CAPITOLO 1

Matrici e Sistemi di Equazioni Lineari

Il primo obiettivo di questo Capitolo è quello di introdurre le matrici a coefficienti complessi e le loro principali operazioni: la somma di due matrici, il prodotto di uno scalare per una matrice e il prodotto righe per colonne di due matrici. In relazione a queste operazioni vengono poi studiati gli operatori di trasposizione e di H-trasposizione, nonché le decomposizioni a blocchi delle matrici, di cui sono illustrati gli esempi più utilizzati.

Il secondo obiettivo del Capitolo è quello di fornire l'algoritmo dell'eliminazione di Gauss per risolvere i sistemi di equazioni lineari; è questo il metodo più antico e più semplice per risolvere un sistema lineare, ma al tempo stesso il più efficiente. Come applicazione notevole dell'eliminazione di Gauss viene presentato l'algoritmo di inversione di una matrice quadrata. Un intero paragrafo è dedicato al problema della ricerca di matrici inverse di una matrice qualunque; si introduce anche la matrice pseudo-inversa, di cui sono provate esistenza e unicità.

La parte finale del Capitolo è dedicata allo studio delle matrici elementari e al loro impiego nella decomposizione LU di una generica matrice. La decomposizione LU è uno dei pilastri portanti dell'Algebra Lineare. Le matrici elementari saranno usate nel Capitolo IV nel corso della trattazione dei determinanti.

1. Matrici e loro operazioni

Questo paragrafo introduce le principali operazioni che si possono eseguire sulle matrici: somma e prodotto di due matrici, prodotto di una matrice per uno scalare. Queste operazioni si basano sulle usuali operazioni di somma e prodotto tra numeri, dove con "numeri" si intenderanno sempre elementi dell'insieme $\mathbb{C}$ dei numeri complessi, o, in particolare, del suo sottoinsieme $\mathbb{R}$ dei numeri reali, o del sottoinsieme $\mathbb{Q}$ dei numeri razionali. I numeri complessi saranno chiamati anche *scalari*, e saranno denotati con lettere greche: α, β, γ, eccetera.

Una *matrice* è una tabella di numeri (o di simboli che rappresentano numeri) disposti in righe e colonne; tali numeri, detti *coefficienti* (o *entrate*) della matrice, sono per comodità racchiusi usualmente tra parentesi. L'unico coefficiente della matrice che si trova nella i-esima riga e nella j-esima colonna si chiama *coefficiente di posto* (i, j). Diremo anche che una matrice con m righe ed n colonne ha *dimensioni* $m \times n$, o semplicemente che è una matrice $m \times n$.

Esempio 1.1. Le seguenti tabelle di numeri

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 2 \\ -\pi \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

costituiscono esempi, rispettivamente, di matrici 2×2 a coefficienti complessi, 2×1 a coefficienti reali e 3×4 a coefficienti razionali. $\square$

Designeremo sempre le matrici con lettere maiuscole "in nero". Una generica matrice $\mathbf{A}$ $m \times n$, in cui sono messe in evidenza la i-esima riga e la j-esima colonna, ha il seguente aspetto:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix};$$

si noti che ogni coefficiente di $\mathbf{A}$ viene contrassegnato da due indici, di cui il primo dice in quale riga si trova il coefficiente e il secondo in quale colonna.

In maniera più compatta, tale matrice verrà anche denotata nel modo seguente

$$\mathbf{A} = [a_{ij}]_{j \leq n}^{i \leq m}$$

o più semplicemente con $\mathbf{A} = [a_{ij}]$ se le dimensioni di $\mathbf{A}$ sono note.

Introduciamo fin dall'inizio alcune speciali classi di matrici.

Matrici quadrate. Sono quelle matrici $m \times n$ per cui $m = n$, tali cioè che il numero delle righe è uguale al numero delle colonne. Tale numero viene chiamato *ordine* della matrice. Nel caso di matrici quadrate si può parlare della *diagonale (principale)* della matrice, costituita dai coefficienti di posto (i, i) $(1 \leq i \leq m)$, che sono detti per l'appunto *coefficienti diagonali.* Si considera anche la *diagonale secondaria*, costituita dai coefficienti di posto $(i, m-i+1)$ $(1 \leq i \leq m)$. La seguente matrice

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix}$$

è un esempio di matrice quadrata di ordine 3, con diagonale $(1, 6, 11)$ e diagonale secondaria $(3, 6, 9)$.

Matrici diagonali. Sono quelle matrici quadrate che hanno tutti i coefficienti al di fuori della diagonale uguali a 0, quindi del tipo

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}.$$

Una tale matrice sarà denotata con $\mathbf{Diag}(d_1, d_2, d_3, \dots, d_n)$. Ovviamente basta un solo indice per contrassegnare i coefficienti sulla diagonale.

Matrici scalari. Una matrice scalare è una matrice diagonale in cui tutti i coefficienti diagonali sono uguali tra di loro, quindi del tipo $\mathbf{Diag}(d, d, d, \dots, d)$. Una tale matrice è individuata dal suo ordine e dal coefficiente che compare sulla diagonale. Tra tutte le matrici, le matrici scalari sono quelle che più si avvicinano ai numeri, e questo anche dal punto di vista delle operazioni (si veda l'Esercizio 1).

Matrici triangolari superiori. Sono quelle matrici quadrate in cui i coefficienti *al di sotto* della diagonale sono nulli. Una tale matrice $\mathbf{A} = [a_{ij}]$ è pertanto definita dalle relazioni

$$a_{ij} = 0 \quad \text{per } i > j.$$

Le matrici triangolari superiori saranno spesso denotate con il simbolo $\mathbf{T}$ o con il simbolo $\mathbf{U}$ (che sta per *upper*). Una generica matrice triangolare superiore 3×3 è quindi del tipo

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}.$$

Se tutti i coefficienti diagonali sono uguali a 1, la matrice si chiama *uni-triangolare superiore.*

Matrici triangolari inferiori. Sono quelle matrici quadrate in cui i coefficienti *al di sopra* della diagonale sono nulli. Una tale matrice $\mathbf{A} = [a_{ij}]$ è pertanto definita dalle relazioni

$$a_{ij} = 0 \quad \text{per } i < j.$$

Una generica matrice triangolare inferiore 3×3 è quindi del tipo

$$\begin{bmatrix} t_{11} & 0 & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & t_{33} \end{bmatrix}.$$

Le matrici triangolari inferiore saranno spesso denotate anche con il simbolo $\mathbf{T}$ o con il simbolo $\mathbf{L}$ (che sta per *lower*). Se tutti i coefficienti diagonali sono uguali a 1, la matrice si chiama *uni-triangolare inferiore.*

Matrici nulle. Una matrice $m \times n$ in cui *tutti* i coefficienti sono uguali a 0 è detta matrice *nulla* ed è denotata col simbolo $\mathbf{O}_{mn}$; gli indici saranno omessi, e la matrice si denoterà semplicemente con $\mathbf{O}$, se sarà chiaro dal contesto quali sono le dimensioni della matrice.

Matrici riga e colonna. Una matrice $1 \times n$, quindi con una sola riga, si chiama *matrice riga*, mentre una matrice $m \times 1$, quindi con una sola colonna, si chiama *matrice colonna.* Le matrici colonna saranno chiamate anche *vettori (colonna)* (con il termine tra parentesi usualmente omesso), e denotati anche con lettere minuscole "in nero": $\mathbf{v}$, $\mathbf{u}$, $\mathbf{w}$, eccetera. Le matrici riga invece saranno chiamate *vettori riga* e denotate, per motivi che appariranno evidenti più avanti, coi seguenti simboli: $\mathbf{v}^T$, $\mathbf{u}^T$, $\mathbf{w}^T$, eccetera. I coefficienti di un vettore (riga o colonna) saranno anche chiamati *coordinate.*

Un esempio di vettore 2×1 è fornito dalla matrice $\mathbf{B}$ nell'Esempio 1.1. Un generico vettore riga con n coordinate sarà scritto nella forma

$$\mathbf{v}^T = [v_1\ v_2\ \dots\ v_n];$$

evidentemente, in un vettore (riga o colonna) basta un solo indice per contrassegnare le coordinate.

Vettori coordinati. Tra i vettori (colonna) ce ne sono alcuni di particolare utilità; essi sono chiamati *vettori coordinati* e sono caratterizzati dal fatto di avere tutte le coordinate nulle tranne una che è uguale a 1; se la coordinata non nulla è la i-esima, il vettore coordinato è denotato con $\mathbf{e}_i$; l'analogo vettore riga coordinato è denotato con $\mathbf{e}_i^T$. Si ha quindi

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \mathbf{e}_i^T = [0\ 0\ \dots\ 0\ 1\ 0\ \dots\ 0]$$

dove gli 1 si trovano all'i-esimo posto.

Introduciamo ora le due più semplici operazioni che si possono eseguire sulle matrici: il prodotto per scalari e la somma.

Prodotto per scalari. Data una matrice $m \times n$ $\mathbf{A} = [a_{ij}]$ e uno scalare α, si definisce la matrice *prodotto dello scalare* α *per la matrice* $\mathbf{A}$ come quella matrice, denotata con $\alpha\mathbf{A}$, che ha come coefficiente di posto (i, j) il numero αa_{ij}; pertanto

$$\alpha\mathbf{A} = [\alpha a_{ij}].$$

Si può definire analogamente il prodotto di $\mathbf{A}$ per lo scalare α come la matrice $\mathbf{A}\alpha = [a_{ij}\alpha]$; tale matrice coincide evidentemente con $\alpha\mathbf{A}$, attesa la commutatività del prodotto tra scalari, pertanto $\alpha\mathbf{A} = \mathbf{A}\alpha$.

Esempio 1.2. Se $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ sono le matrici dell'Esempio 1.1, si ha

$$i\mathbf{A} = \begin{bmatrix} i-1 & 2i+1 \\ 3i & 1 \end{bmatrix}, \qquad \pi^{-1}\mathbf{B} = \begin{bmatrix} 2\pi^{-1} \\ -1 \end{bmatrix}, \qquad -1\mathbf{C} = \begin{bmatrix} -1 & -2 & -3 & -4 \\ -5 & -6 & -7 & -8 \\ -9 & -10 & -11 & -12 \end{bmatrix}$$

□

La matrice $-1\mathbf{C}$ nell'Esempio 1.2 viene denotata più semplicemente con $-\mathbf{C}$. In generale, data la matrice $\mathbf{A} = [a_{ij}]$, si pone

$$-\mathbf{A} = [-a_{ij}];$$

$-\mathbf{A}$ si chiama *matrice opposta* di $\mathbf{A}$. Si osservi che la matrice prodotto di una data matrice per uno scalare produce una matrice delle stesse dimensioni della matrice data. L'operazione di prodotto per scalare gode delle seguenti proprietà, la cui verifica è del tutto evidente:

(p_1) $1\mathbf{A} = \mathbf{A}$;
(p_2) $0\mathbf{A} = \mathbf{O}$;
(p_3) $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$.

Somma di matrici. Date due matrici $\mathbf{A} = [a_{ij}]$ e $\mathbf{B} = [b_{ij}]$ delle medesime dimensioni $m \times n$, si definisce come loro *somma* la matrice

$$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}].$$

Quindi la somma tra matrici è definita nel modo più naturale possibile, e viene detta *somma per componenti.* Sottolineiamo il fatto che, con la definizione data sopra, non si può eseguire la somma tra due matrici che hanno dimensioni diverse (cioè il numero delle righe oppure quello delle colonne, o entrambi, sono differenti).

Esempio 1.3. Date le due matrici 2×2

$$\mathbf{A} = \begin{bmatrix} i-1 & 2i+1 \\ 3i & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} i+1 & 2i-1 \\ i & -1 \end{bmatrix}$$

risulta

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}2i & 4i\\ 4i & 0\end{bmatrix}. \qquad \square$$

ESEMPIO 1.4. La somma di due matrici scalari (rispettivamente: diagonali, triangolari superiori, triangolari inferiori) delle stesse dimensioni è ancora una matrice scalare (rispettivamente: diagonale, triangolare superiore, triangolare inferiore). $\square$

L'operazione di somma tra matrici gode di numerose proprietà, anche in relazione a quella di prodotto per scalari, che elenchiamo di seguito. La verifica di queste proprietà è del tutto evidente.

Se $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ sono tre matrici di dimensioni $m \times n$, e α, β sono scalari, allora

(S_1) $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$,
(S_2) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$,
(S_3) $\mathbf{A}+\mathbf{O}_{mn}=\mathbf{A}$,
(S_4) $\mathbf{A}+(-\mathbf{A})=\mathbf{O}_{mn}$,
(pS_1) $\alpha(\mathbf{A}+\mathbf{B})=\alpha\mathbf{A}+\alpha\mathbf{B}$,
(pS_2) $(\alpha+\beta)\mathbf{A}=\alpha\mathbf{A}+\beta\mathbf{A}$.

Si possono definire diversi tipi di prodotto tra matrici. Ad esempio, si può definire l'operazione di prodotto nel modo più semplice possibile, cioè *per componenti*, in modo analogo a quanto fatto con la somma; tale operazione è detta anche *prodotto di Hadamard-Schur*. Un altro modo più complicato è tramite il cosiddetto *prodotto di Kronecker*, o *prodotto tensoriale*. Pur essendo tali prodotti interessanti per numerose applicazioni, essi sono meno fondamentali del prodotto che ora andremo a studiare e non saranno qui trattati.

Passiamo quindi a definire l'operazione di prodotto tra due matrici di gran lunga più utilizzata, che viene detta *prodotto righe per colonne*. Partiremo dal caso più semplice del prodotto di una sola riga per una sola colonna, per passare poi al caso generale.

Prodotto di vettore riga per vettore colonna. Dati un vettore riga $\mathbf{v}^T$ e un vettore colonna $\mathbf{u}$ con lo stesso numero di coordinate, tali cioè che le loro dimensioni siano rispettivamente $1 \times n$ e $n \times 1$, denotiamo con $v_1, v_2, \dots, v_n$ le coordinate di $\mathbf{v}^T$ e con $u_1, u_2, \dots, u_n$ le coordinate di $\mathbf{u}$. Si chiama *prodotto (riga per colonna)* dei due vettori, e lo si denota con $\mathbf{v}^T\mathbf{u}$, il numero (o matrice 1×1)

$$\mathbf{v}^T\mathbf{u} = v_1u_1 + v_2u_2 + \cdots + v_nu_n.$$

Le seguenti proprietà del prodotto di un vettore riga per un vettore colonna sono del tutto evidenti: $\mathbf{v}^T\mathbf{0}=0$; $\mathbf{v}^T\mathbf{u}=\mathbf{u}^T\mathbf{v}$.

ESEMPIO 1.5. Dato un vettore riga $\mathbf{v}^T = [v_1\ v_2\ \dots\ v_n]$, il prodotto di $\mathbf{v}^T$ per l'i-esimo vettore coordinato $\mathbf{e}_i$ coincide con v_i, la i-esima coordinata di $\mathbf{v}^T$. Il prodotto di $\mathbf{v}^T$ per il vettore colonna $\mathbf{u}$ che ha n coordinate tutte uguali a 1 è

$$\mathbf{v}^T\mathbf{u} = v_1 + v_2 + \ldots + v_n.$$

Il prodotto di $\mathbf{v}^T \neq \mathbf{0}^T$ per un vettore colonna $\mathbf{w} \neq \mathbf{0}$ può dare 0; per esempio, se $\mathbf{v}^T = [1\ 0]$ e $\mathbf{w}^T = [0\ 1]$. $\square$

Prodotto di matrici righe per colonne. Consideriamo ora due matrici $\mathbf{A}$ e $\mathbf{B}$ e definiamo il loro *prodotto righe per colonne*. Va detto subito che

> *il prodotto righe per colonne di* $\mathbf{A}$ *per* $\mathbf{B}$ *è possibile solo se il numero di colonne di* $\mathbf{A}$ *coincide col numero di righe di* $\mathbf{B}$.

Deve quindi risultare $\mathbf{A}$ di dimensioni $m \times n$ e $\mathbf{B}$ di dimensioni $n \times p$. Due matrici che si trovano in questa situazione rispetto alle dimensioni si dicono *conformi per il prodotto.*

La nozione base per definire questo prodotto è quella appena vista di prodotto di vettore riga per vettore colonna. È utile a tal fine porre in evidenza le righe della matrice $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix}.$$

Se $\mathbf{A} = [a_{ij}]$, allora $\mathbf{r}_i^T = [a_{i1}\ a_{i2}\ \dots\ a_{in}]$. Mettiamo anche in evidenza le colonne della matrice $\mathbf{B}$:

$$\mathbf{B} = [\mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_p]$$

dove ciascun vettore colonna $\mathbf{b}_j$ ha n coordinate. Se $\mathbf{B} = (b_{ij})$, allora

$$\mathbf{b}_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \dots \\ b_{nj} \end{bmatrix}.$$

Visto che il numero di coordinate di ciascun vettore riga $\mathbf{r}_i^T$ è uguale al numero delle coordinate di ciascun vettore colonna $\mathbf{b}_j$, è possibile eseguire il prodotto $\mathbf{r}_i^T \mathbf{b}_j$. La matrice prodotto righe per colonne della matrice $\mathbf{A}$ per la matrice $\mathbf{B}$, che viene denotata con $\mathbf{AB}$, è la matrice di dimensioni $m \times p$ definita nel modo seguente:

$$\mathbf{AB} = [\mathbf{r}_i^T \mathbf{b}_j]_{j \leq p}^{i \leq m}.$$

ESEMPIO 1.6. Date le due matrici

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} i \\ -1 \end{bmatrix}$$

risulta

$$\mathbf{AB} = \begin{bmatrix} (1+i)i + (2-i)(-1) \\ 3i + i \end{bmatrix} = \begin{bmatrix} -3+2i \\ 4i \end{bmatrix},$$

mentre il prodotto $\mathbf{BA}$ non si può eseguire. □

D'ora in avanti, quando parleremo di prodotto tra due matrici, intenderemo sempre il prodotto righe per colonne. Questo modo di definire il prodotto sembra alquanto artificioso. Tale artificiosità scompare, non appena si pensi a rappresentare tramite tale prodotto un sistema di m equazioni lineari in n incognite:

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

Il sistema si può rappresentare semplicemente tramite l'uguaglianza del vettore (detto *vettore dei termini noti*)

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

con il vettore che ha come coordinate i primi membri delle m equazioni

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}.$$

Quest'ultimo vettore non è altro che il risultato del prodotto righe per colonne della matrice (detta *matrice dei coefficienti*)

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

per il vettore (detto *vettore delle incognite*)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}.$$

Pertanto il sistema di equazioni si può scrivere in forma compatta, detta *forma matriciale* del sistema:

$$\mathbf{Ax} = \mathbf{b}.$$

Diamo qui di seguito le più importanti proprietà cui soddisfa il prodotto tra matrici, anche in relazione alle due operazioni di somma e di prodotto per scalari.

(P$_1$) Se $\mathbf{A}, \mathbf{B}$ e $\mathbf{C}$ sono tre matrici di dimensioni rispettivamente $m \times n$, $n \times p$ e $p \times q$, allora

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Sia $[a_{ij}] = \mathbf{A}$, $[b_{ij}] = \mathbf{B}$ e $[c_{ij}] = \mathbf{C}$. Sia inoltre $[d_{ij}] = \mathbf{A}(\mathbf{BC})$ e $[e_{ij}] = (\mathbf{AB})\mathbf{C}$. Per la proprietà distributiva risulta:

$$d_{ij} = \sum_{1 \le h \le n} a_{ih} \sum_{1 \le k \le p} b_{hk} c_{kj} = \sum_{1 \le k \le p} c_{kj} \sum_{1 \le h \le n} a_{ih} b_{hk} = e_{ij}.$$

(P$_2$) Se $\mathbf{A}$ è una matrice $n \times p$, allora

$$\mathbf{O}_{mn}\mathbf{A} = \mathbf{O}_{mp}, \qquad \mathbf{AO}_{pq} = \mathbf{O}_{nq}.$$

Per dare la proprietà successiva dobbiamo introdurre un nuovo tipo di matrici scalari: fissato $n \ge 1$, denotiamo con $\mathbf{I}_n$ la matrice scalare $n \times n$ che ha sulla diagonale il numero 1:

$$\mathbf{I}_n = \mathbf{Diag}(1, 1, 1, \dots, 1) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Le matrici $\mathbf{I}_n$ sono chiamate *matrici identità*, per il motivo che appare chiaro dalla proprietà seguente.

(P$_3$) Se $\mathbf{A}$ è una matrice $n \times p$, allora

$$\mathbf{I}_n\mathbf{A} = \mathbf{A} = \mathbf{AI}_p.$$

Questa proprietà è un caso particolare del seguente Esempio 1.8 e la sua dimostrazione è lasciata al lettore.

