dove il numero delle colonne di A è uguale al numero delle righe di A'. Allora la regola per la moltiplicazione per blocchi è:

$$(1.20)\qquad \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right] = \left[\begin{array}{c|c} AA'+BC' & AB'+BD' \\ \hline CA'+DC' & CB'+DD' \end{array}\right].$$

Per esempio,

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ \hline 0 & 1 & 7 \end{array}\right]\left[\begin{array}{cc|cc} 2 & 3 & 1 & 1 \\ 4 & 1 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 2 & 8 & 6 & 1 \\ \hline 4 & 8 & 7 & 0 \end{array}\right].$$

In questo prodotto, il blocco in alto a sinistra è $[1\ \ 0]\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} + [5][0\ \ 1] = [2\ \ 8]$, e così via.

Anche questa regola può essere verificata direttamente a partire dalla definizione di moltiplicazione tra matrici. In generale, la moltiplicazione per blocchi può essere usata ogni volta che due matrici sono scomposte in sottomatrici in modo tale che i prodotti che occorre considerare siano definiti.

Oltre a facilitare i calcoli, la moltiplicazione per blocchi è uno strumento utile per dimostrare proprietà relative alle matrici, per induzione.

2 Riduzione per righe

Sia $A = (a_{ij})$ una matrice $m \times n$ e consideriamo una matrice variabile $n \times p$, diciamo $X = (x_{ij})$. Allora l'equazione matriciale

$$(2.1)\qquad Y = AX$$

definisce la matrice $m \times p$ $Y = (y_{ij})$ come funzione di X. Tale operazione è detta *moltiplicazione a sinistra* per A:

$$(2.2)\qquad y_{ij} = a_{i1}x_{1j} + \cdots + a_{in}x_{nj}.$$

Si noti che nella formula (2.2) l'elemento y_{ij} dipende soltanto da $x_{1j}, \ldots, x_{nj}$, ossia dalla j-esima *colonna* di X e dalla i-esima riga della matrice A. Dunque A opera separatamente su ciascuna colonna di X e possiamo capire il modo in cui A agisce, considerando la sua azione su un vettore colonna X:

$$\boxed{\qquad\quad A \qquad\quad}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

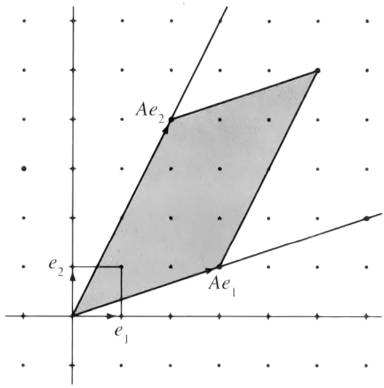
La moltiplicazione a sinistra per A sui vettori colonna può essere interpretata come una funzione dallo spazio dei vettori colonna X di dimensione n nello spazio dei vettori colonna Y di dimensione m, oppure come una collezione di m funzioni di n variabili:

$$y_i = a_{i1}x_1 + \cdots + a_{in}x_n \quad (i = 1, \ldots, m).$$

Essa è detta *trasformazione lineare*, poiché le funzioni sono lineari e omogenee. (Una funzione *lineare* di un insieme di variabili $u_1, \ldots, u_k$ è una funzione della forma $a_1u_1 + \cdots + a_ku_k + c$, dove $a_1, \ldots, a_k, c$ sono scalari. Una funzione siffatta si dice *lineare omogenea* se il termine costante c è zero.)

Una rappresentazione grafica dell'azione della matrice 2×2 $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ è riportata qui sotto. Essa manda il piano in se stesso:

(2.3)



Ritornando all'azione di A su una matrice X di tipo $n \times p$, possiamo interpretare il fatto che A agisce allo stesso modo su ciascuna colonna di X nel modo seguente. Denotiamo con Y_i la i-esima riga di Y, considerata come un *vettore riga*:

$$Y = \begin{bmatrix} \text{---}\; Y_1 \;\text{---} \\ \text{---}\; Y_2 \;\text{---} \\ \vdots \\ \text{---}\; Y_n \;\text{---} \end{bmatrix}.$$

Usando la notazione vettoriale, possiamo calcolare Y_i mediante le righe X_j di X:

(2.4) $\quad Y_i = a_{i1}X_1 + \cdots + a_{in}X_n.$

Si tratta precisamente di una riformulazione di (2.2) e di un altro esempio di moltiplicazione per blocchi. Per esempio, l'ultima riga del prodotto

