiene da A una matrice

$$C = \begin{pmatrix} 0 & \cdot\cdot & 0 & c_{1,i_1} & \cdot\cdot & & & & & & & & & & \\ 0 & & & 0 & \cdot\cdot & 0 & c_{2,i_2} & \cdot\cdot & & & & & & & \\ & & & & & & 0 & \cdot\cdot & 0 & c_{3,i_3} & \cdot\cdot & & & & \\ & & & & & & & & & 0 & \cdot\cdot & 0 & c_{4,i_4} & \cdot\cdot & \\ \cdot & & & & & & & & & & & & & & \\ \cdot & & & & & & & & & & & & & & \\ & & & & & \cdot\cdot & & & & & & & 0 & c_{h,i_h} & \cdot\cdot \\ & & & & & & & & & & & & & 0 & 0 \\ 0 & & & & & & & & \cdot\cdot & & & & & & 0 \end{pmatrix}$$

dove $i_1 < i_2 < \dots < i_h$, $c_{j,i_j} \neq 0$ per $j = 1, \dots, h$, e sotto e a sinistra della linea spezzata vi sono solo zeri. Naturalmente le colonne nulle all'estremità sinistra della matrice e le righe nulle nella sua parte più bassa possono esserci o no. Diremo che una matrice come C è *a scaletta* (inglese: *row-echelon form*). Il procedimento appena illustrato per ottenere, a partire da una matrice data, una matrice a scaletta va sotto il nome di *eliminazione gaussiana* (per righe).

Il rango per righe e il rango per colonne di una matrice a scaletta si calcolano facilmente. Ricordiamo che il rango per righe di una matrice è la dimensione dello spazio vettoriale generato dalle righe, o equivalentemente il numero massimo di righe linearmente indipendenti; il rango per colonne si definisce allo stesso modo, sostituendo ovunque la parola "riga" con "colonna". Indichiamo con $C_1, \dots, C_m$ le righe della matrice C. Dico che il rango per righe di C è pari ad h, cioè al numero delle righe non nulle. In effetti, consideriamo una combinazione lineare

$$(r_1, \dots, r_n) = \sum_{j=1}^{h} \alpha_j C_j$$

delle righe di C, e supponiamola nulla. Notiamo che $r_{i_1} = \alpha_1 c_{1,i_1}$, e quindi che $\alpha_1 = 0$, dato che $c_{1,i_1} \neq 0$. Ma allora $r_{i_2} = \alpha_2 c_{2,i_2}$, da cui $\alpha_2 = 0$. Procedendo allo stesso modo si mostra che anche $\alpha_3, \dots, \alpha_h$ sono nulli. Dunque $C_1, \dots, C_h$ sono indipendenti.

Mostreremo ora che anche il rango per colonne di C vale h. Indicheremo le colonne di C con $C^{(1)}, \dots, C^{(n)}$. Esse sono tutte contenute nel sottospazio h-dimensionale di K^n costituito dalle colonne con gli ultimi $n-h$ elementi nulli. Dunque il rango per colonne di C non supera h; per mostrare che vale esattamente h basta osservare che le colonne $C^{(i_1)}, C^{(i_2)}, \dots, C^{(i_h)}$ sono indipendenti, il che si dimostra procedendo esattamente come si è fatto per calcolare il rango per righe di C.

Applicheremo le considerazioni svolte finora per dimostrare il seguente risultato.

LEMMA (4.1). *Sia A una matrice a elementi in K. Il rango per righe e il rango per colonne di A sono uguali.*

Dimostrazione. Abbiamo osservato che il lemma vale per le matrici a scaletta. Poiché A può essere trasformata in una matrice a scaletta con un numero finito di operazioni elementari per righe, basterà mostrare che ognuna di queste operazioni non altera il rango per righe e quello per colonne. Sia B una matrice ottenuta da A tramite una operazione elementare per righe. Le righe di B sono combinazioni lineari di righe di A, e viceversa. Quindi lo spazio vettoriale generato dalle righe di A coincide con quello generato dalle righe di B. Ne segue che A e B hanno lo stesso rango per righe. Occupiamoci ora del rango per colonne. Scriviamo