(PS$_1$) Se $\mathbf{A}$ è una matrice $m \times n$ e $\mathbf{B}$ e $\mathbf{C}$ sono matrici $n \times p$, allora

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

Sia $[d_{ij}] = \mathbf{A}(\mathbf{B} + \mathbf{C})$ e $[e_{ij}] = \mathbf{AB} + \mathbf{AC}$. Risulta

$$d_{ij} = \sum_{1 \le k \le n} a_{ik}(b_{kj} + c_{kj}) = \sum_{1 \le k \le n} a_{ik} b_{kj} + \sum_{1 \le k \le n} a_{ik} c_{kj} = e_{ij}.$$

(PS_2) Se $\mathbf{A}$ e $\mathbf{B}$ sono matrici $m \times n$ e $\mathbf{C}$ è una matrice $n \times p$, allora

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}.$$

La dimostrazione è analoga a quella della proprietà precedente.

(Pp) Se $\mathbf{A}$ e $\mathbf{B}$ sono matrici conformi per il prodotto e α è uno scalare, allora

$$\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}).$$

La dimostrazione è lasciata come facile esercizio per il lettore.

Esempio 1.7. Il lettore provi come esercizio che il prodotto di matrici scalari (rispettivamente: diagonali, triangolari superiori, triangolari inferiori) dello stesso ordine è ancora una matrice scalare (rispettivamente: diagonale, triangolare superiore, triangolare inferiore), con coefficienti diagonali i prodotti dei corrispondenti coefficienti diagonali dei due fattori. □

I due seguenti esempi descrivono il risultato del prodotto di una matrice per una matrice diagonale, prima moltiplicata a sinistra (*pre-moltiplicazione*) e poi a destra (*post-moltiplicazione*).

Esempio 1.8. Siano

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix}$$

una matrice $m \times n$ in cui sono in evidenza le righe e $\mathbf{D} = \mathbf{Diag}(d_1, d_2, \dots, d_m)$ una matrice diagonale. Una facile verifica mostra che

$$\mathbf{DA} = \begin{bmatrix} d_1\mathbf{r}_1^T \\ d_2\mathbf{r}_2^T \\ \dots \\ d_m\mathbf{r}_m^T \end{bmatrix},$$

quindi la pre-moltiplicazione di $\mathbf{A}$ per una matrice diagonale $\mathbf{D}$ ha l'effetto di moltiplicare ogni riga di $\mathbf{A}$ per il corrispondente elemento diagonale di $\mathbf{D}$.

Per esempio, se

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

e $\mathbf{D} = \mathbf{Diag}(1,2,3)$, risulta

$$\mathbf{DA} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 10 & 12 & 14 & 16 \\ 27 & 30 & 33 & 36 \end{bmatrix}.$$ □

Esempio 1.9. Siano

$$\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$$

una matrice $m \times n$ in cui sono in evidenza le colonne e $\mathbf{D} = \mathbf{Diag}(d_1, d_2, \dots, d_n)$ una matrice diagonale. Si verifica che

$$\mathbf{AD} = [d_1\mathbf{a}_1 \ d_2\mathbf{a}_2 \ \dots \ d_n\mathbf{a}_n]$$

quindi la post-moltiplicazione di $\mathbf{A}$ per una matrice diagonale $\mathbf{D}$ ha l'effetto di moltiplicare ogni colonna di $\mathbf{A}$ per il corrispondente elemento diagonale di $\mathbf{D}$.

Per esempio, se la matrice $\mathbf{A}$ è quella dell'esempio precedente e se si pone $\mathbf{D} = \mathbf{Diag}(1,2,3,4)$, risulta

$$\mathbf{AD} = \begin{bmatrix} 1 & 4 & 9 & 16 \\ 5 & 12 & 21 & 32 \\ 9 & 20 & 33 & 48 \end{bmatrix}.$$ □

Una cosa da dire sul prodotto tra matrici, altrettanto importante che elencarne le proprietà, è quali sono le proprietà cui non soddisfa. A tal fine la cosa più opportuna da fare è quella di fornire esempi di prodotti di particolari matrici che non soddisfano a quelle proprietà. La proprietà più importante cui il prodotto non soddisfa è quella commutativa.

Esempio 1.10. Il prodotto di due matrici quadrate non gode in generale della proprietà commutativa. Prendiamo per esempio due generiche matrici 2×2:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \mathbf{A}' = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}.$$

Allora risulta

$$\mathbf{A}\mathbf{A}' = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}, \qquad \mathbf{A}'\mathbf{A} = \begin{bmatrix} a'a + b'c & a'b + b'd \\ c'a + d'c & c'b + d'd \end{bmatrix}$$

per cui $\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A}$ se e solo se

$$bc' = b'c, \quad ab' + bd' = a'b + b'd, \quad ca' + dc' = c'a + d'c.$$

Il lettore trovi esempi numerici opportuni di matrici 2×2 $\mathbf{A}$ e $\mathbf{A}'$ che non commutano, tali cioè che $\mathbf{A}\mathbf{A}' \neq \mathbf{A}'\mathbf{A}$. □

Data una matrice quadrata $\mathbf{A}$, se ne possono considerare le potenze successive:

$$\mathbf{A}^0 = \mathbf{I}, \qquad \mathbf{A}^1 = \mathbf{A}, \qquad \mathbf{A}^2 = \mathbf{A}\mathbf{A}, \qquad \mathbf{A}^3 = \mathbf{A}\mathbf{A}\mathbf{A}, \qquad \dots$$

Col crescere dell'esponente k, la matrice $\mathbf{A}^k$ può diventare più complicata oppure più semplice rispetto ad $\mathbf{A}$, come mostrano i tre esempi seguenti.

Esempio 1.11. Si consideri la matrice 3×3:

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

che ha cinque coefficienti nulli. Il lettore verifichi che nelle successive potenze $\mathbf{A}^2$, $\mathbf{A}^3$ e $\mathbf{A}^4$ i coefficienti nulli diminuiscono progressivamente, e che $\mathbf{A}^k$ non ha coefficienti nulli per ogni $k \geq 5$. □

Esempio 1.12. Si consideri la matrice 3×3:

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

Si verifichi che $\mathbf{P}^3 = \mathbf{I}_3$. Quindi le successive potenze di $\mathbf{P}$ riproducono ciclicamente le prime tre potenze. □

Esempio 1.13. Si consideri la matrice 3×3:

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

Il lettore verifichi che $\mathbf{J}^2$ ha un unico coefficiente non nullo uguale a 1 e che $\mathbf{J}^k = \mathbf{O}$ per ogni $k \geq 3$. Questo esempio prova anche che in generale il prodotto di matrici non nulle può annullarsi. □

Due matrici $\mathbf{A}$ e $\mathbf{B}$ si possono sia sommare che moltiplicare tra di loro se e solo se sono quadrate dello stesso ordine n. Denoteremo con $\mathrm{M}_n(\mathbb{C})$ l'insieme di tutte le matrici a coefficienti complessi di ordine n, dotato delle operazioni di prodotto per scalari, somma e prodotto (righe per colonne). Chiameremo $\mathrm{M}_n(\mathbb{C})$ l'*algebra* delle matrici complesse $n \times n$. Analogamente, per le matrici a coefficienti reali o razionali, useremo le notazioni $\mathrm{M}_n(\mathbb{R})$ e $\mathrm{M}_n(\mathbb{Q})$ e parleremo dell'algebra delle matrici reali o razionali di ordine n.

Altre notazioni universalmente usate sono quelle di $\mathbb{C}^n$ e $\mathbb{C}_n$, che denotano rispettivamente l'insieme dei vettori colonna e dei vettori riga con n coordinate, dotati del prodotto per scalari e della somma. Analogamente, per i vettori a coordinate reali (o razionali), useremo le notazioni $\mathbb{R}^n$ (o $\mathbb{Q}^n$) e $\mathbb{R}_n$ (o $\mathbb{Q}_n$).

Infine, denoteremo con $\mathrm{M}_{n\times p}(\mathbb{C})$ (rispettivamente, con $\mathrm{M}_{n\times p}(\mathbb{R})$ e $\mathrm{M}_{n\times p}(\mathbb{Q})$) l'insieme delle matrici $n \times p$ a coefficienti complessi (rispettivamente, reali e razionali).

2. Trasposte e H-trasposte, decomposizioni a blocchi

Data una matrice $m \times n$ $\mathbf{A} = [a_{ij}]$, si chiama *trasposta* di $\mathbf{A}$, e la si indica con $\mathbf{A}^T$, la matrice che si ottiene da $\mathbf{A}$ scambiandone le righe con le colonne. Pertanto, il coefficiente di posto (i, j) della matrice trasposta $\mathbf{A}^T$ coincide con a_{ji}. È evidente dalla definizione che la trasposta $\mathbf{A}^T$ ha dimensioni $n \times m$.

Esempio 2.1. Riprendiamo le matrici dell'Esempio 1.1 cui affianchiamo le rispettive trasposte:

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}, \qquad \mathbf{A}^T = \begin{bmatrix} 1+i & 3 \\ 2-i & -i \end{bmatrix};$$

$$\mathbf{B} = \begin{bmatrix} 2 \\ -\pi \end{bmatrix}, \qquad \mathbf{B}^T = [2 \ \ -\pi];$$

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}, \qquad \mathbf{C}^T = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{bmatrix}. \qquad \square$$

Esempio 2.2. La matrice trasposta di un vettore colonna $\mathbf{v}$ non è altro che il vettore riga $\mathbf{v}^T$, che è stato denotato in questo modo nel paragrafo precedente proprio per utilizzare la notazione della trasposizione. □

Per le informazioni basilari sull'operazione di coniugazione dei numeri complessi rimandiamo all'Appendice II. La matrice *coniugata* della matrice $\mathbf{A} = [a_{ij}]$, denotata con $\overline{\mathbf{A}}$, è la matrice ottenuta coniugando tutti i coefficienti di $\mathbf{A}$:

$$\overline{\mathbf{A}} = [\overline{a_{ij}}].$$

Evidentemente la coniugata di una matrice reale coincide con la matrice stessa, e la matrice coniugata di $\overline{\mathbf{A}}$ coincide con $\mathbf{A}$ stessa. Effettuando prima l'operazione di trasposizione e poi quella di coniugazione sulla matrice $\mathbf{A}$ si ottiene la medesima matrice che invertendo l'ordine delle due operazioni; tale matrice si denota con $\mathbf{A}^H$ e si chiama matrice *H-trasposta* di $\mathbf{A}$. Si ha pertanto

$$\mathbf{A}^H = \overline{\mathbf{A}}^T = \overline{\mathbf{A}^T}.$$

Il simbolo "H" deriva dal matematico francese Hermite. Ovviamente, se $\mathbf{A}$ è una matrice reale, $\mathbf{A}^H = \mathbf{A}^T$.

Esempio 2.3. La matrice H-trasposta della matrice $\mathbf{A}$ dell'Esempio 1.1 è

$$\mathbf{A}^H = \begin{bmatrix} 1-i & 3 \\ 2+i & i \end{bmatrix}.$$

Le matrici H-trasposte delle matrici $\mathbf{B}$ e $\mathbf{C}$ dello stesso Esempio 1.1 coincidono con $\mathbf{B}^T$ e $\mathbf{C}^T$, rispettivamente. □

Elenchiamo di seguito le principali proprietà di cui godono le matrici trasposte e H-trasposte; le dimostrazioni di queste proprietà sono tutte banali, a eccezione di quella della proprietà (T_4).

Siano $\mathbf{A}$ e $\mathbf{B}$ matrici delle medesime dimensioni, $\mathbf{C}$ una matrice conforme ad $\mathbf{A}$ per il prodotto, e α uno scalare. Allora

(T_1) $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$;
(T_2) $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$;
(T_3) $(\mathbf{A}^T)^T = \mathbf{A}$;
(T_4) $(\mathbf{AC})^T = \mathbf{C}^T\mathbf{A}^T$;
(H_1) $(\alpha\mathbf{A})^H = \overline{\alpha}\mathbf{A}^H$;
(H_2) $(\mathbf{A}+\mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H$;
(H_3) $(\mathbf{A}^H)^H = \mathbf{A}$;
(H_4) $(\mathbf{AC})^H = \mathbf{C}^H\mathbf{A}^H$.

Dimostriamo la (T_4). Sia $\mathbf{A} = [a_{ij}]_{j\le n}^{i\le m}$ e $\mathbf{C} = [c_{jh}]_{h\le p}^{j\le n}$. L'elemento di posto (i,j) di $(\mathbf{AC})^T$ coincide con quello di posto (j,i) di $\mathbf{AC}$, che è $\sum_{1\le h\le n} a_{jh}c_{hi}$. D'altra parte, l'elemento di posto (i,j) di $\mathbf{C}^T\mathbf{A}^T$ è dato dal prodotto della i-esima riga di $\mathbf{C}^T$ per la j-esima colonna di $\mathbf{A}^T$, quindi è $\sum_{1\le h\le n} c_{hi}a_{jh}$, da cui l'asserto.

È importante rilevare che nelle proprietà (T_4) e (H_4) l'ordine dei due fattori viene invertito passando alle trasposte e H-trasposte. La proprietà (H_4) si ricava immediatamente dalla proprietà (T_4) e dal fatto che $\overline{\mathbf{AC}} = \overline{\mathbf{A}}\,\overline{\mathbf{C}}$, il che risulta evidente non appena si ricordi che il coniugato del prodotto di due numeri complessi è il prodotto dei loro coniugati.

Matrici simmetriche ed hermitiane. Tra tutte le matrici quadrate, quelle che coincidono con la loro trasposta rivestono un particolare interesse; esse vengono chiamate matrici *simmetriche*. Pertanto, una matrice $n \times n$ $\mathbf{A}$ è simmetrica se $\mathbf{A} = \mathbf{A}^T$. In altri termini, una matrice quadrata risulta simmetrica quando, ruotata di mezzo giro attorno alla sua diagonale, coincide con sé stessa.

La matrice $\mathbf{A}$ è invece detta *hermitiana* se coincide con la sua H-trasposta, cioè se $\mathbf{A} = \mathbf{A}^H$.

È evidente che per le matrici reali le due nozioni di matrice simmetrica e di matrice hermitiana vengono a coincidere, giacché la coniugazione non modifica una matrice reale. Invece una matrice complessa non reale può risultare simmetrica senza essere hermitiana, e viceversa.

Mentre i coefficienti diagonali di una matrice simmetrica possono essere del tutto arbitrari, quelli di una matrice hermitiana devono essere numeri reali, perché devono coincidere con i loro coniugati.

Ci sono altri due tipi importanti di matrici, collegate alle operazioni di trasposizione e H-trasposizione. Diremo che una matrice $\mathbf{A}$ è *anti-simmetrica* se $\mathbf{A} = -\mathbf{A}^T$; diremo invece che è *anti-hermitiana* se $\mathbf{A} = -\mathbf{A}^H$.

Segue immediatamente dalla definizione che gli elementi diagonali di una matrice anti-simmetrica sono nulli, mentre gli elementi diagonali di una matrice anti-hermitiana sono tutti immaginari, cioè del tipo ir, con $r \in \mathbb{R}$.

Esempio 2.4. Le seguenti quattro matrici offrono esempi, ordinatamente, di matrice simmetrica, hermitiana, anti-simmetrica e anti-hermitiana:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \quad \text{simmetrica;}$$

$$\mathbf{B} = \begin{bmatrix} 2 & 1-i & i \\ 1+i & 0 & 3+2i \\ -i & 3-2i & 3 \end{bmatrix} \quad \text{hermitiana;}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 4 \\ -3 & -4 & 0 \end{bmatrix} \quad \text{anti-simmetrica;}$$

$$\mathbf{D} = \begin{bmatrix} i & 1-i & i \\ -1-i & 0 & 3+2i \\ i & -3+2i & -2i \end{bmatrix} \quad \text{anti-hermitiana.} \qquad \square$$

In analogia con la forma algebrica di un numero complesso, per cui ogni numero $z \in \mathbb{C}$ si scrive in uno e un solo modo come $z = a + ib$, con a e b numeri reali (a è detto parte reale e ib parte immaginaria di z), per ogni matrice complessa quadrata si ha la seguente decomposizione.

Proposizione 2.5. *Ogni matrice complessa quadrata* $\mathbf{A}$ *si scrive in uno e un solo modo nella forma* $\mathbf{A} = \mathbf{B} + \mathbf{C}$*, con* $\mathbf{B}$ *matrice hermitiana e* $\mathbf{C}$ *matrice anti-hermitiana.*

Dimostrazione. Si ponga $\mathbf{B} = (\mathbf{A} + \mathbf{A}^H)/2$ e $\mathbf{C} = (\mathbf{A} - \mathbf{A}^H)/2$. Una verifica diretta mostra che $\mathbf{B}$ è hermitiana, $\mathbf{C}$ è anti-hermitiana e $\mathbf{A} = \mathbf{B} + \mathbf{C}$. Quanto all'unicità, se $\mathbf{A} = \mathbf{B}' + \mathbf{C}'$, con $\mathbf{B}'$ hermitiana e $\mathbf{C}'$ anti-hermitiana, risulta

$$\mathbf{B} - \mathbf{B}' = \mathbf{C}' - \mathbf{C}.$$

Ma $\mathbf{B} - \mathbf{B}'$ risulta hermitiana, mentre $\mathbf{C}' - \mathbf{C}$ risulta anti-hermitiana; l'unica matrice contemporaneamente hermitiana e anti-hermitiana è la matrice nulla, percio $\mathbf{B} = \mathbf{B}'$ e $\mathbf{C} = \mathbf{C}'$. $\square$

Le matrici $\mathbf{B}$ e $\mathbf{C}$ che compaiono nella Proposizione 2.5 si chiamano, rispettivamente, *parte hermitiana* e *parte anti-hermitiana* della matrice complessa $\mathbf{A}$ (se $\mathbf{A}$ è una matrice reale, allora la parte hermitiana risulta reale simmetrica e la parte anti-hermitiana risulta reale anti-simmetrica). Esiste un motivo ben preciso per cui c'è analogia tra matrici hermitiane e numeri reali da un lato e matrici anti-hermitiane e numeri immaginari dall'altro, che potrà essere compreso solo al Capitolo VI quando avremo a disposizione la teoria degli autovalori.

L'analogia con i numeri complessi cade se si chiede alla parte hermitiana di commutare con la parte anti-hermitiana, come invece accade per a e ib nel numero complesso $z = a + ib$. Ciò infatti caratterizza una particolare classe di matrici, chiamate normali, come mostra la seguente Proposizione. Una matrice $\mathbf{A}$ si dice *normale* se $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$, cioè se $\mathbf{A}$ commuta con la sua H-trasposta. Le matrici normali sono molto importanti per un'altra loro proprietà, che si vedrà al Capitolo VI nel Teorema Spettrale.

Proposizione 2.6. *Sia* $\mathbf{A} = \mathbf{B} + \mathbf{C}$ *la decomposizione della matrice quadrata* $\mathbf{A}$ *nella parte hermitiana* $\mathbf{B}$ *e nella parte anti-hermitiana* $\mathbf{C}$*. Allora* $\mathbf{BC} = \mathbf{CB}$ *se e solo se* $\mathbf{A}$ *è normale.*

Dimostrazione. Poiché $\mathbf{A}^H = \mathbf{B}^H + \mathbf{C}^H = \mathbf{B} - \mathbf{C}$, si ha che $\mathbf{A}\mathbf{A}^H = \mathbf{B}^2 - \mathbf{BC} + \mathbf{CB} - \mathbf{C}^2$, mentre $\mathbf{A}^H\mathbf{A} = \mathbf{B}^2 + \mathbf{BC} - \mathbf{CB} - \mathbf{C}^2$. Perciò $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ se e solo se $-\mathbf{BC} + \mathbf{CB} = \mathbf{BC} - \mathbf{CB}$, se e solo se $\mathbf{BC} = \mathbf{CB}$. $\square$

ESEMPIO 2.7. Data la matrice complessa

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}$$

la sua decomposizione in parte hermitiana e parte anti-hermitiana è:

$$\mathbf{A} = [(\mathbf{A} + \mathbf{A}^H)/2] + [(\mathbf{A} - \mathbf{A}^H)/2] = \frac{1}{2}\begin{bmatrix} 2 & 5-i \\ 5+i & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2i & -1-i \\ 1-i & -2i \end{bmatrix};$$

il lettore verifichi che parte hermitiana e parte anti-hermitiana non commutano, ovvero che $\mathbf{A}$ non è normale. □

Diamo alcune proprietà cui soddisfano le matrici simmetriche ed hermitiane, anti-simmetriche e anti-hermitiane. La loro verifica è lasciata come esercizio.

(a) La somma di due matrici simmetriche, hermitiane, anti-simmetriche o anti-hermitiane è dello stesso tipo.
(b) Se $\mathbf{A}$ è una matrice simmetrica (risp., hermitiana) e α è uno scalare (risp., un numero reale), allora $\alpha\mathbf{A}$ è una matrice simmetrica (risp., hermitiana).
(c) Se $\mathbf{A}$ è una matrice anti-simmetrica (risp., anti-hermitiana) e α è uno scalare (risp., un numero reale), allora $\alpha\mathbf{A}$ è una matrice anti-simmetrica (risp., anti-hermitiana).
(d) $\mathbf{A}$ è una matrice hermitiana se e solo se $i\mathbf{A}$ è una matrice anti-hermitiana.
(e) Il prodotto di due matrici simmetriche (risp., hermitiane) è una matrice simmetrica (risp., hermitiana) se e solo se le due matrici commutano tra di loro.