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

può essere calcolata come

$$3[1 \quad 0]+4[4 \quad 2]-6[3 \quad 2]=[1 \quad -4].$$

Se A è una matrice quadrata, si parla spesso della moltiplicazione a sinistra per A come di una *operazione sulle righe.*

Le matrici più semplici diverse dalla matrice nulla sono le *unità matriciali*, che denotiamo con e_{ij}:

(2.5)
$$e_{ij} = \begin{array}{cc} & j \\ i & \begin{bmatrix} & \cdot & \\ \cdot\cdot & 1 & \cdot\cdot\cdot\cdot \\ & \vdots & \end{bmatrix} \end{array}.$$

Tale matrice e_{ij} ha tutti gli elementi nulli tranne l'elemento di indici i, j che è uguale a 1. (Anche se di solito le matrici vengono denotate con lettere maiuscole, tuttavia per le unità matriciali si usano per tradizione le lettere minuscole.) Le unità matriciali sono utili poiché ogni matrice $A = (a_{ij})$ può essere scritta come una somma del tipo:

$$A = a_{11}e_{11} + a_{12}e_{12} + \cdots + a_{nn}e_{nn} = \sum_{i,j} a_{ij}e_{ij}.$$

Gli indici i, j sotto il simbolo $\sum$ indicano che la somma è estesa a tutti i valori di i e a tutti i valori di j. Per esempio:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & \\ & \end{bmatrix} + \begin{bmatrix} & 2 \\ & \end{bmatrix} + \begin{bmatrix} & \\ 1 & \end{bmatrix} + \begin{bmatrix} & \\ & 4 \end{bmatrix} = 3e_{11} + 2e_{12} + 1e_{21} + 4e_{22}.$$

Una somma siffatta è chiamata *combinazione lineare* delle matrici e_{ij}.

Le unità matriciali sono utili per lo studio dell'addizione di matrici e della moltiplicazione per uno scalare. Tuttavia, nel caso della moltiplicazione tra matrici, sono più utili alcune matrici quadrate chiamate *matrici elementari*. Ve ne sono di tre tipi

(2.6i)
$$\begin{bmatrix} 1 & & & \\ & \cdot & a & \\ & & \cdot & \\ & & & \cdot \\ & & & & 1 \end{bmatrix} \quad \text{oppure} \quad \begin{bmatrix} 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & a & & \cdot \\ & & & & 1 \end{bmatrix} = I + ae_{ij} \quad (i \neq j).$$

Una matrice siffatta ha tutti gli elementi diagonali uguali a 1 e un solo elemento diverso da zero al di fuori della diagonale.

(2.6ii)
$$\begin{bmatrix} 1 & & & & & \\ & \cdot & & & & \\ & & 0 & & 1 & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & 1 & & 0 & \\ & & & & & \cdot \\ & & & & & & 1 \end{bmatrix} = I + e_{ij} + e_{ji} - e_{ii} - e_{jj}.$$

In questo caso, gli elementi diagonali a_{ii} e a_{jj} di I sono uguali a zero, mentre gli elementi a_{ij} e a_{ji} sono uguali a 1. (La formula precedente, espressa mediante le unità matriciali, è piuttosto brutta e pertanto sarà poco usata.)

$$\text{(2.6iii)} \quad \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} = I + (c-1)e_{ii}, \quad (c \neq 0).$$

In questo caso, un solo elemento diagonale della matrice identica è sostituito da un numero c diverso da zero.

Le matrici elementari 2×2 sono

$$\text{(i)} \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}; \quad \text{(ii)} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \text{(iii)} \begin{bmatrix} c & \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & \\ & c \end{bmatrix};$$

dove, come sopra, a è arbitrario e c è un qualunque numero diverso da zero.

Le matrici elementari E agiscono su una matrice X nel modo seguente:

(2.7) Per ottenere la matrice EX, occorre:

Tipo (i): *Sostituire la i-esima riga X_i con $X_i + aX_j$*, oppure *aggiungere $a \cdot X_j$ alla riga X_i*;

Tipo (ii): *Scambiare tra loro le righe X_i e X_j*;

Tipo (iii): *Moltiplicare la riga X_i per uno scalare non nullo c.*

Queste operazioni sono chiamate *operazioni elementari sulle righe*. Dunque la moltiplicazione per una matrice elementare è un'operazione elementare sulle righe. Tali regole di moltiplicazione vanno verificate con cura.