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} = (\, A^{(1)}, \dots, A^{(n)} \,)\,,$$

$$B = \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} = (\, B^{(1)}, \dots, B^{(n)} \,)\,,$$

dove le $A^{(i)}$ sono le colonne di A e le $B^{(i)}$ quelle di B. Mostreremo che, se vale una relazione di dipendenza lineare $\sum \alpha_s A^{(s)} = 0$ tra le colonne di A, la stessa relazione vale tra le colonne di B, cioè che $\sum \alpha_s B^{(s)} = 0$, e viceversa. Di conseguenza le colonne $A^{(i_1)}, \dots, A^{(i_k)}$ sono indipendenti se e solo se lo sono $B^{(i_1)}, \dots, B^{(i_k)}$, e dunque il rango per colonne di A e quello di B coincidono. Per dimostrare quanto affermato conviene procedere separatamente per ognuno dei tre tipi di operazioni per righe. Tratteremo solo delle operazioni di tipo 3), lasciando al lettore il compito di fare altrettanto per gli altri due tipi di operazioni. Supponiamo dunque che la matrice B sia ottenuta da A rimpiazzando la i-esima riga con la riga stessa più c volte la riga j-esima. In altre parole $b_{i,s} = a_{i,s} + ca_{j,s}$ per ogni s, mentre $b_{\ell,s} = a_{\ell,s}$ quando $\ell \neq i$. Dire che $\sum \alpha_s A^{(s)} = 0$ equivale a dire che $\sum \alpha_s a_{\ell,s} = 0$ per ogni ℓ. Ne segue che $\sum \alpha_s b_{\ell,s} = 0$ per $\ell \neq i$, e che

$$\sum_s \alpha_s b_{i,s} = \sum_s \alpha_s a_{i,s} + c \sum_s \alpha_s a_{j,s} = 0\,,$$

cioè che $\sum \alpha_s B^{(s)} = 0$. La dimostrazione è completa.

Il lemma (4.1) dice che si può parlare semplicemente di rango di una matrice, senza dover specificare se per righe o per colonne. Dal lemma segue anche che il rango di una matrice è pari a quello della sua trasposta.

L'eliminazione gaussiana dà un metodo semplice ed efficiente per calcolare l'inversa di una matrice quadrata, quando questa esiste. Sia A una matrice $n \times n$. Costruiamo una matrice $n \times 2n$, che indichiamo con $(\,A \quad I\,)$, giustapponendo ad A la matrice identità $n \times n$. Tramite eliminazione gaussiana (per righe) possiamo ridurre questa matrice in forma a scaletta; il risultato è una matrice $(\,B \quad C\,)$, dove B e C sono matrici $n \times n$. Il rango di B è uguale a quello di A. Dunque, se B ha qualche riga nulla, A non è invertibile. In caso contrario

$$B = \begin{pmatrix} b_{11} & & \dots & & b_{1n} \\ 0 & b_{22} & & \dots & \\ 0 & 0 & b_{33} & \dots & \\ \vdots & & & & \\ 0 & 0 & \dots & 0 & b_{nn} \end{pmatrix}$$

e tutti i b_{ii} sono diversi da zero. Possiamo continuare il processo di eliminazione gaussiana in modo da cancellare tutti gli elementi sopra la diagonale di B. Per farlo possiamo sottrarre dalla prima riga di $(\,B \quad C\,)$ la seconda riga moltiplicata per b_{12}/b_{22}, poi la terza riga moltiplicata per b_{13}/b_{33}, e così via, poi sottrarre dalla seconda riga b_{23}/b_{33} volte la terza, e così via. Il risultato finale è una matrice $(\,B' \quad C'\,)$, dove B' è una matrice diagonale i cui elementi diagonali sono $b_{11}, b_{22}, \dots, b_{nn}$. A questo punto si possono moltiplicare la prima riga per l'inverso di b_{11}, la seconda per l'inverso di b_{22}, e così via. Il risultato finale è una matrice $(\,I \quad D\,)$, dove I è la matrice identità $n \times n$. Dico che D è l'inversa di A. Per mostrarlo ricordiamo innanzitutto che $(\,I \quad D\,)$ è stata ottenuta da $(\,A \quad I\,)$ tramite una successione finita di operazioni elementari per righe. Osseviamo poi che ogni operazione per righe può essere realizzata tramite moltiplicazione a sinistra per una opportuna matrice $n \times n$. Torneremo fra poco su questa affermazione; ora usiamola per concludere. L'osservazione dice che

$$(\,I \quad D\,) = M_1 M_2 \dots M_h \,(\,A \quad I\,)\,,$$

dove le M_i sono opportune matrici $n \times n$. Scrivendo M al posto del prodotto $M_1 \dots M_h$ otteniamo che

$$(\,I \quad D\,) = M\,(\,A \quad I\,) = (\,MA \quad MI\,)\,,$$

e dunque che $MA = I$ e $D = M$. Ciò mostra che D è in effetti l'inversa di A.