    È importante rilevare che il prodotto di due matrici simmetriche (risp., hermitiane) che non commutano tra di loro non risulta mai una matrice simmetrica (risp., hermitiana). La proprietà seguente prende in considerazione una matrice del tutto arbitraria, e da essa ricava due matrici simmetriche, oppure hermitiane, che risulteranno molto usate in futuro.
(f) Sia $\mathbf{X}$ una matrice complessa $m \times n$. Le matrici quadrate $\mathbf{X}\mathbf{X}^T$ e $\mathbf{X}^T\mathbf{X}$ sono entrambi simmetriche e le matrici $\mathbf{X}\mathbf{X}^H$ e $\mathbf{X}^H\mathbf{X}$ sono entrambi hermitiane.

Si osservi che le matrici $\mathbf{X}\mathbf{X}^T$ e $\mathbf{X}^T\mathbf{X}$ hanno in generale dimensioni del tutto differenti. Vediamo un esempio.

ESEMPIO 2.8. Sia $\mathbf{v}^T = [v_1\ v_2\ \ldots\ v_n]$ un vettore riga con n coordinate. Allora la matrice $\mathbf{v}\mathbf{v}^T$ è una matrice simmetrica $n \times n$, con $v_i v_j$ come coefficiente di posto (i,j), mentre la matrice $\mathbf{v}^T\mathbf{v}$ ha dimensioni 1×1, giacché è lo scalare $\sum_{i \le n} v_i^2$. Ad esempio, se $\mathbf{v}^T = [1\ 0\ 2\ -1]$, $\mathbf{v}^T\mathbf{v} = 6$, mentre

$$\mathbf{v}\mathbf{v}^T = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 4 & -2 \\ -1 & 0 & -2 & 1 \end{bmatrix}.$$ □

Sottomatrici. Data una matrice $m \times n$ $\mathbf{A}$, se ne possono considerare delle sue "porzioni". Per esempio, se scegliamo ad arbitrio alcune righe e alcune colonne di $\mathbf{A}$, i coefficienti che stanno nella intersezione di queste righe e colonne formano quella che si chiama una *sottomatrice* di $\mathbf{A}$. La matrice $\mathbf{A}$ è naturalmente una sottomatrice di sé stessa; le sottomatrici diverse da $\mathbf{A}$ sono dette sottomatrici *proprie*.

ESEMPIO 2.9. Si consideri la seguente matrice 4×4:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}.$$

Ciascuna delle seguenti matrici è sottomatrice di $\mathbf{A}$:

$$\begin{bmatrix}1 & 2\\ 9 & 10\end{bmatrix};\quad \begin{bmatrix}1 & 3 & 4\\ 9 & 11 & 12\end{bmatrix};\quad \begin{bmatrix}2 & 4\\ 6 & 8\\ 14 & 16\end{bmatrix};\quad \begin{bmatrix}5 & 6 & 7 & 8\\ 13 & 14 & 15 & 16\end{bmatrix}. \qquad \square$$

Tra le sottomatrici di una matrice quadrata risultano particolarmente utili e usate le *sottomatrici principali*, che si ottengono scegliendo righe e colonne (la cui intersezione produce la sottomatrice) con gli stessi indici. È facile convincersi che i coefficienti diagonali di una sottomatrice principale sono coefficienti diagonali anche nella matrice originaria.

Esempio 2.10. Nessuna delle sottomatrici della matrice $\mathbf{A}$ considerate nel precedente esempio è una sottomatrice principale. Le seguenti sottomatrici invece lo sono:

$$\begin{bmatrix}1 & 3\\ 9 & 11\end{bmatrix};\quad \begin{bmatrix}6 & 7\\ 10 & 11\end{bmatrix};\quad \begin{bmatrix}6 & 7 & 8\\ 10 & 11 & 12\\ 14 & 15 & 16\end{bmatrix}. \qquad \square$$

In certi casi (per esempio nello studio delle matrici hermitiane definite positive (si veda il Capitolo VI) si considerano solo le sottomatrici principali ottenute intersecando le prime k righe e colonne, dove k varia da 1 all'ordine n della matrice; esse vengono chiamate *sottomatrici principali k-esime*. Naturalmente la matrice stessa è la sottomatrice principale n-esima di sé stessa.

Esempio 2.11. Nessuna delle sottomatrici considerate nell'Esempio 2.10 è una sottomatrice principale k-esima della matrice $\mathbf{A}$. Lo sono invece tutte e sole le seguenti matrici:

$$[1];\quad \begin{bmatrix}1 & 2\\ 5 & 6\end{bmatrix};\quad \begin{bmatrix}1 & 2 & 3\\ 5 & 6 & 7\\ 9 & 10 & 12\end{bmatrix};\quad \mathbf{A}. \qquad \square$$

Decomposizioni a blocchi. Un'altra maniera di considerare "porzioni" di matrici è quella di *decomporre a blocchi* una matrice; questa operazione si esegue tracciando delle righe orizzontali e verticali che "tagliano" la matrice in sottomatrici, ciascuna delle quali è formate dalla intersezione di righe e colonne consecutive ed è chiamata *blocco*. Si ottengono così delle righe e delle colonne di blocchi, e si parla di *dimensioni a blocchi* della matrice. Di seguito sono riportati due diversi modi di decomporre a blocchi la matrice dell'Esempio 2.9.

$$\mathbf{A} = \left[\begin{array}{cc|cc}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ \hline 9 & 10 & 11 & 12\\ 13 & 14 & 15 & 16\end{array}\right] = \begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}$$

dove è evidente quali sono i quattro blocchi $\mathbf{A}_{ij}$. La matrice $\mathbf{A}$ così decomposta ha dimensioni a blocchi 2×2;

$$\mathbf{A} = \left[\begin{array}{c|cc|c}1 & 2 & 3 & 4\\ \hline 5 & 6 & 7 & 8\\ 9 & 10 & 11 & 12\\ \hline 13 & 14 & 15 & 16\end{array}\right] = \begin{bmatrix}\mathbf{A}'_{11} & \mathbf{A}'_{12} & \mathbf{A}'_{13}\\ \mathbf{A}'_{21} & \mathbf{A}'_{22} & \mathbf{A}'_{23}\\ \mathbf{A}'_{31} & \mathbf{A}'_{32} & \mathbf{A}'_{33}\end{bmatrix}$$

dove è chiaro quali sono i nove blocchi $\mathbf{A}'_{ij}$. La matrice $\mathbf{A}$ così decomposta ha dimensioni a blocchi 3×3.

L'utilità delle decomposizioni a blocchi sta nel fatto che sulle matrici si possono eseguire le operazioni di somma e prodotto (righe per colonne) *per blocchi*, cioè considerando i blocchi come fossero dei coefficienti. Naturalmente ciò è possibile solo se sono rispettati i seguenti requisiti:

(B_1) le dimensioni a blocchi delle due matrici devono essere le stesse nella somma a blocchi, e conformi per il prodotto nel prodotto a blocchi;

(B_2) le coppie di blocchi delle due matrici che si sommano tra di loro devono essere delle stesse dimensioni, e quelle che si moltiplicano tra di loro devono essere conformi rispetto al prodotto.

Una volta che tali requisiti sono soddisfatti, si può *operare a blocchi*, ottenendo lo stesso risultato che operando sui coefficienti. La giustificazione di questo fatto è del tutto evidente per la somma, mentre, per quanto riguarda il prodotto, dipende essenzialmente dalla proprietà associativa della somma.

Esempio 2.12. Il prodotto a blocchi delle due matrici decomposte a blocchi

$$\mathbf{A} = \left[\begin{array}{c|cc|c} 1 & 2 & 3 & 1 \\ \hline 0 & -1 & 0 & -2 \\ 3 & 0 & 1 & 2 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \quad \mathbf{B} = \left[\begin{array}{c|c} 2 & 0 \\ \hline -1 & 2 \\ 0 & -1 \\ \hline 3 & -1 \end{array}\right] = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \end{bmatrix}$$

produce la matrice

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{A}_{13}\mathbf{B}_{31} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{A}_{13}\mathbf{B}_{32} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} + \mathbf{A}_{23}\mathbf{B}_{31} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} + \mathbf{A}_{23}\mathbf{B}_{32} \end{bmatrix} \\ &= \left[\begin{array}{c|c} 3 & 0 \\ \hline -5 & 0 \\ 12 & -3 \end{array}\right]. \end{aligned}$$ □

La notazione usuale per una matrice decomposta a blocchi di dimensioni a blocchi $m \times n$ è la seguente:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2n} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_{m1} & \mathbf{A}_{m2} & \dots & \mathbf{A}_{mn} \end{bmatrix}.$$

Se ciascun blocco $\mathbf{A}_{ij}$ ha dimensioni $m_i \times n_j$, le dimensioni della matrice $\mathbf{A}$ sono $m \times n$, dove $m = \sum_i m_i$ e $n = \sum_j n_j$.

Molte definizioni date per le matrici si estendono in modo ovvio alle matrici a blocchi. Così, la matrice a blocchi $m \times n$ $\mathbf{A} = (\mathbf{A}_{ij})$ si dice *triangolare superiore a blocchi* se $\mathbf{A}_{ij} = \mathbf{O}$ per $i > j$. Analogamente si hanno le nozioni di matrice triangolare inferiore a blocchi e matrice diagonale a blocchi. Per queste ultime a volte si usa, anziché l'ovvia notazione

$$\mathbf{Diag}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n),$$

la notazione alternativa

$$\mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \dots \oplus \mathbf{A}_n.$$

Anche l'operazione di trasposizione (e di H-trasposizione) si può eseguire a blocchi: basta scambiare tra di loro le righe e le colonne di blocchi, avendo l'avvertenza di trasporre (e di coniugare) ciascun blocco. Per esempio, data la matrice a blocchi

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}$$

la matrice trasposta è:

$$\mathbf{A}^T = \begin{bmatrix} \mathbf{A}_{11}^T & \mathbf{A}_{21}^T \\ \mathbf{A}_{12}^T & \mathbf{A}_{22}^T \\ \mathbf{A}_{13}^T & \mathbf{A}_{23}^T \end{bmatrix}.$$

Per ottenere la matrice H-trasposta a blocchi basta sostituire l'operatore ^{T} con l'operatore ^{H}.

Un modo di decomporre a blocchi una matrice che viene molto usato anche nelle dimostrazioni di risultati teorici, è quello di mettere in evidenza la prima riga e la prima colonna (oppure l'ultima riga e l'ultima colonna); diremo allora che la matrice è *in forma bordata*. Una generica matrice $m \times n$ $\mathbf{A}$ in forma bordata ha questo aspetto:

$$\mathbf{A} = \begin{bmatrix} a & \mathbf{u}^T \\ \mathbf{v} & \mathbf{B} \end{bmatrix}$$

dove a è il coefficiente di posto $(1,1)$, $\mathbf{u}^T$ è la prima riga privata del primo coefficiente, $\mathbf{v}$ è la prima colonna privata del primo coefficiente, e $\mathbf{B}$ si ottiene da $\mathbf{A}$ eliminando prima riga e prima colonna.

Chiudiamo questo paragrafo mostrando alcune rilevanti conseguenze che si ricavano da decomposizioni a blocchi di matrici fatte in modo opportuno.

Consideriamo due matrici $\mathbf{A}$ e $\mathbf{B}$ conformi per il prodotto, diciamo di dimensioni rispettivamente $m \times n$ e $n \times p$. Mettiamo in evidenza in entrambe righe e colonne:

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix} = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_n], \qquad \mathbf{B} = \begin{bmatrix} \mathbf{s}_1^T \\ \mathbf{s}_2^T \\ \dots \\ \mathbf{s}_n^T \end{bmatrix} = [\mathbf{b}_1\ \mathbf{b}_2\ \dots\ \mathbf{b}_p].$$

Si hanno quindi per ciascuna matrice due diverse decomposizioni a blocchi, la prima a blocchi riga e la seconda a blocchi colonna. Possiamo perciò eseguire la moltiplicazione a blocchi in due modi diversi.

(I) Utilizzando per $\mathbf{A}$ la decomposizione a blocchi riga e per $\mathbf{B}$ la decomposizione a blocchi colonna, si ottiene

$$\mathbf{AB} = (\mathbf{r}_i^T \mathbf{b}_j)_{j \leq p}^{i \leq m}$$

che non è altro che la definizione di prodotto righe per colonne.

(II) Utilizzando per $\mathbf{A}$ la decomposizione a blocchi colonna e per $\mathbf{B}$ la decomposizione a blocchi riga si ottiene

$$\mathbf{AB} = \sum_{i \leq n} \mathbf{a}_i \mathbf{s}_i^T$$

che esprime il prodotto di $\mathbf{A}$ per $\mathbf{B}$ come somma di n matrici, ciascuna prodotto di un vettore colonna per un vettore riga.

Si può anche eseguire la moltiplicazione prima considerando $\mathbf{A}$ come un unico blocco e decomponendo $\mathbf{B}$ in blocchi colonna. Si ricava allora

$$\mathbf{A}[\mathbf{b}_1\ \mathbf{b}_2\ \dots\ \mathbf{b}_p] = [\mathbf{A}\mathbf{b}_1\ \mathbf{A}\mathbf{b}_2\ \dots\ \mathbf{A}\mathbf{b}_p].$$

Simmetricamente, si può considerare $\mathbf{B}$ come un unico blocco e decomporre $\mathbf{A}$ in blocchi riga. Si ricava allora

$$\begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{r}_1^T\mathbf{B} \\ \mathbf{r}_2^T\mathbf{B} \\ \dots \\ \mathbf{r}_m^T\mathbf{B} \end{bmatrix}.$$

Quanto appena visto si può esprimere nel modo seguente.

PROPOSIZIONE 2.13. *Nel prodotto (righe per colonne)* $\mathbf{AB}$ *di due matrici:*

(a) *la j-esima colonna di* $\mathbf{AB}$ *coincide con il prodotto* $\mathbf{A}\mathbf{b}_j$ *di* $\mathbf{A}$ *per la j-esima colonna* $\mathbf{b}_j$ *di* $\mathbf{B}$*;*

(b) *la i-esima riga di* $\mathbf{AB}$ *coincide con il prodotto* $\mathbf{r}_i^T\mathbf{B}$ *della i-esima riga* $\mathbf{r}_i^T$ *di* $\mathbf{A}$ *per* $\mathbf{B}$.

Casi particolari del modo descritto in (II) di eseguire il prodotto si hanno quando la matrice $\mathbf{A}$ è un vettore riga $\mathbf{u}^T$, oppure la matrice $\mathbf{B}$ è un vettore (colonna) $\mathbf{v}$. Nel primo caso si ottiene

$$\mathbf{u}^T\mathbf{B} = u_1\mathbf{r}_1^T + u_2\mathbf{r}_2^T + \ldots + u_m\mathbf{r}_m^T$$

e nel secondo caso si ottiene

$$\mathbf{A}\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + \ldots + v_n\mathbf{a}_n;$$

dove naturalmente gli scalari u_i e v_j sono le coordinate di $\mathbf{u}^T$ e di $\mathbf{v}$, rispettivamente; il che si esprime anche dicendo che $\mathbf{u}^T\mathbf{B}$ è combinazione lineare delle righe di $\mathbf{B}$ con coefficienti le coordinate di $\mathbf{u}^T$, e che $\mathbf{A}\mathbf{v}$ è combinazione lineare delle colonne di $\mathbf{A}$ con coefficienti le coordinate di $\mathbf{v}$.

Se come vettori riga e colonna prendiamo i vettori coordinati $\mathbf{e}_i^T$ ed $\mathbf{e}_j$, rispettivamente, si ricava immediatamente la seguente

PROPOSIZIONE 2.14. *Siano* $\mathbf{A}$ *una matrice* $m \times n$, $\mathbf{e}_i^T$ *un vettore coordinato riga* ($i \leq m$) *ed* $\mathbf{e}_j$ *un vettore coordinato colonna* ($j \leq n$). *Allora* $\mathbf{e}_i^T\mathbf{A}$ *coincide con la* i*-esima riga di* $\mathbf{A}$ *e* $\mathbf{A}\mathbf{e}_j$ *coincide con la* j*-esima colonna di* $\mathbf{A}$. □

3. Eliminazione di Gauss per sistemi di equazioni lineari

L'algoritmo dell'Eliminazione di Gauss (che in breve denoteremo con EG) presentato in questo Paragrafo per risolvere i sistemi di equazioni lineari, attribuito a Gauss (e quindi fatto risalire alla fine del XVIII secolo), ha un suo antesignano nell'algoritmo denominato Fang Ch'eng, che si trova nel libro cinese "Nove capitoli sull'arte matematica", risalente al III secolo A.C.; esso è quindi una antichissima conoscenza dell'umanità. L'idea alla base dell'algoritmo è molto semplice: eliminare progressivamente (da qui il nome di "eliminazione") nelle successive equazioni del sistema sempre più incognite, ottenendo un sistema equivalente a quello di partenza (cioè con le medesime soluzioni), ma che si può risolvere molto facilmente con un metodo di "sostituzione all'indietro".

Il modo moderno di presentare la EG consiste nell'operare direttamente sulla matrice dei coefficienti del sistema, cui si aggiunge come ultima colonna la colonna dei termini noti, cioè su quella che si chiama *matrice aumentata* (o ampliata) del sistema. La matrice del sistema semplificato, cui si perviene al termine della EG, si chiama *forma ridotta (di Gauss)* della matrice aumentata del sistema.

Dato il sistema in m equazioni ed n incognite nella sua forma generale vista nel Paragrafo 1

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

che in forma matriciale si scrive

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

dove $\mathbf{A}$ è la matrice dei coefficienti e $\mathbf{b}$ è il vettore dei termini noti, la matrice aumentata del sistema è la matrice a blocchi

$$[\mathbf{A}\ \mathbf{b}].$$

Ricordiamo fatti ben noti riguardanti i sistemi di equazioni lineari: si ottiene un sistema equivalente a quello dato (cioè con esattamente le stesse soluzioni) se si opera sulle sue equazioni in uno dei tre seguenti modi:

(i) una equazione viene moltiplicata per uno scalare non nullo;
(ii) una equazione viene sostituita con la sua somma con un'altra equazione del sistema moltiplicata per uno scalare non nullo;

(iii) si scambiano tra di loro due equazioni.

Si osservi che nell'operazione (ii) la parte "moltiplicata per uno scalare non nullo" è già considerta nell'operazione (i) e potrebbe essere quindi sottaciuta; di fatto, si applica usualmente proprio l'operazione descritta in (ii). Applicando ripetutamente queste tre operazioni, dette *operazioni elementari*, si possono eliminare progressivamente le incognite nelle successive equazioni. Osserviamo inoltre che per ciascuna operazione elementare ne esiste un'altra che riporta la matrice ottenuta nella sua configurazione originale; pertanto queste due operazioni sono una l'inversa dell'altra.

Gli esempi che seguono servono meglio di ogni discorso generale a illustrare come ciò avviene. Nel primo esempio manterremo ancora l'usuale forma del sistema (con le incognite e le somme) e la modificheremo con la EG; tradurremo poi i vari passaggi nella forma matriciale. Dall'esempio successivo opereremo direttamente sulle matrici aumentate.