(2.8) LEMMA *Le matrici elementari sono invertibili e le loro inverse sono anch'esse matrici elementari.*

La dimostrazione è semplicemente un calcolo. L'inversa di una matrice elementare è la matrice corrispondente all'operazione inversa sulle righe. Se $E = I + ae_{ij}$ è del tipo (i), allora $E^{-1} = I - ae_{ij}$ (basta "sottrarre aX_j dalla riga X_i"). Se E è del tipo (ii), allora $E^{-1} = E$. Infine, se E è del tipo (iii), allora E^{-1} è dello stesso tipo, con c^{-1} nella stessa posizione occupata da c in E (basta "moltiplicare la riga X_i per c^{-1}"). ∎

Studiamo ora l'effetto di operazioni elementari sulle righe (2.7) su una matrice A, al fine di ottenere una matrice più semplice A':

$$A \overset{\text{successione di operazioni}}{\longrightarrow\longrightarrow\cdots\longrightarrow} A'.$$

Poiché ogni operazione elementare sulle righe è ottenuta mediante la moltiplicazione per una matrice elementare, possiamo esprimere il risultato di una successione di operazioni siffatte mediante il prodotto per una successione $E_1, \dots, E_k$ di matrici elementari:

$$A' = E_k \cdots E_2 E_1 A. \tag{2.9}$$

Tale procedimento è chiamato *riduzione per righe* oppure *eliminazione di Gauss*. Per esempio, possiamo semplificare la matrice

$$M = \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \tag{2.10}$$

con operazioni elementari del primo tipo per eliminare quanti più elementi possibile:

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 2 & 6 & 3 & 7 \end{bmatrix} \longrightarrow$$

$$\longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix} \longrightarrow\longrightarrow \begin{bmatrix} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}.$$

La riduzione per righe è un metodo utile per risolvere i sistemi di equazioni lineari. Sia dato un sistema di m equazioni lineari in n incognite, diciamo $AX = B$ come in (1.9), essendo A una matrice $m \times n$, X un vettore colonna costituito da incognite e B un vettore colonna assegnato. Per risolvere tale sistema, consideriamo la matrice a blocchi $m \times (n+1)$

$$M = [A \,|\, B] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_n \end{array}\right], \tag{2.11}$$

ed effettuiamo operazioni sulle righe per semplificare M. Si noti che $EM = [EA \,|\, EB]$. Sia

$$M' = [A' \,|\, B']$$

il risultato di successive operazioni sulle righe. L'osservazione fondamentale è la seguente:

(2.12) PROPOSIZIONE *Le soluzioni del sistema $A'X = B'$ coincidono con le soluzioni del sistema $AX = B$.*

Dimostrazione. Poiché M' è ottenuta mediante una successione di operazioni elementari sulle righe, si ha:

$$M' = E_r \cdots E_1 M.$$

Poniamo $P = E_r \cdots E_1$. Tale matrice è invertibile, in virtù del lemma (2.8) e della proposizione (1.18). Inoltre,

$$M' = [A' \mid B'] = [PA \mid PB].$$

Ora, se X è una soluzione del sistema di partenza $AX = B$, si ha: $PAX = PB$, ossia $A'X = B'$. Dunque X è anche una soluzione del nuovo sistema. Viceversa, se $A'X = B'$, allora $AX = P^{-1}A'X = P^{-1}B' = B$, sicché X è anche una soluzione del sistema $AX = B$. ■

Per esempio, consideriamo il sistema:

$$\begin{aligned} x_1 + \phantom{2x_2 +{}} 2x_3 + x_4 &= 5 \\ x_1 + x_2 + 5x_3 + 2x_4 &= 7 \\ x_1 + 2x_2 + 8x_3 + 4x_4 &= 12 \end{aligned} \tag{2.13}$$

La sua matrice completa è la matrice M considerata sopra (2.10); quindi la riduzione per righe di tale matrice mostra che il sistema di equazioni assegnato è equivalente al sistema:

$$\begin{aligned} x_1 + \phantom{x_2 +{}} 2x_3 &= 2 \\ x_2 + 3x_3 &= -1 \\ x_4 &= 3 \end{aligned}$$

Possiamo ricavare immediatamente le soluzioni di questo sistema: basta scegliere un valore arbitrario per x_3 e poi risolvere il sistema nelle variabili x_1, x_2 e x_4. La soluzione generale del sistema (2.13) può essere scritta pertanto nella forma

$$x_3 = c_3, \;\; x_1 = 2 - 2c_3, \;\; x_2 = -1 - 3c_3, \;\; x_4 = 3,$$

dove c_3 è un numero arbitrario.