Torniamo all'osservazione che ogni operazione per righe corrisponde alla moltiplicazione a sinistra per una opportuna matrice. Sia E una matrice $n \times m$, e sia G un'altra matrice ottenuta da E per mezzo di una operazione elementare per righe. Allora si verifica immediatamente che $G = QE$, dove $E = (q_{ij})$ è una matrice $n \times n$ di uno dei seguenti tipi:

- se G è ottenuta scambiando la riga i-esima con la j-esima

$$\begin{aligned} &q_{ij} = q_{ji} = 1 \\ &q_{hk} = 1 \quad \text{se } h = k,\ h \neq i, j \\ &q_{hk} = 0 \quad \text{altrimenti;} \end{aligned}$$

- se G è ottenuta moltiplicando la riga i-esima per a

$$\begin{aligned} &q_{ii} = a \\ &q_{hk} = 1 \quad \text{se } h = k,\ h \neq i \\ &q_{hk} = 0 \quad \text{altrimenti;} \end{aligned}$$

- se G è ottenuta sommando alla riga i-esima a volte la j-esima

$$\begin{aligned} &q_{ij} = a \\ &q_{hk} = 1 \quad \text{se } h = k \\ &q_{hk} = 0 \quad \text{altrimenti.} \end{aligned}$$

Allo stesso modo, ogni operazione elementare per colonne corrisponde alla moltiplicazione a destra per una opportuna matrice quadrata; lasciamo al lettore il compito di trovare quali siano queste matrici.

Terminiamo questa sezione con un semplice esempio esplicito di eliminazione gaussiana. Ci proponiamo di trovare l'inversa della matrice

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

Partiamo dunque dalla matrice

$$(\,A \quad I\,) = \begin{pmatrix} 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 3 & -5 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Ognuna delle matrici elencate qui sotto è ottenuta dalla precedente per mezzo di operazioni elementari per righe; a destra di ognuna sono indicate le operazioni effettuate per ottenerla.

$$\begin{pmatrix} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 3 & -5 & 1 & 0 & 0 & 1 \end{pmatrix} \qquad \text{Scambio della prima e della seconda riga.}$$

$$\begin{pmatrix} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & -14 & -5 & 0 & -3 & 1 \end{pmatrix} \qquad \text{Sottrazione dalla terza riga di tre volte la prima.}$$

$$\begin{pmatrix} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & -12 & 7 & -3 & 1 \end{pmatrix} \qquad \text{Aggiunta alla terza riga di sette volte la seconda.}$$

$$\begin{pmatrix} 1 & 0 & \frac{7}{2} & -\frac{3}{2} & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & -12 & 7 & -3 & 1 \end{pmatrix} \qquad \text{Sottrazione dalla prima riga di 3/2 volte la seconda.}$$

$$\begin{pmatrix} 1 & 0 & 0 & \frac{13}{24} & \frac{1}{8} & \frac{7}{24} \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & -12 & 7 & -3 & 1 \end{pmatrix} \qquad \text{Aggiunta alla prima riga di 7/24 volte la terza.}$$

$$\left(\begin{array}{cccccc} 1 & 0 & 0 & \frac{13}{24} & \frac{1}{8} & \frac{7}{24} \\ 0 & 2 & 0 & \frac{5}{12} & \frac{1}{4} & -\frac{1}{12} \\ 0 & 0 & -12 & 7 & -3 & 1 \end{array}\right)$$

Sottrazione dalla seconda riga di $1/12$ della terza.

$$\left(\begin{array}{cccccc} 1 & 0 & 0 & \frac{13}{24} & \frac{1}{8} & \frac{7}{24} \\ 0 & 1 & 0 & \frac{5}{24} & \frac{1}{8} & -\frac{1}{24} \\ 0 & 0 & 1 & -\frac{7}{12} & \frac{1}{4} & -\frac{1}{12} \end{array}\right)$$

Moltiplicazione dalla seconda riga per $1/2$ e della terza per $-1/12$.

Il lettore potrà verificare direttamente che

$$D = \begin{pmatrix} \frac{13}{24} & \frac{1}{8} & \frac{7}{24} \\ \frac{5}{24} & \frac{1}{8} & -\frac{1}{24} \\ -\frac{7}{12} & \frac{1}{4} & -\frac{1}{12} \end{pmatrix}$$

è l'inversa di A.