Esempio 3.1. Si consideri il sistema 3×3 (cioè di tre equazioni in tre incognite)

$$\begin{cases} 2x_1 + 6x_2 + 3x_3 = 4 \\ x_1 - 2x_2 + \dfrac{1}{2}x_3 = 1 \\ -x_1 + x_2 - \dfrac{1}{2}x_3 = \dfrac{2}{5} \end{cases}$$

che ha come matrice aumentata

$$[\mathbf{A}\ \mathbf{b}] = \left[\begin{array}{ccc|c} 2 & 6 & 3 & 4 \\ 1 & -2 & \frac{1}{2} & 1 \\ -1 & 1 & -\frac{1}{2} & \frac{2}{5} \end{array}\right].$$

Moltiplichiamo la prima equazione per $1/2$ ottenendo il sistema equivalente

$$\begin{cases} x_1 + 3x_2 + \dfrac{3}{2}x_3 = 2 \\ x_1 - 2x_2 + \dfrac{1}{2}x_3 = 1 \\ -x_1 + x_2 - \dfrac{1}{2}x_3 = \dfrac{2}{5}. \end{cases}$$

Sommando ora alla seconda equazione la prima moltiplicata per -1, e sommando alla terza equazione la prima moltiplicata per 1, si ricava il nuovo sistema equivalente

$$\begin{cases} x_1 + 3x_2 + \dfrac{3}{2}x_3 = 2 \\ -5x_2 - x_3 = -1 \\ 4x_2 + x_3 = \dfrac{12}{5}. \end{cases}$$

Moltiplichiamo in questo sistema la seconda equazione per $-1/5$, ottenendo il sistema equivalente

$$\begin{cases} x_1 + 3x_2 + \dfrac{3}{2}x_3 = 2 \\ x_2 + \dfrac{1}{5}x_3 = \dfrac{1}{5} \\ 4x_2 + x_3 = \dfrac{12}{5}. \end{cases}$$

Nel sistema così modificato, sommando alla terza equazione la seconda moltiplicata per -4, si ricava il sistema equivalente

$$\begin{cases} x_1 + 3x_2 + \dfrac{3}{2}x_3 = 2 \\ \qquad x_2 + \dfrac{1}{5}x_3 = \dfrac{1}{5} \\ \qquad\qquad \dfrac{1}{5}x_3 = \dfrac{8}{5}. \end{cases}$$

Infine, moltiplicando la terza equazione per 5, si perviene al sistema equivalente

$$\begin{cases} x_1 + 3x_2 + \dfrac{3}{2}x_3 = 2 \\ \qquad x_2 + \dfrac{1}{5}x_3 = \dfrac{1}{5} \\ \qquad\qquad x_3 = 8. \end{cases}$$

A questo punto la EG è terminata. Le soluzioni del sistema si possono facilmente trovare partendo dalla terza equazione, che fornisce il valore di x_3, poi sostituendo tale valore nella seconda equazione e ricavando il valore di x_2, poi sostituendo tali valori nella prima equazione e ricavando il valore di x_1. Si dice che si esegue la *sostituzione all'indietro*. Si ricava così la soluzione:

$$x_3 = 8, \qquad x_2 = -\frac{7}{5}, \qquad x_1 = -\frac{29}{5}.$$

Scriviamo le matrici aumentate dei sistemi ottenuti modificando quello originario

$$\left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 1 & -2 & \frac{1}{2} & 1 \\ -1 & 1 & -\frac{1}{2} & \frac{2}{5} \end{array}\right]; \quad \left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & -5 & -1 & -1 \\ 0 & 4 & 1 & \frac{12}{5} \end{array}\right]; \quad \left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 4 & 1 & \frac{12}{5} \end{array}\right];$$
$$\left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & \frac{1}{5} & \frac{8}{5} \end{array}\right]; \quad \left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 1 & 8 \end{array}\right] = [\mathbf{U}\ \mathbf{c}].$$

Si può notare che le operazioni fatte sulle equazioni dei sistemi si potevano fare direttamente sulle righe delle matrici aumentate, rendendo più semplici le notazioni. Abbiamo denotato la matrice a blocchi ottenuta alla fine con $[\mathbf{U}\ \mathbf{c}]$; $\mathbf{U}$ si chiama forma ridotta (di Gauss) della matrice $\mathbf{A}$, e $[\mathbf{U}\ \mathbf{c}]$ è forma ridotta della matrice aumentata $[\mathbf{A}\ \mathbf{b}]$. Si noti che la matrice $\mathbf{U}$ è uni-triangolare superiore. □

Esempio 3.2. Modifichiamo leggermente il sistema dell'esempio precedente, aggiungendo una incognita:

$$\begin{cases} 2x_1 + 6x_2 + 3x_3 + 2x_4 = 4 \\ x_1 - 2x_2 + \dfrac{1}{2}x_3 + \dfrac{9}{4}x_4 = 1 \\ -x_1 + x_2 - \dfrac{1}{2}x_3 - x_4 = \dfrac{2}{5}. \end{cases}$$

Il sistema ha come matrice aumentata

$$[\mathbf{B}\ \mathbf{b}] = \left[\begin{array}{cccc|c} 2 & 6 & 3 & 2 & 4 \\ 1 & -2 & \frac{1}{2} & \frac{9}{4} & 1 \\ -1 & 1 & -\frac{1}{2} & -1 & \frac{2}{5} \end{array}\right].$$

Le medesime operazioni elementari eseguite nell'Esempio 3.1 producono successivamente le matrici

$$\left[\begin{array}{cccc|c} 1 & 3 & \frac{3}{2} & 1 & 2 \\ 1 & -2 & \frac{1}{2} & \frac{9}{4} & 1 \\ -1 & 1 & -\frac{1}{2} & -1 & \frac{2}{5} \end{array}\right]; \quad \left[\begin{array}{cccc|c} 1 & 3 & \frac{3}{2} & 1 & 2 \\ 0 & -5 & -1 & \frac{5}{4} & -1 \\ 0 & 4 & 1 & 0 & \frac{12}{5} \end{array}\right];$$

$$\left[\begin{array}{cccc|c} 1 & 3 & \frac{3}{2} & 1 & 2 \\ 0 & 1 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{5} \\ 0 & 4 & 1 & 0 & \frac{12}{5} \end{array}\right]; \quad \left[\begin{array}{cccc|c} 1 & 3 & \frac{3}{2} & 1 & 2 \\ 0 & 1 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{5} \\ 0 & 0 & \frac{1}{5} & 1 & \frac{8}{5} \end{array}\right];$$

$$\left[\begin{array}{cccc|c} 1 & 3 & \frac{3}{2} & 1 & 2 \\ 0 & 1 & \frac{1}{5} & -\frac{1}{4} & \frac{1}{5} \\ 0 & 0 & 1 & 5 & 8 \end{array}\right] = [\mathbf{V}\ \mathbf{c}].$$

L'ultima riga di $[\mathbf{V}\ \mathbf{c}]$ corrisponde all'equazione

$$x_3 + 5x_4 = 8.$$

Diamo all'incognita x_4 un qualunque valore h; ne risulta

$$x_3 = 8 - 5h, \quad x_2 = -\frac{7}{5} + \frac{5}{4}h, \quad x_1 = -\frac{29}{5} + \frac{11}{4}h;$$

pertanto il sistema ha infinite soluzioni dipendenti dal parametro h, che è conveniente scrivere in questa forma:

$$\begin{bmatrix} -\frac{29}{5} \\ -\frac{7}{5} \\ 8 \\ 0 \end{bmatrix} + h \begin{bmatrix} \frac{11}{4} \\ \frac{5}{4} \\ -5 \\ 1 \end{bmatrix}. \qquad \square$$

Le matrici in forma ridotta $\mathbf{U}$ e $\mathbf{V}$ cui si è pervenuti nei due esempi precedenti hanno una configurazione che viene detta *a scala per righe*; essa è caratterizzata dal fatto che ogni riga a partire dalla seconda ha un numero di zeri iniziali superiore alla riga precedente (qualora questa non sia la riga nulla, nel qual caso è essa stessa nulla).

Per un sistema che ha la matrice aumentata in forma a scala per righe è chiaro che le soluzioni si trovano facilmente con la sostituzione all'indietro, dando possibilmente valori arbitrari a certe variabili.

Esempio 3.3. Modifichiamo in altro modo il sistema dell'Esempio 3.1, cambiando il coefficiente della variabile x_3 nella terza equazione. Consideriamo quindi il sistema

$$\begin{cases} 2x_1 + 6x_2 + \ \ 3x_3 = 4 \\ \ \ x_1 - 2x_2 + \ \frac{1}{2}x_3 = 1 \\ -x_1 + \ \ x_2 - \frac{7}{10}x_3 = \frac{2}{5} \end{cases}$$

che ha come matrice aumentata

$$[\mathbf{A}' \ \mathbf{b}] = \left[\begin{array}{ccc|c} 2 & 6 & 3 & 4 \\ 1 & -2 & \frac{1}{2} & 1 \\ -1 & 1 & -\frac{7}{10} & \frac{2}{5} \end{array}\right].$$

Le operazioni elementari eseguite nell'Esempio 3.1 producono le matrici

$$\left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 1 & -2 & \frac{1}{2} & 1 \\ -1 & 1 & -\frac{7}{10} & \frac{2}{5} \end{array}\right]; \quad \left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & -5 & -1 & -1 \\ 0 & 4 & \frac{4}{5} & \frac{12}{5} \end{array}\right];$$

$$\left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 4 & \frac{4}{5} & \frac{12}{5} \end{array}\right]; \quad \left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & \frac{8}{5} \end{array}\right];$$

$$\left[\begin{array}{ccc|c} 1 & 3 & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 1 \end{array}\right] = [\mathbf{U}' \ \mathbf{c}].$$

L'ultima riga corrisponde all'equazione

$$0x_4 = 1$$

che evidentemente non ha soluzione. Quindi il sistema $\mathbf{A}'\mathbf{x} = \mathbf{b}$ non ha soluzioni. □

Nei tre esempi precedenti non si è mai resa necessaria l'operazione elementare (iii) di scambio di due equazioni. Vediamo allora un esempio in cui bisogna eseguire tale operazione.

Esempio 3.4. Consideriamo il sistema $\mathbf{C}\mathbf{x} = \mathbf{d}$, dove

$$[\mathbf{C} \ \mathbf{d}] = \left[\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 5 & -1 & 3 & 4 \end{array}\right].$$

Sommando alla seconda riga l'opposto della prima, e alla terza riga la prima moltiplicata per -2, si ricava la matrice

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \end{array}\right].$$

È evidente che per trasformare la matrice in forma a scala per righe bisogna scambiare tra di loro la seconda e la terza riga; con questo scambio di righe si ha la matrice

$$[\mathbf{W}\ \mathbf{w}] = \left[\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

che è in forma ridotta. La variabile cui diamo in questo caso un valore arbitrario k è x_3; le soluzioni dipendenti da tale parametro sono, come si verifica facilmente

$$x_1 = -2 - 12k; \quad x_2 = 1 + 5k; \quad x_3 = k; x_4 = 1$$

che vanno convenientemente scritte nella forma

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix} + k \begin{bmatrix} -12 \\ 5 \\ 1 \\ 0 \end{bmatrix}. \qquad \square$$

Osservando il diverso comportamento delle matrici in forma ridotta cui si perviene nei quattro esempi precedenti, si capisce subito da cosa dipende il fatto che un sistema abbia una e una sola soluzione, oppure infinite soluzioni, oppure nessuna soluzione; non esitono infatti altre possibilità (si dice anche che i sistemi di equazioni lineari sono di tipo 0, 1, ∞).

È utile a tal fine introdurre la seguente definizione: in una matrice in forma a scala per righe una colonna si dice *dominante* se contiene il primo coefficiente non nullo di qualche riga. Le variabili corrispondenti alle colonne dominanti sono pure chiamate *variabili dominanti.*

Guardando agli esempi precedenti, nell'Esempio 3.1 sono colonne dominanti di $[\mathbf{U}\ \mathbf{c}]$ le prime tre, mentre l'ultima colonna non è dominante. Nell'Esempio 3.2 sono colonne dominanti di $[\mathbf{V}\ \mathbf{c}]$ le prime tre, mentre le ultime due colonne non lo sono. Nell'Esempio 3.3 sono colonne dominanti di $[\mathbf{U}'\ \mathbf{c}]$ le prime due e l'ultima, mentre la terza colonna non lo è. Nell'Esempio 3.4 sono colonne dominanti di $[\mathbf{W}\ \mathbf{w}]$ la prima, la seconda e la quarta, mentre la terza e la quinta colonna non lo sono.

Possiamo allora dire che per la forma ridotta $[\mathbf{U}\ \mathbf{c}]$ della matrice aumentata $[\mathbf{A}\ \mathbf{b}]$ di un generico sistema $\mathbf{Ax} = \mathbf{b}$ accade uno e uno solo dei tre casi seguenti.

- [1] Tutte le colonne sono dominanti tranne l'ultima (corrispondente alla colonna dei termini noti); in tal caso il sistema ammette una e una sola soluzione.
- [∞] L'ultima colonna non è dominante ed esiste almeno una colonna non dominante; in tal caso il sistema ammette infinite soluzioni, che dipendono da tanti parametri quante sono le colonne non dominanti.

  Le variabili del sistema corrispondenti alle colonne non dominanti della forma ridotta si chiamano *variabili libere.*
- [0] L'ultima colonna è dominante; in tal caso il sistema non ammette soluzioni.

Se $\mathbf{A}$ è una matrice non nulla $m \times n$, il caso [1] può accadere solo se $m \geq n$, il caso [∞] solo se $n > 1$, mentre il caso [0] può accadere qualunque siano m ed n. Osserviamo che la EG, così come la abbiamo descritta, è un algoritmo che non lascia discrezionalità se, quando necessitano scambi di righe, si fissa *a priori* il modo in cui lo scambio va eseguito, per esempio utilizzando la prima riga utile successiva a quella da scambiare; per altro, ogni altro scambio con righe utili è legittimo.

Esplicitiamo meglio cosa si intende con "riga utile". Uno scambio di righe si rende necessario quando si vuole eliminare una certa variabile x_j nelle equazioni a partire da una di esse, diciamo dalla i-esima; se il coefficiente di x_j nella i-esima equazione è 0, si cerca nella matrice aumentata una riga al di sotto della i-esima in cui il coefficiente di x_j è diverso da 0. Una qualunque di queste è una "riga utile" al fine di eseguire lo scambio con la riga i-esima. Può però accadere che in tutte le righe al di sotto della i-esima il coefficiente di x_j sia uguale a 0. In tal caso la colonna j-esima della matrice aumentata non ha bisogno di nessuna modifica e si passa quindi alla colonna successiva, cioè alla variabile x_{j+1}.

Vogliamo schematizzare la forma che una generica matrice in forma ridotta $[\mathbf{U}\ \mathbf{c}]$ assume quando si è giunti al termine della EG, a partire dalla matrice aumentata di un sistema lineare.

Nel caso [1], in cui il sistema $m \times n$ ammette una e una sola soluzione, la forma ridotta è del tipo

$$[\mathbf{U}\ \mathbf{c}] = \left[\begin{array}{cccccc|c} 1 & + & + & \dots & + & + & c_1 \\ 0 & 1 & + & \dots & + & + & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & + & c_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 1 & c_n \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{array}\right]$$

dove i simboli "+" stanno a indicare coefficienti arbitrari.

Nel caso $[\infty]$, in cui il sistema ammette infinite soluzioni, la forma ridotta è del tipo

$$[\mathbf{U}\ \mathbf{c}] = \left[\begin{array}{ccccccc|c} \mathbf{0}^T & 1 & \dots & + & \dots & + & \dots & c_1 \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & + & \dots & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & c_k \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \end{array}\right]$$

dove compare effettivamente qualcuno dei vettori riga $\mathbf{0}^T$ in qualcuna delle prime k righe (ovvero, ci sono variabili libere) e ci sono $k \leq n$ righe non nulle.

Infine, nel caso [0], in cui il sistema non ammette soluzioni, la forma a scala per righe è del tipo

$$[\mathbf{U}\ \mathbf{c}] = \left[\begin{array}{ccccccc|c} \mathbf{0}^T & 1 & \dots & + & \dots & + & \dots & c_1 \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & + & \dots & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & c_k \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & c_{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \end{array}\right]$$

dove c_{k+1} è diverso da 0.

Nella EG si possono anche non eseguire le moltiplicazioni che rendono uguali a 1 i primi coefficienti non nulli di ogni riga; in tal caso i numeri $d_i \neq 0$ che compaiono come primi coefficienti non nulli di ogni riga si chiamano *pivot*, per il fatto che vengono utilizzati per annullare gli elementi della loro colonna che si trovano sotto di loro.

Il numero k di righe non nulle nella forma ridotta $\mathbf{U}$ (che nel caso [1] coincide con n) si chiama *rango* di $\mathbf{U}$; si noti che il rango di $\mathbf{U}$ coincide col numero delle sue colonne dominanti.

La definizione di rango di una matrice generica che daremo più avanti è del tutto diversa da quella data ora per una matrice in forma ridotta, ma risulterà a essa equivalente. Seguirà da quanto vedremo sugli spazi vettoriali che ogni forma ridotta cui si perviene al termine della EG a partire da una data matrice ha sempre lo stesso rango. Si può pertanto definire il rango di una generica matrice $\mathbf{A} \neq \mathbf{O}$ come il numero di righe non nulle comune a tutte le forme ridotte di $\mathbf{A}$.

Esempio 3.5. Consideriamo il sistema $\mathbf{Ax} = \mathbf{b}$, dove

$$[\mathbf{A}\ \mathbf{b}] = \left[\begin{array}{cccc|c} 0 & 0 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 2 & 0 & \alpha \end{array}\right]$$

e dove α è un parametro, cioè un simbolo che può assumere un arbitrario valore numerico. La prima riga non si può utilizzare per porre 0 nella prima colonna al di sotto del coefficiente di posto $(1, 1)$; si può scambiare la prima riga con la seconda oppure con la terza. Nel primo caso la EG porge successivamente le tre matrici

$$\left[\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 & \alpha \end{array}\right]; \quad \left[\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & \alpha - 1 \end{array}\right];$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & \alpha - 2 \end{array}\right].$$

Nel secondo caso invece la EG porge successivamente le tre matrici

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 0 & \alpha \\ 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]; \quad \left[\begin{array}{cccc|c} 1 & -1 & 2 & 0 & \alpha \\ 0 & 0 & -1 & -1 & 1 - \alpha \\ 0 & 0 & 1 & 1 & 1 \end{array}\right];$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 0 & \alpha \\ 0 & 0 & 1 & 1 & \alpha - 1 \\ 0 & 0 & 0 & 0 & 2 - \alpha \end{array}\right].$$

Si perviene quindi a due forme ridotte differenti di $[\mathbf{A}\ \mathbf{b}]$. Però in entrambi le forme ridotte il sistema ha soluzione se e solo se $\alpha = 2$, e in tal caso la prima forma ridotta porge con la sostituzione all'indietro le soluzioni

$$x_4 = k, \quad x_3 = 1 - k, \quad x_2 = h, \quad x_1 = 1 + k - (1 - k) + h$$

mentre la seconda forma ridotta porge le soluzioni

$$x_4 = k, \quad x_3 = 1 - k, \quad x_2 = h, \quad x_1 = 2 - 2(1 - k) + h;$$

i due insiemi di soluzioni evidentemente coincidono, giacché $1 + k - (1 - k) + h = 2k + h = 2 - 2(1 - k) + h$. □

4. Matrici inverse e matrice pseudo-inversa

Dato un numero complesso z, esiste un test semplicissimo per vedere se z ammette un inverso, cioè un numero z' tale che $zz' = 1$: basta e occorre che z sia diverso da 0. Tale nulmero z' è inoltre univocamente individuato da z, per il fatto che vale la "legge dell'annullamento del prodotto" (che dice che il prodotto di numeri non nulli è non nullo): $zz' = 1 = zz''$ implica $z(z' - z'') = 0$, implica $z' = z''$.

Il problema della invertibilità per le matrici è più complesso. Anzitutto, essendo la moltiplicazione non commutativa, occorre distinguere tra inversa destra e inversa sinistra: data la matrice $m \times n$ $\mathbf{A}$, si chiama *inversa destra* di $\mathbf{A}$ una matrice $\mathbf{R}$ tale che $\mathbf{AR} = \mathbf{I}_m$; si chiama *inversa sinistra* di $\mathbf{A}$ una matrice $\mathbf{L}$ tale che $\mathbf{LA} = \mathbf{I}_n$ (i simboli $\mathbf{R}$ ed $\mathbf{L}$ stanno a indicare le iniziali dei termini inglesi "right" e "left"). Notiamo subito che sia $\mathbf{R}$ che $\mathbf{L}$ sono matrici $n \times m$.

Una matrice che sia contemporaneamente inversa destra e inversa sinistra della matrice $\mathbf{A}$ si chiama *inversa bilatera*, o più semplicemente *la inversa* di $\mathbf{A}$, e ciò per il buon motivo che essa, se esiste, è unica, come prova la seguente proposizione.

PROPOSIZIONE 4.1. *Sia* $\mathbf{A}$ *una matrice* $m \times n$. *Se* $\mathbf{A}$ *ha sia inversa destra* $\mathbf{R}$ *che inversa sinistra* $\mathbf{L}$, *allora* $\mathbf{R} = \mathbf{L}$. *Ne consegue che* $\mathbf{R} = \mathbf{L}$ *è l'unica inversa (destra, sinistra e bilatera) di* $\mathbf{A}$.

DIMOSTRAZIONE. $\mathbf{L} = \mathbf{L}\mathbf{I}_m = \mathbf{L}(\mathbf{A}\mathbf{R}) = (\mathbf{L}\mathbf{A})\mathbf{R} = \mathbf{I}_n\mathbf{R} = \mathbf{R}$. □

La matrice inversa della matrice $\mathbf{A}$, qualora esista, viene denotata con $\mathbf{A}^{-1}$. Vedremo tra poco che $\mathbf{A}^{-1}$ può esistere solo nel caso in cui $\mathbf{A}$ è quadrata; in tal caso la matrice $\mathbf{A}$ si dice *invertibile* (o anche *non-singolare*). È importante rilevare che il termine *invertibile* non si può riferire a matrici dotate solo di inversa destra o sinistra.