Ritorniamo ora alla riduzione per righe di una matrice arbitraria. Non è difficile vedere che, mediante una successione di operazioni sulle righe, ogni matrice A può essere ridotta ad una matrice della forma:

$$(2.14) \qquad A = \boxed{\begin{matrix} 1 & *..* & 0 & *..* & 0 & *...* & 0 & \\ & & 1 & *..* & 0 & *...* & 0 & \\ & & & & 1 & *...* & 0 & \cdots \\ & & & & & & 1 & \\ & & & & & & & \ddots \end{matrix}}$$

dove $*$ denota un numero qualsiasi e lo spazio vuoto è costituito da zeri. Essa è chiamata una *matrice a scala per righe*. Per esempio

$$\begin{bmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

è una matrice a scala per righe. Tale risulta la forma finale della riduzione della matrice (2.10). La definizione di matrice a scala per righe, o più semplicemente di matrice a scala, è la seguente:

(2.15) (a) *Il primo elemento non nullo in ciascuna riga è* 1. *Tale elemento è chiamato* pivot.

(b) *Il primo elemento non nullo della* $(i+1)$*-esima riga si trova alla destra del primo elemento non nullo della* i*-esima riga.*

(c) *Gli elementi al di sopra di un pivot sono nulli.*

Per effettuare una riduzione per righe, occorre svolgere le seguenti operazioni: Trovare la prima colonna che contiene un elemento non nullo (se ciò non è possibile, allora $A = 0$ e 0 è una matrice a scala per righe); scambiare le righe usando un'operazione elementare del tipo (ii), portando un elemento non nullo nella prima riga; normalizzare tale elemento trasformandolo in 1, usando un'operazione del tipo (iii); eliminare gli altri elementi nella sua colonna mediante una successione di operazioni del tipo (i). Il risultato finale sarà una matrice avente la forma a blocchi:

$$\left[\begin{array}{ccc|c|ccc} 0 & \dots & 0 & 1 & * & \dots\dots & * \\ \hline 0 & \dots & 0 & 0 & * & \dots\dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & * & \dots\dots & * \end{array}\right],$$

che possiamo scrivere come $\left[\begin{array}{c|c|c} & 1 & B \\ \hline & & D \end{array}\right] = A'$.

A questo punto si può operare su D (che è più piccola di A) così come si è fatto su A, e andare avanti così. Formalmente, stiamo applicando il principio di induzione completa [cfr. app. (2.6)] sul numero di righe della matrice. Per ipotesi induttiva possiamo supporre che ogni matrice avente meno righe di A sia riducibile alla forma a scala per righe; quindi, in particolare, la matrice D può essere ridotta ad una tale matrice, diciamo D''. Le operazioni sulle righe che portano D in D'' non alterano gli altri blocchi che costituiscono A'. Quindi A' può essere ridotta alla matrice

$$\left[\begin{array}{c|c|c} & 1 & B \\ \hline & & D'' \end{array}\right] = A''$$

che soddisfa le proprietà 2.15(a) e (b). Osserviamo ora che gli elementi di B al di sopra dei pivot di D'' possono essere eliminati, e concludiamo la riduzione di A alla forma a scala per righe. ∎

Si può dimostrare che la matrice a scala per righe ottenuta a partire da una matrice assegnata A mediante riduzione per righe è unica, ossia non dipende dalla particolare sequenza di operazioni. Tuttavia, questo fatto non è molto importante, sicché ne ometteremo la dimostrazione.

Se A' è ridotta alla forma a scala per righe, possiamo risolvere un sistema di equazioni $A'X = B'$ immediatamente. Per esempio, supponiamo che

$$[A' \mid B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Il sistema $A'X = B'$ non ammette soluzioni, poiché la terza equazione è $0 = 1$. D'altra parte,

$$[A' \mid B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

ammette soluzioni. Scegliendo arbitrariamente x_2, x_4, possiamo risolvere la prima equazione nell'incognita x_1 e la seconda nell'incognita x_3. Questo è il procedimento che abbiamo usato per risolvere il sistema (2.13).

La regola generale è data dalla:

(2.16) PROPOSIZIONE *Sia $M' = [A' \mid B']$ una matrice a scala per righe. Allora il sistema di equazioni $A'X = B'$ ammette soluzioni se e soltanto se l'ultima colonna B' non contiene pivot. In tal caso, può essere assegnato un valore arbitrario all'incognita x_i, se la colonna i-esima non contiene pivot.* ∎

È chiaro che ogni sistema lineare *omogeneo* $AX = 0$ ammette la soluzione banale $X = 0$. Tuttavia, considerando ancora la forma a scala per righe, possiamo concludere che, se vi sono più incognite che equazioni, allora il sistema omogeneo $AX = 0$ ammette una soluzione *non banale*:

(2.17) COROLLARIO *Ogni sistema $AX = 0$ di m equazioni lineari omogenee in n incognite, con $m < n$, ammette una soluzione X in cui almeno un elemento x_i è diverso da zero.*

Infatti, sia $A'X = 0$ il sistema associato alla matrice a scala per righe A' e sia r il numero dei pivot di A', sicché $r \leq m$. Allora, dalla proposizione (2.16) segue che possiamo assegnare valori arbitrari a $n - r$ incognite x_i. ∎

Utilizzeremo ora la riduzione per righe per caratterizzare le matrici quadrate invertibili.

(2.18) PROPOSIZIONE *Sia A una matrice quadrata. Le seguenti condizioni sono tra loro equivalenti:*

(a) *A può essere ridotta alla matrice identica mediante una successione di operazioni elementari sulle righe.*

(b) *A è un prodotto di matrici elementari.*

(c) *A è invertibile.*

(d) *Il sistema di equazioni lineari omogenee $AX = 0$ ammette soltanto la soluzione banale $X = 0$.*

Dimostrazione. Procederemo dimostrando le implicazioni (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ $\Rightarrow$ (d) $\Rightarrow$ (a). Per far vedere che (a) implica (b), supponiamo che A sia riducibile alla matrice identica mediante operazioni sulle righe: $E_k \cdots E_1 A = I$. Moltiplicando ambo i membri di questa equazione a sinistra per $E_1^{-1} \cdots E_k^{-1}$, si ottiene $A = E_1^{-1} \cdots E_k^{-1}$. Poiché l'inversa di una matrice elementare è una matrice elementare, ciò prova che A è un prodotto di matrici elementari. Inoltre, poiché un prodotto di matrici elementari è invertibile, (b) implica (c). Se A è invertibile, possiamo moltiplicare ambo i membri dell'equazione matriciale $AX = 0$ per A^{-1} e ottenere così $X = 0$. Dunque il sistema di equazioni omogenee $AX = 0$ ammette soltanto la soluzione banale. Ciò prova che (c) implica (d).

Per dimostrare l'ultima implicazione, ossia (d) $\Rightarrow$ (a), consideriamo una matrice *quadrata* a scala per righe M. Notiamo la seguente dicotomia:

(2.19) *Sia M una matrice quadrata a scala per righe. Allora, o M è la matrice identica, oppure la sua ultima riga è nulla.*

Ciò si vede facilmente (cfr. (2.15)).

Supponiamo che la proprietà (a) non valga per una data matrice A. Allora A può essere ridotta, mediante operazioni sulle righe, ad una matrice A' avente l'ultima riga nulla. In tal caso, il sistema lineare $A'X = 0$ contiene al più $n - 1$ equazioni non banali e pertanto, in virtù del corollario (2.17), ammette una soluzione non banale. Poiché il sistema $AX = 0$ è equivalente al sistema $A'X = 0$, anch'esso ammette una soluzione non banale. Ne segue che, se (a) non vale, neppure (d) vale; dunque (d) implica (a). Ciò completa la dimostrazione della proposizione (2.18). ■

(2.20) COROLLARIO *Se una matrice quadrata A contiene una riga nulla, A non è invertibile.* ■

La riduzione per righe fornisce un metodo per calcolare l'inversa di una matrice invertibile A. In pratica, si riduce A all'identità mediante operazioni sulle righe:

$$E_k \cdots E_1 A = I,$$

come sopra; moltiplicando poi ambo i membri di tale equazione a destra per A^{-1}, si ha:

$$E_k \cdots E_1 I = A^{-1}.$$

Possiamo enunciare pertanto il risultato seguente:

(2.21) COROLLARIO *Sia A una matrice invertibile. Per calcolare la sua inversa A^{-1}, basta effettuare operazioni elementari sulle righe $E_1, \ldots, E_k$ su A, riducendola all'identità. La stessa successione di operazioni, applicata a I, fornisce A^{-1}.* ■

(2.22) *Esempio*

Cerchiamo l'inversa della matrice

$$A = \begin{bmatrix} 5 & 4 \\ 6 & 5 \end{bmatrix}.$$

Per calcolarla, formiamo la matrice a blocchi 2×4

$$[A \mid I] = \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 6 & 5 & 0 & 1 \end{array}\right].$$