Oltre al problema di distinguere tra inverse destre e sinistre, esiste anche il problema che non basta più che una matrice sia diversa dalla matrice nulla perché sia dotata di inversa. Inoltre, come terzo problema, si ha anche la possibilità che una matrice ammetta più di una inversa destra o sinistra, cosa prevedibile per il fatto che la "legge dell'annullamento del prodotto" non vale per il prodotto tra matrici, come si è visto nell'Esempio 1.13.

I seguenti semplici esempi convinceranno subito dell'esistenza di queste possibilità.

ESEMPIO 4.2. Il vettore riga $\mathbf{v}^T = [a\ b]$, con a e b non entrambi nulli, è una matrice con infinite inverse destre e nessuna inversa sinistra. Infatti, se $\mathbf{u} = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$, risulta

$$\mathbf{v}^T\mathbf{u} = 1 \iff ax + by = 1$$

e tale equazione nelle incognite x e y ha infinite soluzioni. Inoltre, $\mathbf{u}\mathbf{v}^T = \mathbf{I}_2$ comporta che

$$xa = 1, \quad xb = 0, \quad ya = 0, \quad yb = 1,$$

il che è evidentemente assurdo. □

ESEMPIO 4.3. La matrice 2×2

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

non ha nè inversa destra nè inversa sinistra. Se infatti esistesse una matrice

$$\mathbf{X} = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

per cui $\mathbf{X}\mathbf{A} = \mathbf{I}_2$, risulterebbe $x + y = 1$ e $x + y = 0$, che è manifestamente assurdo; se invece risultasse $\mathbf{A}\mathbf{X} = \mathbf{I}_2$, si avrebbe $x + z = 1$ e $x + z = 0$, pure assurdo. □

Stabiliamo preliminarmente un collegamento tra l'esistenza di soluzioni di un sistema lineare $\mathbf{A}\mathbf{x} = \mathbf{b}$ e l'esistenza di inverse della matrice dei coefficienti $\mathbf{A}$.

Se la matrice $\mathbf{A}$ ammette una inversa destra $\mathbf{R}$, il vettore $\mathbf{R}\mathbf{b}$ risulta essere una soluzione, perché $\mathbf{A}(\mathbf{R}\mathbf{b}) = (\mathbf{A}\mathbf{R})\mathbf{b} = \mathbf{I}\mathbf{b} = \mathbf{b}$.

Se invece la matrice $\mathbf{A}$ ammette una inversa sinistra $\mathbf{L}$, e se il sistema ammette soluzioni, allora la soluzione è unica: infatti, da $\mathbf{A}\mathbf{u} = \mathbf{b} = \mathbf{A}\mathbf{v}$, moltiplicando a sinistra ambo i membri per $\mathbf{L}$ si ricava che $\mathbf{u} = \mathbf{L}\mathbf{A}\mathbf{u} = \mathbf{L}v = \mathbf{L}\mathbf{A}\mathbf{v} = \mathbf{v}$.

Possiamo pertanto enunciare il seguente risultato, in cui il punto (c) segue ovviamente dai punti (a) e (b).

PROPOSIZIONE 4.4. *Il sistema di equazioni lineari* $\mathbf{A}\mathbf{x} = \mathbf{b}$

(a) *ammette almeno una soluzione se la matrice* $\mathbf{A}$ *ammette inversa destra;*
(b) *ammette al più una soluzione se la matrice* $\mathbf{A}$ *ammette inversa sinistra;*
(c) *ammette una e una sola soluzione se la matrice* $\mathbf{A}$ *ammette inversa bilatera.* □

Osserviamo che nel punto (c) la soluzione è $\mathbf{A}^{-1}\mathbf{b}$. Osserviamo inoltre che i tre punti considerati nella precedente proposizione forniscono condizioni che sono solo sufficienti per l'esistenza e l'unicità di soluzioni, e questo per il fatto che il vettore $\mathbf{b}$ è fissato. Proveremo invece che, se $\mathbf{Ax} = \mathbf{b}$ ha soluzione per ogni vettore $\mathbf{b}$, allora $\mathbf{A}$ ammette inversa destra (Teorema 4.10), mentre, se il particolare sistema $\mathbf{Ax} = \mathbf{0}$ ha al più una soluzione, allora $\mathbf{A}$ ammette inversa sinistra (Teorema 4.12).

Inverse destre e sinistre. Affrontiamo per primo il problema dell'esistenza di inverse destre e di inverse sinistre. Per le matrici quadrate si ha una situazione del tutto particolare. Infatti, la seguente Proposizione 4.6 mostra che, non appena la matrice quadrata $\mathbf{A}$ ha una inversa destra (risp., sinistra), questa è anche inversa sinistra (risp., destra). La sua prova inaugura una tecnica dimostrativa per induzione che fa uso della forma bordata delle matrici; questa tecnica sarà ampiamente utilizzata in seguito. Faremo inoltre uso del punto (i) del lemma seguente, che si riferisce al primo passo dell'eliminazione di Gauss sulla matrice $\mathbf{A}$; il lemma sarà dimostrato nel Paragrafo 6.

LEMMA 4.5. *Data una qualunque matrice* $\mathbf{A}$

(i) *esiste una matrice invertibile* $\mathbf{E}$ *tale che* $\mathbf{EA}$ *ha la seguente forma bordata*

$$\mathbf{EA} = \begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix};$$

(ii) *esiste una matrice invertibile* $\mathbf{F}$ *tale che* $\mathbf{FA} = \mathbf{U}$, *dove* $\mathbf{U}$ *è una forma ridotta di* $\mathbf{A}$. □

PROPOSIZIONE 4.6. *Sia* $\mathbf{A}$ *una matrice quadrata di ordine* n. *Allora una inversa destra di* $\mathbf{A}$ *è anche inversa sinistra, e viceversa.*

DIMOSTRAZIONE. Sia $\mathbf{AR} = \mathbf{I}_n$ e ragioniamo per induzione su n. Se $n = 1$ l'asserto è ovvio. Sia allora $n > 1$ e l'asserto vero per $n - 1$. Utilizziamo la forma bordata di $\mathbf{EA}$ data dal Lemma 4.5, che vale per una opportuna matrice invertibile $\mathbf{E}$. Osserviamo che

$$\mathbf{AR} = \mathbf{I}_n \implies \mathbf{EAR} = \mathbf{E} \implies \mathbf{EARE}^{-1} = \mathbf{EE}^{-1} = \mathbf{I}_n.$$

Decomponiamo anche $\mathbf{RE}^{-1}$ in forma bordata

$$\mathbf{RE}^{-1} = \begin{bmatrix} b & \mathbf{y}^T \\ \mathbf{z} & \mathbf{Y} \end{bmatrix}.$$

Dall'uguaglianza

$$\begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix} \begin{bmatrix} b & \mathbf{y}^T \\ \mathbf{z} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}_{n-1} \end{bmatrix}$$

si ricavano le uguaglianze

$$ab + \mathbf{x}^T\mathbf{z} = 1, \quad a\mathbf{y}^T + \mathbf{x}^T\mathbf{Y} = \mathbf{0}^T, \quad \mathbf{Xz} = \mathbf{0}, \quad \mathbf{XY} = \mathbf{I}_{n-1}.$$

Poiché le matrici $\mathbf{X}$ e $\mathbf{Y}$ hanno ordine $n - 1$, per l'ipotesi induttiva risulta $\mathbf{YX} = \mathbf{I}_{n-1}$, quindi $\mathbf{Y} = \mathbf{X}^{-1}$ e $\mathbf{z} = \mathbf{I}_{n-1}\mathbf{z} = \mathbf{YXz} = \mathbf{Y0} = \mathbf{0}$. Ne consegue che $ab = 1$ e $\mathbf{y}^T = -a^{-1}\mathbf{x}^T\mathbf{X}^{-1}$, quindi

$$\mathbf{RE}^{-1} = \begin{bmatrix} a^{-1} & -a^{-1}\mathbf{x}^T\mathbf{X}^{-1} \\ \mathbf{0} & \mathbf{X}^{-1} \end{bmatrix}.$$

Basta allora a questo punto eseguire la moltiplicazione $\mathbf{RA} = \mathbf{RE}^{-1}\mathbf{EA}$ a blocchi:

$$\begin{aligned} \begin{bmatrix} a^{-1} & -a^{-1}\mathbf{x}^T\mathbf{X}^{-1} \\ \mathbf{0} & \mathbf{X}^{-1} \end{bmatrix} \begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix} &= \begin{bmatrix} a^{-1}a & a^{-1}\mathbf{x}^T - a^{-1}\mathbf{x}^T\mathbf{X}^{-1}\mathbf{X} \\ \mathbf{0} & \mathbf{X}^{-1}\mathbf{X} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}_{n-1} \end{bmatrix} \end{aligned}$$

per ottenere che $\mathbf{RA} = \mathbf{I}_n$.

Per quanto riguarda il viceversa, se $\mathbf{LA} = \mathbf{I}_n$, utilizzando quanto appena visto si ricava:

$$\mathbf{LA} = \mathbf{I}_n \implies (\mathbf{LA})^T = \mathbf{I}_n^T \implies \mathbf{A}^T\mathbf{L}^T = \mathbf{I}_n \implies$$
$$\implies \mathbf{L}^T\mathbf{A}^T = \mathbf{I}_n \implies (\mathbf{AL})^T = \mathbf{I}_n \implies \mathbf{AL} = \mathbf{I}_n. \quad \square$$

Il risultato che segue mostra come una matrice può avere inversa destra solo se è "orizzontale", cioè con un numero di colonne non inferiore a quello delle righe.

PROPOSIZIONE 4.7. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *dotata di inversa destra. Allora* $m \leq n$.

DIMOSTRAZIONE. Sia $\mathbf{AR} = \mathbf{I}_m$. Se fosse $m > n$ si potrebbero decomporre a blocchi $\mathbf{A}$ ed $\mathbf{R}$ nel modo seguente

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}; \qquad \mathbf{R} = [\mathbf{R}_1 \ \mathbf{R}_2]$$

con $\mathbf{A}_1$ ed $\mathbf{R}_1$ blocchi quadrati di ordine n. Moltiplicando $\mathbf{A}$ ed $\mathbf{R}$ a blocchi e uguagliando a $\mathbf{I}_m$, pure decomposta a blocchi in modo conforme, si ottiene:

$$\mathbf{A}_1\mathbf{R}_1 = \mathbf{I}_n, \quad \mathbf{A}_1\mathbf{R}_2 = \mathbf{O}, \quad \mathbf{A}_2\mathbf{R}_1 = \mathbf{O}, \quad \mathbf{A}_2\mathbf{R}_2 = \mathbf{I}_{m-n}.$$

Dalla Proposizione 4.6 sappiamo che $\mathbf{A}_1\mathbf{R}_1 = \mathbf{R}_1\mathbf{A}_1$, quindi

$$\mathbf{A}_2\mathbf{R}_1 = \mathbf{O} \implies \mathbf{A}_2\mathbf{R}_1\mathbf{A}_1 = \mathbf{O} \implies \mathbf{A}_2 = \mathbf{O},$$

che contraddice l'uguaglianza $\mathbf{A}_2\mathbf{R}_2 = \mathbf{I}_{m-n}$. $\square$

Come immediata conseguenza si ricava il risultato simmetrico per le inverse sinistre di matrici $m \times n$, che mostra come esse possono esistere solo per matrici "verticali", cioè con un numero di righe non inferiore a quello delle colonne.

COROLLARIO 4.8. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *dotata di inversa sinistra. Allora* $m \geq n$.

DIMOSTRAZIONE. $\mathbf{A}$ ammette come inversa sinistra la matrice $\mathbf{L}$ se e solo se $\mathbf{A}^T$ ammette come inversa destra la matrice $\mathbf{L}^T$. Si applichi allora ad $\mathbf{A}^T$ la Proposizione 4.7. $\square$

Diamo un'altra diretta conseguenza dei due ultimi risultati e della Proposizione 4.1.

COROLLARIO 4.9. *Se una matrice* $\mathbf{A}$ *ha inversa destra e inversa sinistra, allora è una matrice quadrata invertibile.*

DIMOSTRAZIONE. Sia $\mathbf{A}$ $m \times n$. Per la Proposizione 4.7 risulta $m \leq n$ e per il Corollario 4.8 risulta $m \geq n$, quindi $m = n$. La Proposizione 4.1 assicura poi che $\mathbf{A}$ è invertibile. $\square$

Diamo ora il risultato che contiene le principali caratterizzazioni delle matrici che ammettono inversa destra.

TEOREMA 4.10. *Per una matrice* $m \times n$ $\mathbf{A}$, *le seguenti proprietà sono equivalenti:*

(a) $\mathbf{A}$ *ammette inversa destra;*
(b) *il sistema* $\mathbf{Ax} = \mathbf{b}$ *ammette almeno una soluzione per ogni scelta del vettore* $\mathbf{b}$*;*
(c) $\mathbf{A}$ *ha rango* m.

DIMOSTRAZIONE. (a) $\Rightarrow$ (b) Segue dalla Proposizione 4.4 (a).

(b) $\Rightarrow$ (c) Se per assurdo $\mathbf{A}$ avesse rango $< m$, una sua forma ridotta $\mathbf{U}$ avrebbe l'ultima riga nulla. Indicato al solito con $\mathbf{e}_m$ l'm-esimo vettore coordinato con m coordinate, nella matrice $[\mathbf{U}\ \mathbf{e}_m]$ l'ultima colonna sarebbe dominante quindi si sarebbe nel caso [0] del Paragrafo precedente e il sistema con matrice aumentata $[\mathbf{U}\ \mathbf{e}_m]$ non avrebbe soluzioni. Procedendo a ritroso, a partire dlla matrice $[\mathbf{U}\ \mathbf{e}_m]$, con le operazioni elementari inverse di quelle eseguite sulla matrice $\mathbf{A}$ per ottenere la forma ridotta $\mathbf{U}$, si perviene alla matrice aumentata $[\mathbf{A}\ \mathbf{b}]$ di un sistema $\mathbf{Ax} = \mathbf{b}$, equivalente al sistema $\mathbf{Ux} = \mathbf{e}_m$, che non ha soluzione, il che è assurdo. Quindi $\mathbf{A}$ ha rango m.

(c) $\Rightarrow$ (a) Si tratta di provare che esiste una matrice $n \times m$ $\mathbf{R} = [\mathbf{r}_1\ \mathbf{r}_2\ \dots\ \mathbf{r}_n]$, di cui sono state messe in evidenza le colonne, tale che $\mathbf{AR} = \mathbf{I}_m$, cioè $\mathbf{Ar}_j = \mathbf{e}_j$ per ogni $j \leq m$. Bisogna trovare quindi soluzioni per ciascuno dei sistemi $\mathbf{Ax}_j = \mathbf{e}_j$. La EG sulle matrici aumentate $[\mathbf{A}\ \mathbf{e}_j]$ produce le forme ridotte $[\mathbf{U}\ \mathbf{c}_j]$ in cui l'ultima colonna non è mai dominante, giacché l'ultima riga di $\mathbf{U}$ non è nulla. Pertanto i sistemi considerati hanno soluzione e la matrice $\mathbf{R}$ esiste. $\square$

L'equivalenza di (a) e (c) nel Teorema 4.10 e dalla Proposizione 4.6 ha una conseguenza immediata.

COROLLARIO 4.11. *Una matrice quadrata di ordine m è invertibile se e solo se ha rango m.* $\square$

Il risultato simmetrico del precedente per le matrici che ammettono inversa sinistra è fornito dal teorema seguente.

TEOREMA 4.12. *Per una matrice $m \times n$ $\mathbf{A}$, le seguenti proprietà sono equivalenti:*

(a) $\mathbf{A}$ *ammette inversa sinistra;*
(b) *il sistema* $\mathbf{Ax} = \mathbf{0}$ *ammette solo la soluzione nulla;*
(c) $\mathbf{A}$ *ha rango n;*
(d) *la matrice* $\mathbf{A}^H\mathbf{A}$ *è invertibile.*

DIMOSTRAZIONE. (a) $\Rightarrow$ (b) Segue dalla Proposizione 4.4 (b).

(b) $\Leftrightarrow$ (c) $\mathbf{A}$ ha rango $< n$ se e solo se la matrice aumentata $[\mathbf{A}\ \mathbf{0}]$ ha forma ridotta $[\mathbf{U}\ \mathbf{0}]$ con meno di n righe non nulle. Ciò equivale al fatto che almeno una colonna di $\mathbf{U}$ non è dominante, cioè il sistema $\mathbf{Ax} = \mathbf{0}$ ricade nel caso $[\infty]$ del Paragrafo 3 e ha infinite soluzioni (oltre a quella nulla).

(c) $\Rightarrow$ (d) Essendo $\mathbf{A}^H\mathbf{A}$ matrice quadrata di ordine n, per il Corollario 4.11 è sufficiente provare che essa ha rango n. Per l'equivalenza di (b) con (c) applicata alla matrice $\mathbf{A}^H\mathbf{A}$, basta provare che il sistema $\mathbf{A}^H\mathbf{Ax} = \mathbf{0}$ ammette solo la soluzione nulla. Sia $\mathbf{v}$ una soluzione; risulta allora $\mathbf{v}^H\mathbf{A}^H\mathbf{Av} = 0$. Posto $\mathbf{Av} = \mathbf{y} = (y_1\ \dots\ y_m)^T$, si ha $\mathbf{y}^H\mathbf{y} = 0$; ma $\mathbf{y}^H\mathbf{y} = \overline{y}_1 y_1 + \dots + \overline{y}_m y_m = 0$ implica che $y_i = 0$ per ogni i e quindi $\mathbf{y} = \mathbf{0}$. Ma allora $\mathbf{Av} = \mathbf{0}$ e da (b) segue che $\mathbf{v} = \mathbf{0}$, come desiderato.

(d) $\Rightarrow$ (a) Una inversa sinistra di $\mathbf{A}$ è $(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. $\square$

Per completare la caratterizzazione delle matrici che ammettono inversa destra, manca l'analogo del punto (d) del Teorema 4.12, che però siamo ora in grado di dimostrare.

COROLLARIO 4.13. *Una matrice $m \times n$ $\mathbf{A}$ ha inversa destra se e solo se la matrice $\mathbf{AA}^H$ è invertibile.*

DIMOSTRAZIONE. Se $\mathbf{A}$ ha inversa destra, allora $\mathbf{A}^H$ ha inversa sinistra. Per il Teorema 4.12, $(\mathbf{A}^H)^H\mathbf{A}^H$ è invertibile. Ma $(\mathbf{A}^H)^H\mathbf{A}^H = \mathbf{AA}^H$, da cui l'asserto.

Viceversa, se $\mathbf{AA}^H$ è invertibile, una inversa destra di $\mathbf{A}$ è $\mathbf{A}^H(\mathbf{AA}^H)^{-1}$. $\square$

È utile osservare che nelle condizioni sulle matrici $\mathbf{A}\mathbf{A}^H$ e $\mathbf{A}^H\mathbf{A}$ nel Teorema 4.10 e nel Corollario 4.13 non si può sostituire la matrice H-trasposta $\mathbf{A}^H$ con la matrice trasposta $\mathbf{A}^T$, naturalmente a meno che $\mathbf{A}$ non sia una matrice reale. Infatti, pure essendo $\mathbf{A}$ non nulla, può risultare $\mathbf{A}\mathbf{A}^T = \mathbf{O}$ (e analogamente $\mathbf{A}^T\mathbf{A} = \mathbf{O}$), come mostra il seguente esempio.

Esempio 4.14. Si consideri il vettore riga

$$\mathbf{A} = [1 + i \ 1 - i]$$

per il quale risulta

$$\mathbf{A}^T = \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}, \qquad \mathbf{A}^H = \begin{bmatrix} 1-i \\ 1+i \end{bmatrix}.$$

Un facile calcolo mostra che

$$\begin{aligned} \mathbf{A}\mathbf{A}^T &= 1 - 1 + 2i + 1 - 1 - 2i = 0, \\ \mathbf{A}\mathbf{A}^H &= (1+i)(1-i) + (1-i)(1+i) = 4. \end{aligned}$$ □

I due Teoremi 4.10 e 4.12 congiuntamente con la Proposizione 4.6 e il Corollario 4.13 porgono come immediata conseguenza il seguente teorema, in cui il punto (e) impiega la nozione di determinante, che sarà introdotta nel Capitolo IV, ed è pertanto inserito qui a puro titolo informativo.

Teorema 4.15. *Sia* $\mathbf{A}$ *una matrice quadrata di ordine* n. *Le seguenti proprietà sono equivalenti:*

(a) $\mathbf{A}$ *è invertibile;*
(b) *il sistema* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *ammette almeno una soluzione per ogni scelta del vettore* $\mathbf{b}$*;*
(c) *il sistema* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *ha come unica soluzione* $\mathbf{x} = \mathbf{0}$*;*
(d) $\mathbf{A}$ *ha rango* n*;*
(e) *il determinante di* $\mathbf{A}$ *è diverso da* 0. □

Il modo più semplice per vedere se una matrice quadrata è invertibile risulta quello fornito nel punto (d) del Teorema 4.15: si controlla che una forma ridotta della matrice non abbia righe nulle. L'algoritmo di inversione presentato nel prossimo Paragrafo permetterà poi di trovare esplicitamente l'inversa.