Effettuiamo operazioni sulle righe per ridurre A all'identità, considerando sempre il blocco a destra, così da ottenere alla fine A^{-1} a destra, in virtù del corollario (2.21).

$$[A \mid I] = \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 6 & 5 & 0 & 1 \end{array}\right] \qquad \text{Sottrarre (riga 1) da (riga 2)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 1 & 1 & -1 & 1 \end{array}\right] \qquad \text{Sottrarre } 4 \cdot \text{(riga 2) da (riga 1)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & -4 \\ 1 & 1 & -1 & 1 \end{array}\right] \qquad \text{Sottrarre (riga 1) da (riga 2)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & -4 \\ 0 & 1 & -6 & 5 \end{array}\right] = [I \mid A^{-1}].$$

Dunque

$$A^{-1} = \begin{bmatrix} 5 & -4 \\ -6 & 5 \end{bmatrix}.$$

(2.23) PROPOSIZIONE *Sia A una matrice quadrata, dotata di un'inversa sinistra B ($BA = I$), oppure di un'inversa destra ($AB = I$). Allora A è invertibile e B è la sua inversa.*

Dimostrazione. Supponiamo che $AB = I$. Effettuiamo la riduzione per righe su A. In base a (2.19), esistono matrici elementari $E_1, \ldots, E_k$ tali che $A' = E_k \ldots E_1 A$ è la matrice identica oppure ha l'ultima riga nulla. Allora $A'B = E_k \ldots E_1$, la quale è una matrice invertibile. Quindi l'ultima riga di $A'B$ non è nulla e, di conseguenza, tale risulta anche l'ultima riga di A'. Dunque $A' = I$. Per la (2.18), A è invertibile, e le equazioni $I = E_k \ldots E_1 A$ e $AB = I$ mostrano che $A^{-1} = E_k \ldots E_1 = B$ (cfr. (1.17)). Vediamo ora il caso $BA = I$. Scambiando A e B nell'argomentazione precedente si conclude che B è invertibile e A è la sua inversa. Dunque anche A è invertibile. ∎

Per la maggior parte di questa trattazione, avremmo potuto considerare le colonne piuttosto che le righe. Abbiamo scelto di lavorare con le righe per applicare i risultati ai sistemi di equazioni lineari; altrimenti avremmo potuto anche usare le colonne. L'operazione che scambia tra loro righe e colonne è la *trasposizione* di matrici. La *trasposta* di una matrice $m \times n$ A è la matrice $n \times m$ $A^{\mathrm{t}} = (b_{ij})$, dove

$$b_{ij} = a_{ji}.$$

Per esempio,

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}^{\mathrm{t}} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \quad \text{e} \quad [1 \quad 2 \quad 3]^{\mathrm{t}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Le regole di calcolo per la matrice trasposta sono date in:

(2.24) (a) $(A + B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$.

(b) $(cA)^{\mathrm{t}} = cA^{\mathrm{t}}$.

(c) $(AB)^{\mathrm{t}} = B^{\mathrm{t}} A^{\mathrm{t}}$ (!)

(d) $(A^{\mathrm{t}})^{\mathrm{t}} = A$.

Usando le formule (2.24c, d), possiamo dedurre le proprietà della *moltiplicazione a destra*, XP, dalle corrispondenti proprietà per la moltiplicazione a sinistra.

Le matrici elementari (2.6) agiscono mediante moltiplicazione a destra come le seguenti *operazioni elementari sulle colonne*:

(2.25) (a) *Sommare* $a \cdot$ (colonna i) *alla* (colonna j).

(b) *Scambiare tra loro la* (colonna i) *e la* (colonna j).

(c) *Moltiplicare la* (colonna i) *per* $c \neq 0$.

3 Determinanti

Ad ogni matrice quadrata A è associato un numero chiamato il suo *determinante*. In questo paragrafo definiremo il determinante e dimostreremo alcune sue proprietà. Il determinante di una matrice A sarà denotato con det A.

$$= \mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\,S\,\text{---} \\ \text{---}(S-R)\text{---} \\ \vdots \end{bmatrix} = \mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\,S\,\text{---} \\ \text{---}(-R)\text{---} \\ \vdots \end{bmatrix} = -\mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\,S\,\text{---} \\ \text{---}\,R\,\text{---} \\ \vdots \end{bmatrix}.$$

(3.7′) *Se due righe di una matrice A sono uguali, allora* $\mathrm{d}(A) = 0$.

Infatti, scambiando righe adiacenti più volte, si ottiene una matrice A' avente due righe adiacenti uguali. In base a (3.7), si ha: $\mathrm{d}(A') = 0$, e da (3.9) segue che $\mathrm{d}(A) = \pm\mathrm{d}(A')$.

Utilizzando (3.7′), le dimostrazioni di (3.8) e (3.9) forniscono i seguenti risultati:

(3.8′) *Aggiungendo un multiplo di una riga a un'altra riga, il determinante non cambia.*

(3.9′) *Scambiando tra loro due righe, il determinante viene moltiplicato per* -1.

Inoltre, da (3.6) discende il risultato seguente:

(3.10) *Se una riga di A è nulla, allora* $\mathrm{d}(A) = 0$.

Infatti, se una riga è nulla, allora la matrice A non cambia moltiplicando tale riga per 0. Ma allora, in base a (3.6), $\mathrm{d}(A)$ viene moltiplicato per 0. Dunque $\mathrm{d}(A) = 0 \cdot \mathrm{d}(A) = 0$.

Le regole (3.8′), (3.9′) e (3.6) descrivono l'effetto sul determinante di un'operazione elementare sulle righe (2.7), sicché esse possono essere riscritte in termini di matrici elementari. Esse si enunciano dicendo che $\mathrm{d}(EA) = \mathrm{d}(A)$, se E è una matrice elementare del primo tipo, che $\mathrm{d}(EA) = -\mathrm{d}(A)$, se E è del secondo tipo, e [cfr. (3.6)] che $\mathrm{d}(EA) = c \cdot \mathrm{d}(A)$, se E è del terzo tipo. Applichiamo ora tali regole per calcolare $\mathrm{d}(E)$, essendo E una matrice elementare. Poniamo $A = I$. Allora, poiché $\mathrm{d}(I) = 1$, le regole determinano $\mathrm{d}(EI) = \mathrm{d}(E)$:

(3.11) Il determinante di una matrice elementare è:

(i) Primo tipo (*aggiungere un multiplo di una riga a un'altra*): $\mathrm{d}(E) = 1$, in virtù di (3.8′).

(ii) Secondo tipo (*scambio di righe*): $\mathrm{d}(E) = -1$, in virtù di (3.9′).

(iii) Terzo tipo (*moltiplicare una riga per una costante non nulla*): $\mathrm{d}(E) = c$, in virtù di (3.6).

Inoltre, se utilizziamo ancora le regole (3.8′), (3.9′) e (3.6), applicandole a una matrice arbitraria A e usando i valori per $\mathrm{d}(E)$ che abbiamo appena determinato, otteniamo il seguente risultato:

(3.12) *Sia E una matrice elementare e sia A una matrice arbitraria. Allora si ha:*

$$\mathrm{d}(EA) = \mathrm{d}(E) \cdot \mathrm{d}(A).$$

Ricordiamo che, in base a (2.19), ogni matrice quadrata A può essere ridotta, mediante operazioni elementari sulle righe, a una matrice A', la quale o è l'identità I oppure ha l'ultima riga nulla:

$$A' = E_k \cdots E_1 A.$$

Sappiamo da (3.5) e (3.10) che si ha, rispettivamente: $\mathrm{d}(A') = 1$ oppure $\mathrm{d}(A') = 0$. Utilizzando (3.12) e l'induzione, si ottiene:

(3.13) $$\mathrm{d}(A') = \mathrm{d}(E_k) \cdots \mathrm{d}(E_1)\mathrm{d}(A).$$

D'altra parte, conosciamo $\mathrm{d}(E_i)$, in virtù di (3.11), e pertanto possiamo utilizzare questa formula per calcolare $\mathrm{d}(A)$.

(3.14) TEOREMA (Definizione assiomatica del determinante) *La funzione determinante* (3.4) *è l'unica funzione che soddisfa alle proprietà* (3.5-3.7).

Dimostrazione. Abbiamo usato soltanto tali proprietà per giungere alle equazioni (3.11) e (3.13), ed esse determinano $\mathrm{d}(A)$. Poiché lo sviluppo per minori (3.4) soddisfa (3.5-3.7), esso coincide con (3.13). ∎

Ritorniamo ora alla notazione usuale $\det A$ per il determinante di una matrice.

(3.15) COROLLARIO *Una matrice quadrata A è invertibile se e soltanto se* $\det A \neq 0$.