Le proprietà delle matrici inverse di matrici quadrate sono raccolte nel seguente elenco; le facili dimostrazioni sono lasciate come esercizio.

Siano $\mathbf{A}$ e $\mathbf{B}$ matrici quadrate invertibili di ordine n. Allora:

(a) $\mathbf{AB}$ è invertibile e $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
(b) $\mathbf{A}^{-1}$ è invertibile e $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
(c) $\mathbf{A}^T$ e $\mathbf{A}^H$ sono invertibili e $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$, $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$.

È utile infine sapere cosa accade per le inverse di matrici triangolari e per matrici decomposte a blocchi. Nei due esempi seguenti vengono date informazioni su tali inverse.

Esempio 4.16. Una matrice triangolare superiore (risp., inferiore) $\mathbf{T} = (t_{ij})$ è invertibile se e solo se tutti i coefficienti diagonali t_{ii} sono diversi da zero. Se ciò accade, la matrice inversa $\mathbf{T}^{-1}$ è ancora triangolare superiore (risp., inferiore) e ha come coefficienti diagonali gli inversi dei corrispondenti coefficienti diagonali di $\mathbf{T}$. La dimostrazione si esegue facilmente per induzione sull'ordine della matrice.

Lo stesso vale per matrici triangolari a blocchi, con blocchi diagonali quadrati, sostituendo le condizioni sui coefficienti diagonali con analoghe condizioni sui blocchi diagonali. In particolare, data la matrice triangolare a blocchi

$$\mathbf{T} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{O} & \mathbf{V} \end{bmatrix}$$

con $\mathbf{X}$ e $\mathbf{V}$ blocchi quadrati invertibili, risulta

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{X}^{-1} & -\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} \\ \mathbf{O} & \mathbf{V}^{-1} \end{bmatrix}.$$

Infatti

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{O} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{X}^{-1} & -\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} \\ \mathbf{O} & \mathbf{V}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\mathbf{X}^{-1} & -\mathbf{X}\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} + \mathbf{Y}\mathbf{V}^{-1} \\ \mathbf{O} & \mathbf{V}\mathbf{V}^{-1} \end{bmatrix}$$

e l'ultima matrice a blocchi è evidentemente la matrice identità. □

L'esempio seguente tratta il caso non banale più semplice di matrici inverse, quello 2×2. Si confronti con l'Esempio 4.3.

Esempio 4.17. La matrice 2×2

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

ha inversa se e solo se $\Delta = ad - bc \neq 0$, e in tal caso l'inversa è la matrice

$$\mathbf{A}^{-1} = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

La verifica che, nel caso in cui $\Delta \neq 0$, la suddetta matrice è l'inversa di $\mathbf{A}$ è immediata. Il fatto che l'invertibilità di $\mathbf{A}$ comporta che $\Delta \neq 0$ si vede nel modo seguente. Se

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = I_2$$

allora $ax + bz = 1$, $cx + dz = 0$ e $cy + dw = 1$. Nel caso in cui $c \neq 0$, sottraendo dalla prima uguaglianza moltiplicata per c la seconda moltiplicata per a si ricava $(bc-ad)z = c$; nel caso in cui $d \neq 0$, sottraendo dalla prima uguaglianza moltiplicata per d la seconda moltiplicata per b si ricava $(ad - bc)x = d$. Poiché non può essere $c = 0 = d$, perché $cy + dw = 1$, in ogni caso risulta $\Delta \neq 0$.

Si vedrà al Capitolo IV che il numero Δ non è altro che il determinante della matrice $\mathbf{A}$. □

Pseudo-inversa. Si è visto che una matrice ha inversa—destra, sinistra o bilatera—solo in casi particolari. Esiste una nozione che generalizza quella di matrice inversa che è applicabile a una qualunque matrice, ed è quella di matrice pseudo-inversa. Il lettore è avvertito del fatto che sono state date molteplici nozioni che generalizzano quella di matrice inversa; quella che presentiamo qui prende anche il nome di "pseudo-inversa di Moore-Penrose" (dall'americano E. H. Moore, che la introdusse negli anni '30, e dall'inglese R. Penrose, che la divulgò negli anni '50).

Data una qualunque matrice $m \times n$ $\mathbf{A}$, si chiama *pseudo-inversa* di $\mathbf{A}$ una matrice $\mathbf{A}^+$ che soddisfa alle quattro condizioni:

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}, \qquad \mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+, \qquad \mathbf{A}\mathbf{A}^+ = (\mathbf{A}\mathbf{A}^+)^H, \qquad \mathbf{A}^+\mathbf{A} = (\mathbf{A}^+\mathbf{A})^H.$$

Si osservi che la martrice $\mathbf{A}^+$ ha necessariamente dimensioni $n \times m$. Naturalmente si pone il problema dell'esistenza di una tale matrice $\mathbf{A}^+$, e, qualora essa esista, si pone il problema della sua unicità. Cominciamo col problema dell'unicità.

Proposizione 4.18. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *e siano* $\mathbf{B}$ *e* $\mathbf{C}$ *due matrici che soddisfano alle quattro condizioni cui deve soddisfare una matrice pseudo-inversa. Allora* $\mathbf{B} = \mathbf{C}$.

DIMOSTRAZIONE. Utilizzeremo nel seguito, oltre alle proprietà delle matrici H-trasposte, tutte le quattro condizioni cui soddisfano $\mathbf{B}$ e $\mathbf{C}$. Il lettore controlli a ogni passaggio quale delle condizioni viene utilizzata.

$$\begin{aligned}\mathbf{B} &= \mathbf{BAB} = (\mathbf{BA})^H\mathbf{B}\\ &= \mathbf{A}^H\mathbf{B}^H\mathbf{B} = (\mathbf{ACA})^H\mathbf{B}^H\mathbf{B}\\ &= \mathbf{A}^H\mathbf{C}^H\mathbf{A}^H\mathbf{B}^H\mathbf{B} = (\mathbf{CA})^H(\mathbf{BA})^H\mathbf{B}\\ &= \mathbf{CABAB} = \mathbf{CAB}\\ &= \mathbf{CACAB} = \mathbf{C}(\mathbf{AC})^H(\mathbf{AB})^H\\ &= \mathbf{CC}^H\mathbf{A}^H\mathbf{B}^H\mathbf{A}^H = \mathbf{CC}^H(\mathbf{ABA})^H\\ &= \mathbf{CC}^H\mathbf{A}^H = \mathbf{C}(\mathbf{AC})^H\\ &= \mathbf{CAC} = \mathbf{C}.\end{aligned}$$

□

Per quanto riguarda l'esistenza della pseudo-inversa, la prima cosa da farsi è quella di esaminare cosa accade nei tre casi, già studiati nella prima parte di questo Paragrafo, in cui esiste l'inversa bilatera, oppure esistono inverse destre o inverse sinistre.

(i) Suppponiamo che la matrice $\mathbf{A}$ sia invertibile, quindi quadrata. È immediato verificare che $\mathbf{A}^{-1}$ soddisfa alle quattro condizioni della pseudo-inversa, perciò $\mathbf{A}^+ = \mathbf{A}^{-1}$.

(ii) Suppponiamo che la matrice $\mathbf{A}$ non sia quadrata e abbia inversa destra; quindi risulta $m < n$. Il Teorema 4.10 (d) assicura che esiste la matrice $\mathbf{A}^H(\mathbf{AA}^H)^{-1}$, che è ovviamente inversa destra di $\mathbf{A}$. Vogliamo verificare che

$$\mathbf{A}^+ = \mathbf{A}^H(\mathbf{AA}^H)^{-1}.$$

Poiché ogni inversa destra soddisfa banalmente a tre delle condizioni richieste, basta controllare che $\mathbf{A}^H(\mathbf{AA}^H)^{-1}\mathbf{A}$ è hermitiana. Si ha infatti

$$(\mathbf{A}^H(\mathbf{AA}^H)^{-1}\mathbf{A})^H = \mathbf{A}^H((\mathbf{AA}^H)^{-1})^H\mathbf{A} = \mathbf{A}^H(\mathbf{AA}^H)^{-1}\mathbf{A}.$$

Si può verificare direttamente (si veda l'Esercizio **??**), o dedurre dalla Proposizione 4.18, che $\mathbf{A}^H(\mathbf{AA}^H)^{-1}$ è l'unica tra le infinite inverse destre $\mathbf{R}$ di $\mathbf{A}$ a soddisfare alla condizione che $\mathbf{RA}$ è hermitiana.

(iii) Supponiamo che la matrice $\mathbf{A}$ non sia quadrata e abbia inversa sinistra; quindi risulta $m > n$. Con ragionamento in tutto analogo al precedente, partendo dal Teorema 4.11 (d) si prova che

$$\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H.$$

L'esistenza della matrice pseudo-inversa nel caso di una matrice qualunque $\mathbf{A}$ si deduce agevolmente dai due precedenti casi (ii) e (iii), una volta che si sappia decomporre la matrice $\mathbf{A}$ in modo opportuno.

PROPOSIZIONE 4.19. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *e sia* $\mathbf{A} = \mathbf{BC}$ *una sua fattorizzazione tale che* $\mathbf{B}$ *ammette inversa sinistra e* $\mathbf{C}$ *ammette inversa destra. Allora* $\mathbf{A}^+$ *esiste e risulta*

$$\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^+ = (\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H\mathbf{B}^H(\mathbf{BB}^H)^{-1}.$$

DIMOSTRAZIONE. Dai casi (ii) e (iii) precedentemente esaminati sappiamo che

$$\mathbf{B}^+ = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H, \qquad \mathbf{C}^+ = \mathbf{C}^H(\mathbf{CC}^H)^{-1}.$$

Si tratta allora di verificare le uguaglianze

$$\mathbf{BC} = \mathbf{BCC}^+\mathbf{B}^+\mathbf{BC}, \qquad \mathbf{C}^+\mathbf{B}^+ = \mathbf{C}^+\mathbf{B}^+\mathbf{BCC}^+\mathbf{B}^+$$

che risultano ovvie, essendo $\mathbf{B}^+$ inversa sinistra di $\mathbf{B}$ e $\mathbf{C}^+$ inversa destra di $\mathbf{C}$; bisogna inoltre verificare che sono hermitiane le matrici $\mathbf{BCC}^+\mathbf{B}^+$ e $\mathbf{C}^+\mathbf{B}^+\mathbf{BC}$, il che è pure ovvio, perché

$$\mathbf{BCC}^+\mathbf{B}^+ = \mathbf{BB}^+, \qquad \mathbf{C}^+\mathbf{B}^+\mathbf{BC} = \mathbf{C}^+\mathbf{C}. \qquad \square$$

Va osservato che se si ha una arbitraria fattorizzazione $\mathbf{A} = \mathbf{BC}$ della matrice $\mathbf{A}$, non è vero che $\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^+$; si vedano però le proprietà (pi_6) e (pi_7) elencate di seguito.

Vedremo al termine del Paragrafo 6 che una generica matrice $\mathbf{A}$ ha sempre fattorizzazioni soddisfacenti all'ipotesi della Proposizione 4.19, che saranno chiamate "decomposizioni a rango pieno". Possiamo perciò concludere con il seguente risultato.

TEOREMA 4.20. *Data una qualunque matrice* $\mathbf{A}$*, la sua pseudo-inversa* $\mathbf{A}^+$ *esiste ed è unica.* $\square$

Diamo un semplice esempio di matrice pseudo-inversa.

ESEMPIO 4.21. Si consideri la matrice 2×3

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

È immediato verificare che vale la fattorizzazione

$$\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} [2\ 0\ 1\ 3]$$

dove la matrice $\mathbf{B} = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ ha inversa sinistra e la matrice $\mathbf{C} = [2\ 0\ 1\ 3]$ ha inversa destra. Un facile calcolo mostra che

$$\mathbf{B}^+ = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H = \left([2\ 1]\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)^{-1}[2\ 1] = [2/5\ 1/5]$$

$$\mathbf{C}^+ = \mathbf{C}^H(\mathbf{CC}^H)^{-1} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix}\left([2\ 0\ 1\ 3]\begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix}\right)^{-1} = \begin{bmatrix} 2/14 \\ 0 \\ 1/14 \\ 3/14 \end{bmatrix}.$$

Pertanto la pseudo-inversa di $\mathbf{A}$ risulta essere:

$$\mathbf{A}^+ = \begin{bmatrix} 2/14 \\ 0 \\ 1/14 \\ 3/14 \end{bmatrix}[2/5\ 1/5] = \begin{bmatrix} 4/70 & 2/70 \\ 0 & 0 \\ 2/70 & 1/70 \\ 6/70 & 3/70 \end{bmatrix} = 70^{-1}\mathbf{A}^T. \qquad \square$$

Forniamo di seguito alcune proprietà della matrice psudo-inversa, la cui verifica, sempre molto facile, è lasciata come esercizio.

(pi_1) $(\mathbf{A}^+)^+ = \mathbf{A}$;
(pi_2) $(\mathbf{A}^+)^H = (\mathbf{A}^H)^+$;
(pi_3) $\mathbf{O}_{mn}^+ = \mathbf{O}_{nm}$;
(pi_4) $(\alpha\mathbf{A})^+ = \alpha^{-1}\mathbf{A}^+$ per ogni scalare $\alpha \neq 0$;
(pi_5) $\mathbf{A}^H = \mathbf{A}^+\mathbf{AA}^H = \mathbf{A}^H\mathbf{AA}^+$;
(pi_6) se $\mathbf{A} = \mathbf{BC}$ con $\mathbf{B}$ soddisfacente all'uguaglianza $\mathbf{B}^H\mathbf{B} = \mathbf{I}$, allora $\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^H$;
(pi_7) se $\mathbf{A} = \mathbf{BC}$ con $\mathbf{C}$ soddisfacente all'uguaglianza $\mathbf{CC}^H = \mathbf{I}$, allora $\mathbf{A}^+ = \mathbf{C}^H\mathbf{B}^+$.

A chiusura del Paragrafo, osserviamo che la proprietà della matrice pseudoinversa $\mathbf{A}^+$ di una matrice $m \times n$ $\mathbf{A}$ che più la avvicina all'inversa è la seguente: l'applicazione $f_\mathbf{A} \colon \mathbb{C}^n \to \mathbb{C}^m$ indotta dalla moltiplicazione per $\mathbf{A}$ e l'applicazione $f_{\mathbf{A}^+} \colon \mathbb{C}^m \to \mathbb{C}^n$ indotta dalla moltiplicazione per $\mathbf{A}^+$ sono due funzioni che risultano una l'inversa dell'altra se ristrette rispettivamente allo spazio delle righe $C(\mathbf{A}^H)$ e allo spazio delle colonne $C(\mathbf{A})$ di $\mathbf{A}$ (si veda il Paragrafo 4 del Capitolo II). Per verificare questa proprietà è sufficiente sapere che $\mathbf{y} \in C(\mathbf{A})$ se e solo se $\mathbf{y} = \mathbf{A}\mathbf{x}$ per un vettore $\mathbf{x} \in \mathbb{C}^n$, e similmente $\mathbf{x} \in C(\mathbf{A}^H)$ se e solo se $\mathbf{x} = \mathbf{A}^H\mathbf{y}$ per un vettore $\mathbf{y} \in \mathbb{C}^m$. Atteso ciò, risulta, per ogni vettore $\mathbf{y} \in C(\mathbf{A})$ e ogni vettore $\mathbf{x} \in C(\mathbf{A}^H)$:

$$f_\mathbf{A}(f_{\mathbf{A}^+}(\mathbf{y}) = \mathbf{A}\mathbf{A}^+\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x} = \mathbf{y}$$

$$f_{\mathbf{A}^+}(f_\mathbf{A}(\mathbf{x}) = \mathbf{A}^+\mathbf{A}\mathbf{A}^H\mathbf{y} = \mathbf{A}^H\mathbf{y} = \mathbf{x}.$$

5. Eliminazione di Gauss per il calcolo delle matrici inverse

Si è visto nel Paragrafo 4 che, se si conosce una inversa destra $\mathbf{R}$ della matrice $\mathbf{A}$, si ricava subito una soluzione $\mathbf{v}$ del sistema lineare $\mathbf{A}\mathbf{x} = \mathbf{b}$: essa è data da $\mathbf{v} = \mathbf{R}\mathbf{b}$. In particolare, se la matrice $\mathbf{A}$ è quadrata e invertibile, l'unica soluzione del sistema è $\mathbf{A}^{-1}\mathbf{b}$. Pertanto, pur non essendo l'esistenza di una inversa destra di $\mathbf{A}$ condizione necessaria per la risolubilità del sistema, sembra che un modo per risolverlo, alternativo all'algoritmo della EG, sia quello di cercare una inversa destra della matrice dei coefficienti.

In realtà è vero il contrario, cioè la ricerca di una inversa destra (o bilatera, nel caso quadrato) di una matrice viene ricondotta alla soluzione di una serie di sistemi lineari; tramite l'*algoritmo di Gauss-Jordan*, che perfeziona la EG, questa serie di sistemi viene trattata in blocco e produce la inversa.

Vediamo dapprima su di un esempio come funziona tale algoritmo per una matrice quadrata, dopo di che cercheremo di capire perché esso produce la matrice inversa; successivamente lo estenderemo alla matrici rettangolari.

Esempio 5.1. Sia data la matrice 3×3:

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

Si consideri la matrice a blocchi 3×6:

$$\begin{bmatrix} \mathbf{A} \mid \mathbf{I}_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ -1 & 0 & -2 & 0 & 0 & 1 \end{array}\right]$$

ottenuta affiancando a destra alla matrice $\mathbf{A}$ la matrice identità $\mathbf{I}_3$. Lo scopo dell'algoritmo di Gauss-Jordan è di modificare con le operazioni elementari tale matrice a blocchi fino a ottenere una matrice del tipo

$$\begin{bmatrix} \mathbf{I}_3 \mid \mathbf{B} \end{bmatrix}.$$

La matrice $\mathbf{B}$ risulterà essere la matrice inversa della matrice $\mathbf{A}$.

Si esegue quindi la usuale EG sulla matrice $\begin{bmatrix} \mathbf{A} \mid \mathbf{I}_3 \end{bmatrix}$ pervenendo alla matrice

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right].$$

Il blocco di sinistra è una forma ridotta di $\mathbf{A}$; ora inizia la parte nuova dell'algoritmo. Si procede quindi con quella che si chiama *eliminazione all'indietro* che, tramite le solite operazioni elementari, pone degli zeri *al di sopra* degli elementi diagonali, partendo dall'ultima riga e risalendo fino alla prima.

Utilizzando come pivot il coefficiente di posto $(3,3)$, si sostituisce quindi alla seconda riga la sua somma con la terza moltiplicata per $-1/2$, e alla prima riga la sua somma con la terza moltiplicata per -2; si ottiene la matrice

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & -3 & -2 & -4 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right].$$

Ora, utilizzando il coefficiente di posto $(2,2)$, si sostituisce alla prima riga la sua somma con la seconda moltiplicata per 1, ottenendo la matrice

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & -2 & -5 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right].$$

La matrice ha assunto quindi la forma desiderata $\left[\mathbf{I}_3 \mid \mathbf{B}\right]$, con

$$\mathbf{B} = \begin{bmatrix} -4 & -2 & -5 \\ -1 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix}.$$

La matrice $\mathbf{B}$ così ottenuta coincide con l'inversa $\mathbf{A}^{-1}$: il lettore verifichi infatti che $\mathbf{AB} = \mathbf{I}_3$. $\square$

Cerchiamo ora di capire, in tutta generalità, perché l'algoritmo di Gauss-Jordan produce la matrice inversa. La spiegazione è fornita nella dimostrazione della seguente proposizione.

PROPOSIZIONE 5.2. *Sia* $\mathbf{A}$ *una matrice quadrata* $n \times n$. *Se* $\mathbf{A}$ *è invertibile, la sua inversa* $\mathbf{A}^{-1}$ *si ottiene a partire dalla matrice a blocchi* $\left[\mathbf{A} \mid \mathbf{I}_n\right]$, *operando su di essa con l'algoritmo di Gauss-Jordan fino a trasformarla nella forma* $\left[\mathbf{I}_n \mid \mathbf{B}\right]$. *Allora la matrice inversa* $\mathbf{A}^{-1}$ *coincide con la matrice* $\mathbf{B}$.