Infatti, ciò segue dalle formule (3.11), (3.13) e (2.18). Da (3.11) segue che $\det E_i \neq 0$ per ogni i. Pertanto, se A' è come in (3.13), allora $\det A \neq 0$ se e soltanto se $\det A' \neq 0$, e ciò accade se e soltanto se $A' = I$. In base a (2.18), $A' = I$ se e soltanto se A è invertibile. ∎

Siamo in grado ora di dimostrare una delle proprietà più importanti della

(3.16) TEOREMA *Siano A, B due matrici $n \times n$ arbitrarie. Allora si ha*

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

Dimostrazione: Notiamo che, se A è una matrice elementare, questa formula coincide con (3.12).

Caso 1: A è invertibile. In base a (2.18b), A è un prodotto di matrici elementari: $A = E_1 \cdots E_k$. Utilizzando (3.12) e l'induzione, si ha: $\det A = (\det E_1) \cdots (\det E_k)$ e

$$\det(AB) = \det(E_1 \cdots E_k B) = (\det E_1) \cdots (\det E_k)(\det B) = (\det A)(\det B).$$

Caso 2: A non è invertibile. Allora, in virtù di (3.15), si ha: $\det A = 0$; pertanto, per dimostrare il teorema in questo caso, basta far vedere che $\det(AB) = 0$. In base a (2.18), A può essere ridotta a una matrice $A' = E_k \cdots E_1 A$ avente l'ultima riga nulla. Allora anche l'ultima riga di $A'B$ è nulla; pertanto si ha:

$$0 = \det(A'B) = \det(E_k \cdots E_1 AB) = (\det E_k) \cdots (\det E_1)(\det AB).$$

Poiché $\det E_i \neq 0$, ne segue che $\det(AB) = 0$. ∎

(3.17) COROLLARIO *Se A è invertibile, si ha:*

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

Dimostrazione. Basta osservare che $(\det A)(\det A^{-1}) = \det I = 1$. ∎

Osservazione. Si potrebbe pensare, in modo naturale, di cercare di definire il determinante di una matrice utilizzando le regole (3.11) e (3.16). Tali regole determinano certamente $\det A$ per ogni matrice invertibile A, poiché possiamo scrivere una matrice siffatta come un prodotto di matrici elementari. Tuttavia c'è un problema: vi sono molti modi di scrivere una matrice assegnata come un prodotto di matrici elementari e non è affatto chiaro che due prodotti diversi avrebbero lo stesso determinante. In realtà, non è per niente facile sviluppare questa idea fino in fondo.

La dimostrazione della proposizione seguente è un buon esercizio.

(3.18) PROPOSIZIONE *Denotiamo con A^t la trasposta di A. Allora si ha*

$$\det A = \det A^t. \blacksquare$$

(3.19) COROLLARIO *Le proprietà* (3.6-3.10) *continuano a valere, se la parola* riga *è sostituita ovunque dalla parola* colonna. ■

4 Matrici di permutazione

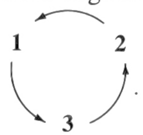
Un'applicazione biiettiva p di un insieme S in sé è una *permutazione* dell'insieme:

(4.1) $$p : S \longrightarrow S.$$

Per esempio,

(4.2) $$\begin{aligned} \mathbf{1} &\mapsto \mathbf{3} \\ \mathbf{2} &\mapsto \mathbf{1} \\ \mathbf{3} &\mapsto \mathbf{2} \end{aligned}$$

è una permutazione dell'insieme $\{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$. Essa è una permutazione *ciclica*, poiché opera nel modo seguente:



Vi sono parecchi modi di denotare le permutazioni. In questo paragrafo useremo la notazione funzionale, cosicché $p(\mathbf{x})$ denota il valore della permutazione p sull'elemento $\mathbf{x}$. Così, se p è la permutazione data in (4.2), allora

$$p(\mathbf{1}) = \mathbf{3}, \quad p(\mathbf{2}) = \mathbf{1}, \quad p(\mathbf{3}) = \mathbf{2}.$$

Una *matrice di permutazione* P è una matrice con la seguente proprietà: l'operazione di moltiplicazione a sinistra per P è una permutazione delle righe di una matrice. Le matrici elementari del secondo tipo (2.6ii) costituiscono gli esempi più semplici. Esse corrispondono alle permutazioni chiamate *trasposizioni*, le quali scambiano tra loro due righe di una matrice, lasciando fisse le altre. Inoltre,

(4.3) $$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

è una matrice di permutazione. Essa agisce su un vettore colonna $X = (x, y, z)^t$ nel modo seguente:

$$PX = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}.$$