DIMOSTRAZIONE. Mettiamo in evidenza le colonne della matrice $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1} = [\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_n].$$

L'uguaglianza $\mathbf{AA}^{-1} = \mathbf{I}_n$ si può scrivere come una serie di n uguaglianze

$$\mathbf{Av}_i = \mathbf{e}_i \qquad (i = 1, 2, \dots, n)$$

giacché $\mathbf{Av}_i$ è la i-esima colonna della matrice prodotto $\mathbf{AA}^{-1}$ ed $\mathbf{e}_i$ è la i-esima colonna della matrice identità. Perciò trovare $\mathbf{A}^{-1}$ significa risolvere gli n sistemi lineari, ciascuno in n equazioni ed n incognite: $\mathbf{Ax} = \mathbf{e}_i$, $(i = 1, 2, \dots, n)$.

Questi sistemi sono tra di loro "parenti stretti", poiché condividono la matrice dei coefficienti, mentre i loro termini noti sono gli n vettori coordinati $\mathbf{e}_1, \dots, \mathbf{e}_n$. Le operazioni elementari che si eseguono sulle matrici aumentate degli n sistemi

$$\left[\mathbf{A} \mid \mathbf{e}_1\right], \quad \left[\mathbf{A} \mid \mathbf{e}_2\right], \quad \left[\mathbf{A} \mid \mathbf{e}_n\right]$$

sono sempre le stesse, per cui è conveniente considerare una unica matrice, detta *matrice pluri-aumentata* di $\mathbf{A}$

$$\left[\mathbf{A} \mid \mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n\right]$$

Se si eseguono in un sol colpo su tale matrice le operazioni elementari che si eseguirebbero per gli n sistemi, quindi se si opera la EG cui si fanno seguire la eliminazione all'indietro, si ottiene la matrice

$$\left[\mathbf{I}_n \mid \mathbf{B}\right] = \left[\mathbf{I}_n \mid \mathbf{b}_1 \quad \dots \quad \mathbf{b}_n\right]$$

dove i vettori $\mathbf{b}_i$ sono le colonne della matrice $\mathbf{B}$. Ciò significa che gli n sistemi $\mathbf{A}\mathbf{x} = \mathbf{e}_i$ sono stati trasformati negli n sistemi equivalenti

$$\mathbf{I}_n\mathbf{x} = \mathbf{b}_i \qquad (i = 1, 2, \dots, n).$$

Evidentemente le soluzioni di questi sistemi sono proprio i vettori $\mathbf{b}_i$, perciò per ogni $i = 1, 2, \dots, n$ risulta $\mathbf{A}\mathbf{b}_i = \mathbf{e}_i$. Ne consegue che $\mathbf{v}_i = \mathbf{b}_i$ per ogni $i = 1, 2, \dots, n$, pertanto $\mathbf{A}^{-1} = \mathbf{B}$. $\square$

Nel corso dell'algoritmo di Gauss-Jordan applicato a una matrice invertibile può accadere che necessitino scambi di righe; ciò non ha però alcun effetto negativo sull'algoritmo, perché le soluzioni non sono alterate dagli scambi di riga.

Se si applica l'algoritmo di Gauss-Jordan a una matrice non invertibile $n \times n$ $\mathbf{A}$, la prima parte dell'algoritmo consistente nella usuale EG modificherà la matrice pluri-aumentata $\begin{bmatrix}\mathbf{A} \mid \mathbf{I}_n\end{bmatrix}$ in una matrice $\begin{bmatrix}\mathbf{U} \mid \mathbf{C}\end{bmatrix}$, dove $\mathbf{U}$ non sarà più triangolare superiore con elementi diagonali non-nulli (come nell'Esempio 5.1), perché altrimenti completando l'algoritmo si otterrebbe l'inversa che non esiste. La matrice $\mathbf{U}$ sarà solo a scala per righe con almeno una riga nulla; perciò la successiva parte dell'algoritmo, cioè l'eliminazione all'indietro, non potrà in alcun modo trasformare $\mathbf{U}$ in $\mathbf{I}_n$.

Una volta acquisito l'algoritmo di Gauss-Jordan, si ha un metodo costruttivo per vedere se una matrice quadrata ha inversa e, qualora ciò accada, di calcolare effettivamente tale inversa. Questo metodo è di gran lunga più economico rispetto ad altri possibili metodi, quali quello del calcolo della matrice aggiunta che fa uso dei determinanti (si veda il Capitolo IV).

Il metodo che abbiamo a disposizione può essere modificato facilmente per calcolare le matrici inverse destre di una matrice $m \times n$ di rango $m < n$, qualora esse si presentino in una forma opportuna.

Proposizione 5.3. *Sia $\mathbf{A} = \begin{bmatrix}\mathbf{B} \mid \mathbf{C}\end{bmatrix}$ una matrice $m \times n$, con $m < n$, decomposta in due blocchi: il blocco $\mathbf{B}$ quadrato $m \times m$ invertibile, e il blocco $\mathbf{C}$ $m \times (n-m)$. Allora le matrici inverse destre della matrice $\mathbf{A}$ sono tutte e sole le matrici $\mathbf{R}(\mathbf{Y})$ che si ottengono nel modo seguente: si sceglie una arbitraria matrice $(n-m) \times m$ $\mathbf{Y}$ e si pone*

$$\mathbf{R}(\mathbf{Y}) = \begin{bmatrix}\mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y}\end{bmatrix}.$$

Dimostrazione. Il prodotto a blocchi di $\mathbf{A}$ con $\mathbf{R}(\mathbf{Y})$ porge

$$\begin{aligned}\mathbf{A} \cdot \mathbf{R}(\mathbf{Y}) &= \begin{bmatrix}\mathbf{B} \mid \mathbf{C}\end{bmatrix} \begin{bmatrix}\mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y}\end{bmatrix} \\ &= \mathbf{B}\mathbf{B}^{-1} - \mathbf{B}\mathbf{B}^{-1}\mathbf{C}\mathbf{Y} + \mathbf{C}\mathbf{Y} \\ &= \mathbf{I}_m - \mathbf{C}\mathbf{Y} + \mathbf{C}\mathbf{Y} \\ &= \mathbf{I}_m.\end{aligned}$$

Pertanto tutte le matrici $\mathbf{R}(\mathbf{Y})$ definite in questo modo sono inverse destre di $\mathbf{A}$. Viceversa, se $\mathbf{R}$ è inversa destra di $\mathbf{A}$, decomponendola a blocchi nella forma

$$\mathbf{R} = \begin{bmatrix}\mathbf{X} \\ \mathbf{Y}\end{bmatrix}$$

con $\mathbf{X}$ e $\mathbf{Y}$ blocchi di dimensioni $m \times m$ e $(n-m) \times m$, rispettivamente, si ha:

$$\begin{aligned}&\mathbf{A}\mathbf{R} = \mathbf{I}_m \implies \mathbf{B}\mathbf{X} + \mathbf{C}\mathbf{Y} = \mathbf{I}_m \implies \\ &\qquad\implies \mathbf{B}^{-1}\mathbf{B}\mathbf{X} + \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} = \mathbf{B}^{-1} \implies \mathbf{X} = \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y}\end{aligned}$$

quindi $\mathbf{R} = \mathbf{R}(\mathbf{Y})$. $\square$

Diamo un esempio del metodo descritto nella precedente proposizione.

Esempio 5.4. Si vogliono trovare tutte le inverse destre della matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 5 & 2 \end{bmatrix}.$$

Decomponiamo $\mathbf{A}$ a blocchi:

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} \mid \mathbf{C} \end{bmatrix}$$

dove

$$\mathbf{B} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Poiché $\mathbf{B}$ è invertibile, si può applicare la Proposizione 5.3. Sia quindi $\mathbf{Y} = [a\ b]$ una qualunque matrice 1×2; tenuto conto che

$$\mathbf{B}^{-1} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$$

(si ricordi l'Esempio 4.17) le inverse destre della matrice $\mathbf{A}$ sono allora le matrici

$$\begin{bmatrix} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y} \end{bmatrix}$$

dove

$$\mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} - \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} [a\ b].$$

Pertanto tutte e sole le matrici inverse destre della matrice $\mathbf{A}$ sono della forma

$$\begin{bmatrix} 5+a & -3+b \\ -3-a & 2-b \\ a & b \end{bmatrix}$$

con a e b scalari arbitrari. □

Per risolvere completamente il problema del calcolo dell'inversa destra di una matrice $m \times n$ di rango $m < n$, basterà trovare un metodo per trasformarla in una matrice soddisfacente all'ipotesi della Proposizione 5.3 e, una volta trovata l'inversa destra di quest'ultima matrice, risalire all'inversa destra della matrice di partenza. Questo metodo richiede alcuni risultati sulle matrici di permutazione che saranno provati più avanti, perciò non lo svilupperemo qui (si veda l'Esercizio 1.33 di questo Capitolo). Mostriamo invece su di un esempio un modo più semplice e diretto per il calcolo delle inverse destre, basato sulla soluzione di sistemi lineari.

Esempio 5.5. Si vogliono trovare tutte le inverse destre della matrice

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

È facile verificare che il blocco di $\mathbf{A}$ formato dalle prime tre colonne non è invertibile, perciò la Proposizione 5.3 non è applicabile. Si tratta di trovare tutte le matrici 5×3 $\mathbf{R} = [\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3]$ tali che $\mathbf{A}\mathbf{R} = \mathbf{I}_3$ o, equivalentemente, si tratta di risolvere i tre sistemi lineari $\mathbf{A}\mathbf{x} = \mathbf{e}_j$ $(j = 1, 2, 3)$. Conviene a tal fine considerare la matrice pluri-aumentata

$$\begin{bmatrix} \mathbf{A} \mid \mathbf{I}_3 \end{bmatrix} = \left[\begin{array}{ccccc|ccc} 1 & -2 & -1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & 2 & 0 & -1 & -2 & 0 & 0 & 1 \end{array}\right].$$

La EG su tale matrice produce la forma ridotta

$$\left[\begin{array}{ccccc|ccc} 1 & -2 & -1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & -1 & -2 \end{array}\right].$$

Con facili calcoli si vede che il sistema $\mathbf{Ax} = \mathbf{e}_1$ ha le infinite soluzioni dipendenti dai parametri h_1, k_1:

$$x_5 = -2, \quad x_4 = h_1, \quad x_3 = -1 - h_1, \quad x_2 = k_1, \quad x_1 = 4 - h_1 + 2k_1.$$

Analogamente, il sistema $\mathbf{Ax} = \mathbf{e}_2$ ha le infinite soluzioni dipendenti dai parametri h_2, k_2:

$$x_5 = -1, \quad x_4 = h_2, \quad x_3 = -h_2, \quad x_2 = k_2, \quad x_1 = 2 - h_2 + 2k_2.$$

Infine, il sistema $\mathbf{Ax} = \mathbf{e}_3$ ha le infinite soluzioni dipendenti dai parametri h_3, k_3:

$$x_5 = -2, \quad x_4 = h_3, \quad x_3 = -1 - h_3, \quad x_2 = k_3, \quad x_1 = 3 - h_3 + 2k_3.$$

In definitiva, le infinite inverse destre della matrice $\mathbf{A}$ sono tutte e sole le matrici dipendenti da 6 parametri:

$$\begin{bmatrix} 4 - h_1 + 2k_1 & 2 - h_2 + 2k_2 & 3 - h_3 + 2k_3 \\ k_1 & k_2 & k_3 \\ -1 - h_1 & -h_2 & -1 - h_3 \\ h_1 & h_2 & h_3 \\ -2 & -1 & -2 \end{bmatrix}.$$

□

Resta infine da esaminare come si trovano le inverse sinistre di una matrice $m \times n$ $\mathbf{A}$ di rango n, con $n > m$. Basta per ciò osservare che una matrice $\mathbf{L}$ è inversa sinistra di $\mathbf{A}$ se e solo se la matrice $\mathbf{L}^T$, trasposta di $\mathbf{L}$, è inversa destra della matrice $\mathbf{A}^T$. Perciò si determinano le inverse destre di $\mathbf{A}^T$ col procedimento sopra descritto e poi le si traspone, ottenendo tutte e sole le inverse sinistre di $\mathbf{A}$.

6. Matrici elementari e decomposizione LU

Scopo di questo Paragrafo è mostrare come la EG su di una matrice $\mathbf{A}$ produce una fattorizzazione della matrice stessa del tipo $\mathbf{A} = \mathbf{LU}$, dove la matrice $\mathbf{U}$ è una forma ridotta di $\mathbf{A}$, mentre la matrice $\mathbf{L}$ è una matrice triangolare inferiore invertibile, qualora non siano intervenuti scambi di righe. In presenza di scambi di righe questa fattorizzazione si ricava non per $\mathbf{A}$, ma per $\mathbf{PA}$, dove $\mathbf{P}$ è una opportuna matrice di permutazione (la cui definizione sarà data tra poco). Il simbolo $\mathbf{L}$ deriva dal termine inglese “lower”, che significa “inferiore”. La fattorizzazione prodotta si chiama *decomposizione* **LU** della matrice $\mathbf{A}$ (o di $\mathbf{PA}$).

Tale decomposizione di $\mathbf{A}$ deriva dal fatto che ogni singola operazione elementare che si esegue nel corso della EG corrisponde alla pre-moltiplicazione per una opportuna matrice invertibile; queste matrici si chiamano appunto *matrici elementari*. Cominciamo pertanto col presentare queste matrici, che sono di tre tipi diversi.

Matrici elementari $\mathbf{E}_{ij}(\alpha)$. Queste matrici corrispondono all'operazione elementare più importante della EG, che consiste nel sommare una riga con un multiplo non nullo di una riga che le sta sopra; è questa operazione che permette di far diventare 0 i coefficienti sotto ai pivot. Esse sono matrici quadrate, definite a partire da uno scalare $\alpha \neq 0$ e da una coppia di indici distinti $i \neq j$. Una matrice elementare $m \times m$ $\mathbf{E}_{ij}(\alpha)$ differisce dalla matrice identità I_m solo nel coefficiente di posto (i, j), dove anziché 0 compare lo scalare α. Conviene scrivere la matrice $\mathbf{E}_{ij}(\alpha)$ a blocchi riga e blocchi colonna:

$$\mathbf{E}_{ij}(\alpha) = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_i^T + \alpha\mathbf{e}_j^T \\ \vdots \\ \mathbf{e}_m^T \end{bmatrix} = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_j + \alpha\mathbf{e}_i\ \ldots\ \mathbf{e}_m]$$

dove la riga $\mathbf{e}_i^T + \alpha\mathbf{e}_j^T$ compare all'i-esimo posto, mentre la colonna $\mathbf{e}_j + \alpha\mathbf{e}_i$ compare al posto j-esimo. Si noti che $\mathbf{E}_{ij}(\alpha)$ è uni-triangolare (inferiore, se $i > j$).

Matrici elementari $\mathbf{E}_{ij}$. Queste matrici corrispondono all'operazione di scambio di due righe nella EG e sono chiamate *matrici di trasposizione* (o *di scambio*). Sono matrici quadrate definite a partire da una coppia (i, j) di indici distinti, che indicano le due righe che vengono scambiate. Una matrice elementare $m \times m$ di questo tipo differisce dalla matrice identità I_m per il fatto che sono scambiate tra di loro in I_m la riga i-esima e la riga j-esima o, equivalentemente, la colonna i-esima e la colonna j-esima. Si noti che tali matrici non sono matrici triangolari.

Matrici elementari $\mathbf{E}_i(\alpha)$. Queste matrici corrispondono all'operazione elementare di moltiplicazione della riga i-esima per lo scalare non nullo α. Sono matrici diagonali (già incontrate nel Paragrafo 1), che hanno sulla diagonale tutti 1 tranne che nel posto i-esimo, dove compare il coefficiente $\alpha \neq 0$.

ESEMPIO 6.1. Esempi di matrici elementari dei tre tipi descritti nel caso di dimensione 3 sono:

$$\mathbf{E}_{31}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \qquad \mathbf{E}_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{E}_2(5) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \square$$

Diamo di seguito alcune proprietà cui soddisfano le matrici elementari.

(I) Tutte le matrici elementari sono matrici invertibili e risulta

$$\mathbf{E}_{ij}(\alpha)^{-1} = \mathbf{E}_{ij}(-\alpha); \qquad \mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij}; \qquad \mathbf{E}_i(\alpha)^{-1} = \mathbf{E}_i(\alpha^{-1}).$$

Verifichiamo la prima uguaglianza:

$$\mathbf{E}_{ij}(\alpha)\mathbf{E}_{ij}(-\alpha) = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \cdots \\ \mathbf{e}_i^T + \alpha\mathbf{e}_j^T \\ \cdots \\ \mathbf{e}_m^T \end{bmatrix} [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_j - \alpha\mathbf{e}_i\ \ldots\ \mathbf{e}_m]$$

dove la riga $\mathbf{e}_i^T + \alpha\mathbf{e}_j^T$ compare all'i-esimo posto, mentre la colonna $\mathbf{e}_j - \alpha\mathbf{e}_i$ compare al posto j-esimo. Vogliamo provare che questo prodotto coincide con $\mathbf{I}_m$. Tenuto conto che $\mathbf{e}_i^T\mathbf{e}_i = 1$ e $\mathbf{e}_i^T\mathbf{e}_j = 0$ se $i \neq j$, si vede che il coefficiente del prodotto di posto (i, j) risulta uguale a

$$(\mathbf{e}_i^T + \alpha\mathbf{e}_j^T)(\mathbf{e}_j - \alpha\mathbf{e}_i) = -\alpha + \alpha = 0.$$

La verifica per gli altri coefficienti è immediata.

Le verifiche della seconda e terza uguaglianza sono facili esercizi lasciati al lettore.

(II) Per le trasposte delle matrici elementari si ha:

$$\mathbf{E}_{ij}(\alpha)^T = \mathbf{E}_{ji}(\alpha); \quad \mathbf{E}_{ij}^T = \mathbf{E}_{ij}; \quad \mathbf{E}_i(\alpha)^T = \mathbf{E}_i(\alpha).$$

Le verifiche sono immediate.

(III) Un prodotto di matrici elementari del tipo $\mathbf{E}_{ij}(\alpha)$ con $i > j$ è una matrice uni-triangolare inferiore.

Segue dal fatto che, se $i > j$, ogni matrice $\mathbf{E}_{ij}(\alpha)$ è una matrice uni-triangolare inferiore, e dall'Esempio 1.7.

(IV) Un prodotto di matrici elementari del tipo $\mathbf{E}_{ij}$ è una matrice che si ottiene dalla matrice identità permutandone le righe (o le colonne) in modo opportuno.

La dimostrazione è immediata conseguenza della successiva Proposizione 6.3 (b) (o della Proposizione 6.5 (b)).

Le matrici considerate nel precedente punto (IV), cioè le matrici che si ottengono moltiplicando tra di loro un certo numero di matrici di trasposizione, sono chiamate *matrici di permutazione*. Esse sono invertibili (in quanto prodotto di matrici invertibili) e sono caratterizzate dalla proprietà di avere in ogni riga e in ogni colonna tutti gli elementi nulli, tranne uno che è uguale a 1.

Esempio 6.2. Le permutazioni di n oggetti sono $n!$. Perciò esistono $3! = 6$ matrici di permutazione 3×3 distinte, che sono

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

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \qquad \square$$

Vogliamo ora esaminare quale è il risultato della pre-moltiplicazione per una matrice elementare su di una arbitraria matrice. L'Esempio 1.8 ci dice cosa accade pre-moltiplicando la matrice $\mathbf{A}$ per una matrice del tipo $\mathbf{E}_i(\alpha)$: poiché $\mathbf{E}_i(\alpha)$ è una matrice diagonale, $\mathbf{E}_i(\alpha)\mathbf{A}$ coincide con $\mathbf{A}$ tranne che per la i-esima riga, che si ottiene dalla i-esima riga di $\mathbf{A}$ moltiplicandola per α. Se per esempio nella i-esima riga di $\mathbf{A}$ c'è un certo pivot $d_i \neq 0$, pre-moltiplicando $\mathbf{A}$ per $\mathbf{E}_i(d_i^{-1})$ si rende quel pivot uguale a 1. Ciò costituisce il punto (c) della seguente Proposizione 6.3; ciò che accade pre-moltiplicando per una matrice elementare degli altri due tipi è descritto nei punti (a) e (b).

Proposizione 6.3. *Sia* $\mathbf{A}$ *una matrice* $m \times n$.

(a) *La pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_{ij}(\alpha)$ *modifica la sola riga* i*-esima di* $\mathbf{A}$*, che diventa uguale alla somma della riga* i*-esima con la riga* j*-esima moltiplicata per* α.
(b) *La pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_{ij}$ *scambia tra di loro la riga* i*-esima e la riga* j*-esima di* $\mathbf{A}$ *e lascia invariate le altre righe.*
(c) *La pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_i(\alpha)$ *moltiplica la riga* i*-esima di* $\mathbf{A}$ *per* α *e lascia invariate le altre righe.*

Dimostrazione. (a) Eseguiamo il prodotto $\mathbf{E}_{ij}(\alpha)\mathbf{A}$ decomponendo $\mathbf{E}_{ij}(\alpha)$ in blocchi riga:

$$\mathbf{E}_{ij}(\alpha)\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \alpha\mathbf{e}_j^T + \mathbf{e}_i^T \\ \vdots \\ \mathbf{e}_m^T \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T\mathbf{A} \\ \mathbf{e}_2^T\mathbf{A} \\ \vdots \\ (\alpha\mathbf{e}_j^T + \mathbf{e}_i^T)\mathbf{A} \\ \vdots \\ \mathbf{e}_m^T\mathbf{A} \end{bmatrix}.$$

L'asserto segue subito non appena si ricordi che nella Proposizione 2.13 (b) si è visto che $\mathbf{e}_i^T\mathbf{A}$ coincide con la i-esima riga di $\mathbf{A}$.

(b) Analoga alla precedente, tenendo conto che la i-esima riga di $\mathbf{E}_{ij}$ coincide con $\mathbf{e}_j^T$ e la j-esima riga coincide con $\mathbf{e}_i^T$.

(c) Visto sopra. □

Osservazione 6.4. Possiamo dare ora la dimostrazione del Lemma 4.5, cioè del fatto che, data una matrice $\mathbf{A}$, esiste una matrice invertibile $\mathbf{E}$ tale che

$$\mathbf{EA} = \begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix}.$$

Se la prima colonna di $\mathbf{A}$ è nulla, si ponga $\mathbf{E} = \mathbf{I}$, altrimenti $\mathbf{E}$ non è altro che il prodotto delle matrici elementari utilizzate per mettere 0 sotto al primo pivot a.

Per quanto riguarda la post-moltiplicazione di una matrice $\mathbf{A}$ per matrici elementari, vale un risultato analogo alla Proposizione 6.3, che riguarda non più le righe, ma le colonne di $\mathbf{A}$. La sua dimostrazione è analoga a quella della Proposizione 6.3, utilizzando la Proposizione 2.13 (a), oppure si può ricavare passando alle matrici trasposte, ed è lasciata come esercizio.

Proposizione 6.5. *Sia $\mathbf{A}$ una matrice $m \times n$.*

(a) *La post-moltiplicazione di $\mathbf{A}$ per la matrice $n \times n$ $\mathbf{E}_{ij}(\alpha)$ modifica la sola colonna j-esima di $\mathbf{A}$, che diventa uguale alla somma della colonna j-esima con la colonna i-esima moltiplicata per α.*

(b) *La post-moltiplicazione di $\mathbf{A}$ per la matrice $n \times n$ $\mathbf{E}_{ij}$ scambia tra di loro la colonna i-esima e la colonna j-esima di $\mathbf{A}$ e lascia invariate le altre colonne.*

(c) *La post-moltiplicazione di $\mathbf{A}$ per la matrice $n \times n$ $\mathbf{E}_i(\alpha)$ moltiplica la colonna i-esima di $\mathbf{A}$ per α e lascia invariate le altre colonne.* □

La Proposizione 6.3 consente di tradurre in fattorizzazione matriciale la EG eseguita su una matrice.

Teorema 6.6. *Sia $\mathbf{A}$ una matrice $m \times n$. La EG su $\mathbf{A}$, qualora non necessiti di scambi di righe, produce una decomposizione $\mathbf{A} = \mathbf{LU}$, dove $\mathbf{U}$ è una forma ridotta di $\mathbf{A}$ ed $\mathbf{L}$ è una matrice triangolare inferiore invertibile.*

Dimostrazione. La EG presentata nel Paragrafo 4, qualora non siano necessari scambi di righe, consiste nell'applicare successivamente un certo numero di operazioni elementari corrispondenti alla pre-moltiplicazione per matrici elementari di tipo $\mathbf{E}_{ij}(\alpha)$ e di tipo $\mathbf{E}_i(\alpha)$ per scalari $\alpha \neq 0$. Se denotiamo con $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_r$ tali matrici elementari, nell'ordine con cui vengono applicate, risulterà

$$\mathbf{E}_r \cdot \dots \cdot \mathbf{E}_2 \cdot \mathbf{E}_1 \cdot \mathbf{A} = \mathbf{U}$$

dove $\mathbf{U}$ è una forma ridotta di $\mathbf{A}$; si noti che, poiché le matrici elementari operano a sinistra di $\mathbf{A}$, l'ordine di scrittura è inverso di quello con cui agiscono. Sia

$$\mathbf{E} = \mathbf{E}_r \cdot \dots \cdot \mathbf{E}_2 \cdot \mathbf{E}_1$$

il prodotto delle matrici elementari utilizzate. La matrice $\mathbf{E}$, in quanto prodotto di matrici triangolari inferiori invertibili, è triangolare inferiore invertibile. Poniamo $\mathbf{L} = \mathbf{E}^{-1}$; tale matrice è ancora triangolare inferiore, per l'Esempio 4.16, e risulta $\mathbf{A} = \mathbf{LU}$, come si voleva provare. □

Vediamo su di un esempio già utilizzato come avviene la fattorizzazione di una matrice tramite la EG.

Esempio 6.7. Riprendiamo il sistema lineare esaminato nell'Esempio 3.2, di cui consideriamo solo la matrice dei coefficienti (e non quella aumentata):

$$\mathbf{B} = \begin{bmatrix} 2 & 6 & 3 & 2 \\ 1 & -2 & 1/2 & 9/4 \\ -1 & 1 & -1/2 & -1 \end{bmatrix}.$$

Le operazioni elementari applicate in successione a partire dalla matrice $\mathbf{B}$ (si veda l'Esempio 3.1) consistevano nel:

- moltiplicare la prima riga per $1/2$;
- sommare alla seconda riga la prima moltiplicata per -1;
- sommare alla terza riga la prima moltiplicata per 1;
- moltiplicare la seconda riga per $-1/5$;
- sommare alla terza riga la seconda moltiplicata per -4;
- moltiplicare la terza riga per 5.

A tali operazioni corrispondono ordinatamente le pre-moltiplicazioni per le matrici elementari

$$\begin{gathered} \mathbf{E}_1 = \mathbf{E}_1(1/2), \quad \mathbf{E}_2 = \mathbf{E}_{21}(-1), \quad \mathbf{E}_3 = \mathbf{E}_{31}(1), \\ \mathbf{E}_4 = \mathbf{E}_2(-1/5), \quad \mathbf{E}_5 = \mathbf{E}_{32}(-4), \quad \mathbf{E}_6 = \mathbf{E}_3(5). \end{gathered}$$

La forma ridotta cui si perviene al termine della EG nell'Esempio 3.2 (denotata con $\mathbf{V}$) è la seguente:

$$\mathbf{V} = \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}.$$

Un calcolo diretto mostra che

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_6\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1 \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & 0 & 0 \\ 1/10 & -1/5 & 0 \\ 1/2 & 4 & 5 \end{bmatrix}. \end{aligned}$$

Pertanto la matrice $\mathbf{E}$, una volta invertita, dà la matrice triangolare inferiore $\mathbf{L}$ che fattorizza $\mathbf{A}$; il calcolo con l'algoritmo di Gauss-Jordan della matrice $\mathbf{L} = \mathbf{E}^{-1}$ porge

$$\mathbf{L} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -5 & 0 \\ -1 & 4 & 1/5 \end{bmatrix}.$$

Si osservi che

$$\begin{aligned} \mathbf{L} &= \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{E}_4^{-1}\mathbf{E}_5^{-1}\mathbf{E}_6^{-1} \\ &= \mathbf{E}_1(2)\mathbf{E}_{21}(1)\mathbf{E}_{31}(-1)\mathbf{E}_2(-5)\mathbf{E}_{32}(4)\mathbf{E}_3(1/5), \end{aligned}$$

dove i coefficienti non nulli che compaiono ordinatamente dall'alto verso il basso nella prima colonna di $\mathbf{L}$, poi nella seconda colonna e infine nella terza, coincidono con gli scalari che compaiono ordinatamente tra parentesi (da sinistra verso destra) nelle matrici elementari il cui prodotto da appunto la matrice $\mathbf{L}$. Il lettore verifichi che $\mathbf{A} = \mathbf{LU}$. $\square$

A questo punto sorgono naturali due domande. Si possono descrivere facilmente gli elementi della matrice $\mathbf{L}$, a partire dalle operazioni elementari eseguite? E cosa accade se nella EG necessitano permutazioni?

Cominciamo con l'esaminare la seconda domanda. Supponiamo che nel corso della EG sulla matrice $\mathbf{A}$ compaiano, alternate a operazioni elementari corrispondenti a matrici elementari del tipo $\mathbf{E}_{ij}(\alpha)$ ed $\mathbf{E}_i(\alpha)$, anche un certo numero di scambi di righe; supponiamo inoltre che a tali scambi di righe corrispondano ordinatamente le matrici di trasposizione $\mathbf{E}_{i_1 j_1}, \mathbf{E}_{i_2 j_2}, \dots, \mathbf{E}_{i_k j_k}$. Denotiamo con $\mathbf{P}$ la matrice di permutazione ottenuta facendo il prodotto di queste matrici di trasposizione nell'ordine con cui operano a sinistra su $\mathbf{A}$, quindi

$$\mathbf{P} = \mathbf{E}_{i_k j_k} \cdot \dots \mathbf{E}_{i_2 j_2} \cdot \mathbf{E}_{i_1 j_1}.$$

Osserviamo che una matrice di permutazione, in quanto prodotto di matrici invertibili, è invertibile. In queste notazioni vale la seguente proposizione.

Proposizione 6.8. *La EG sulla matrice* $\mathbf{PA}$ *non necessita di scambi di righe. Ovvero, se si eseguono subito su* $\mathbf{A}$ *ordinatamente tutti gli scambi di righe che si utilizzano nel corso della EG e se si procede poi sulla matrice ottenuta con la EG, non sono più necessari scambi di righe.*

Dimostrazione. Siano $\mathbf{E}_{i_1 j_1}, \mathbf{E}_{i_2 j_2}, \dots, \mathbf{E}_{i_k j_k}$ le matrici di scambio che intervengono ordinatamente nella EG; possiamo supporre $i_h < j_h$ per ogni $h \leq k$. Si rende necessario far operare $\mathbf{E}_{i_1 j_1}$ perché dopo avere sistemato le colonne fino alla $i_1 - 1$-esima inclusa, al posto (i_1, i_1) troviamo lo 0 e vogliamo sostituirlo con l'elemento di posto (i_1, j_1). Prima di $\mathbf{E}_{i_1 j_1}$ hanno operato matrici del tipo $\mathbf{E}_{rs}(\alpha)$ e $\mathbf{E}_r(\alpha)$ con indice di colonna $s < i_1$. È facile verificare (si veda l'Esercizio ??) che, se $s < i < j$, risulta

$$\mathbf{E}_{ij}\mathbf{E}_{is}(\alpha) = \mathbf{E}_{js}(\alpha)\mathbf{E}_{ij}, \qquad \mathbf{E}_{ij}\mathbf{E}_{js}(\alpha) = \mathbf{E}_{is}(\alpha)\mathbf{E}_{ij}$$

mentre se $r \neq i, j$, allora

$$\mathbf{E}_{ij}\mathbf{E}_{rs}(\alpha) = \mathbf{E}_{rs}(\alpha)\mathbf{E}_{ij}.$$

Pertanto si può permutare $\mathbf{E}_{i_1 j_1}$ con tutte le matrici elementari che hanno operato prima di essa, alterando possibilmente gli indici di riga di tali matrici ma non gli indici delle colonne, che restano sempre minori di i_1. Effettuiamo la stessa procedura sulla matrice $\mathbf{E}_{i_2 j_2}$. Permutiamola quindi con tutte le matrici di tipo $\mathbf{E}_{rs}(\alpha)$ e $\mathbf{E}_r(\alpha)$ che operano prima di essa, di cui quelle che precedevano $\mathbf{E}_{i_1 j_1}$ sono state modificate. Ma tutte tali matrici hanno indici di colonna minori di i_2 e tale proprietà si conserva dopo le permutazioni per le formule viste sopra. Continuando in questo modo, si portano a operare per prime sulla matrice $\mathbf{A}$ ordinatamente le matrici $\mathbf{E}_{i_1 j_1}, \mathbf{E}_{i_2 j_2}, \dots, \mathbf{E}_{i_k j_k}$, il cui prodotto nell'ordine rovesciato con cui sono premoltiplicate ad $\mathbf{A}$ è $\mathbf{P} = \mathbf{E}_{i_k j_k} \cdot \dots \mathbf{E}_{i_2 j_2} \cdot \mathbf{E}_{i_1 j_1}$. Le matrici che ora pre-moltiplicano $\mathbf{PA}$, il cui risultato è la forma ridotta $\mathbf{U}$, sono matrici elementari del tipo $\mathbf{E}_{rs}(\alpha)$ con $r > s$ e $\mathbf{E}_r(\alpha)$, da cui l'asserto. $\square$

Dalla Proposizione 6.8 si può trarre immediatamente il seguente risultato.

Teorema 6.9. *Sia* $\mathbf{A}$ *una matrice* $m \times n$. *Qualora la EG su* $\mathbf{A}$ *necessiti di scambi di righe, si possono eseguire subito tali scambi su* $\mathbf{A}$ *ottenendo la nuova matrice* $\mathbf{PA}$, *ove* $\mathbf{P}$ *denota una matrice di permutazione. La EG su* $\mathbf{PA}$ *produce una fattorizzazione* $\mathbf{PA} = \mathbf{LU}$, *dove* $\mathbf{U}$ *è una forma ridotta di* $\mathbf{A}$ *ed* $\mathbf{L}$ *è una matrice triangolare inferiore invertibile. Pertanto* $\mathbf{A} = \mathbf{P}^{-1}\mathbf{LU}$. $\square$

Si noti che nel Teorema 6.9 la matrice $\mathbf{P}^{-1}\mathbf{L}$ è una matrice invertibile e che

$$\mathbf{P} = \mathbf{E}_{i_k j_k} \cdot \ldots \mathbf{E}_{i_2 j_2} \cdot \mathbf{E}_{i_1 j_1} \implies \mathbf{P}^{-1} = \mathbf{E}_{i_1 j_1}\mathbf{E}_{i_2 j_2} \cdot \ldots \cdot \mathbf{E}_{i_k j_k} = \mathbf{P}^T$$

in virtù del fatto che $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij} = \mathbf{E}_{ij}^T$ per ogni matrice di scambio $\mathbf{E}_{ij}$ (si veda l'Esercizio ??). Il Teorema 6.9 fornisce inoltre la dimostrazione del Lemma 4.5, (ii).

Forniamo un esempio della situazione descritta nel Teorema 6.9.

Esempio 6.10. Consideriamo la matrice

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

Sommando alla seconda riga l'opposto della prima, alla terza riga la prima moltiplicata per -2, e alla quarta riga l'opposto della prima, si ricava la matrice

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

Eseguiamo prima lo scambio di seconda e terza riga, poi nella matrice ottenuta lo scambio tra terza e quarta riga, ottenendo la matrice

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

dalla quale, moltiplicando la terza riga per $1/3$, si ottiene la forma ridotta

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Le operazioni che sono state eseguite corrispondono alla pre-moltiplicazione di $\mathbf{A}$ per opportune matrici elementari; più precisamente, la matrice ottenuta alla fine è la matrice

$$\mathbf{E}_3(1/3)\mathbf{E}_{34}\mathbf{E}_{23}\mathbf{E}_{41}(-1)\mathbf{E}_{31}(-2)\mathbf{E}_{21}(-1)\mathbf{A}.$$

Il Teorema6.9 ci dice che pre-moltiplicando $\mathbf{A}$ per la matrice di permutazione

$$\mathbf{P} = \mathbf{E}_{34}\mathbf{E}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

la nuova matrice

$$\mathbf{PA} = \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 2 & 5 & -1 & 3 & 4 \\ 1 & 2 & 5 & 1 & 2 \\ 1 & 2 & 2 & 2 & 2 \end{bmatrix}$$

non avrà bisogno, nell'eseguire la EG, di alcuno scambio di righe. In effetti, la EG su $\mathbf{PA}$ porge

$$\mathbf{E}_3(1/3)\mathbf{E}_{41}(-1)\mathbf{E}_{31}(-1)\mathbf{E}_{21}(-2)\mathbf{PA} = \mathbf{U}.$$

Quindi si ha la decomposizione

$$\mathbf{PA} = \mathbf{LU}$$

dove $\mathbf{L}$ è la matrice triangolare inferiore data da

$$\mathbf{L} = \mathbf{E}_{21}(2)\mathbf{E}_{31}(1)\mathbf{E}_{41}(1)\mathbf{E}_{3}(3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Ne consegue che per $\mathbf{A}$ si ha la seguente fattorizzazione:

$$\begin{aligned} \mathbf{A} &= \mathbf{P}^{-1}\mathbf{L}\mathbf{U} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \end{aligned}$$

Si noti che il primo fattore nell'ultimo prodotto non è una matrice triangolare inferiore; resta però una matrice invertibile. □

Ritorniamo ora alla prima domanda posta in precedenza: è possibile determinare facilmente i coefficienti della matrice triangolare inferiore $\mathbf{L}$ che fattorizza la matrice $\mathbf{A}$? In virtù del Teorema 6.9, ci limiteremo a considerare il caso in cui non necessitano scambi di righe; se infatti intervengono scambi, fattorizzeremo la matrice $\mathbf{PA}$ per la quale la EG non necessita di scambi di righe.

La risposta alla precedente domanda è positiva, come suggerisce l'osservazione alla fine dell'Esempio 6.7; essa dipende dal seguente lemma tecnico. Per renderne più scorrevole l'enunciato, conveniamo di denotare con $\mathbf{E}_{ii}(\alpha)$ la matrice elementare $\mathbf{E}_{i}(\alpha)$, per cui tra le matrici elementari $\mathbf{E}_{ij}(\alpha)$ considerate nel lemma potrà succedere che $i = j$. Introduciamo inoltre la seguente definizione: diremo che la coppia ordinata di numeri interi positivi (i_1, j_1) è minore della coppia (i_2, j_2) nell'ordine *anti-lessicografico* se $j_1 < j_2$, oppure, nel caso in cui $j_1 = j_2$, se $i_1 < i_2$.

Lemma 6.11. *Se la matrice triangolare inferiore* $\mathbf{L} = [l_{ij}]$ *è prodotto di matrici elementari triangolari inferiori*

$$\mathbf{L} = \mathbf{E}_{i_1 j_1}(\alpha_1)\mathbf{E}_{i_2 j_2}(\alpha_2)\dots\mathbf{E}_{i_r j_r}(\alpha_r)$$

con gli scalari $\alpha_h \neq 0$, *e le coppie ordinate di indici* (i_h, j_h) $(h \leq r)$ *si susseguono in modo crescente nell'ordine anti-lessicografico, allora* $l_{i_h j_h} = \alpha_h$ *per ogni* $h \leq r$*; i restanti coefficienti sulla diagonale di* $\mathbf{L}$ *sono uguali a* 1 *e i restanti coefficienti al di sotto della diagonale sono uguali a* 0.

Dimostrazione. Facciamo induzione su r. Per $r = 1$, la matrice $\mathbf{L} = \mathbf{E}_{i_1 j_1}(\alpha_1)$ ha l'elemento $l_{i_1 j_1}$ coincidente con α_1 e per il resto coincide con la matrice identica, come richiesto. Sia $r > 1$ e l'asserto vero per $s = r - 1$. Allora $\mathbf{L}_s = \mathbf{E}_{i_1 j_1}(\alpha_1)\mathbf{E}_{i_2 j_2}(\alpha_2)\dots\mathbf{E}_{i_s j_s}(\alpha_s)$ soddisfa a quanto richiesto. Eseguendo il prodotto

$$\mathbf{L}_s\mathbf{E}_{i_r j_r}(\alpha_r)$$

si modifica $\mathbf{L}_s$ perché si somma alla sua colonna j_r-esima la colonna i_r-esima moltiplicata per α_r. Ma poiché $j_r < i_r$, ciò equivale a sostituire il coefficiente nullo di posto (i_r, j_r) di $\mathbf{L}_s$ con α_r, da cui l'asserto. □

Esempio 6.12. Applichiamo il Lemma 6.11 al seguente prodotto di matrici elementari 4×4, che si trovano nell'ordine anti-lessicografico crescente richiesto nel lemma:

$$\mathbf{E}_{21}(-3)\mathbf{E}_{41}(2)\mathbf{E}_{22}(7)\mathbf{E}_{32}(5)\mathbf{E}_{42}(-7)\mathbf{E}_{33}(3)\mathbf{E}_{43}(4)\mathbf{E}_{44}(5) =$$

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

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

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

Dal Lemma 6.11 ricaviamo la seguente conseguenza.

PROPOSIZIONE 6.13. *Sia* $\mathbf{A}$ *una matrice per cui la EG non necessita di scambi di righe. La matrice triangolare inferiore invertibile* $\mathbf{L}$ *ottenuta nel Teorema 6.6 ha come coefficiente di posto* (i, j)*:*

(a) *lo scalare* $-\alpha$ *se* $i > j$*, dove* $\mathbf{E}_{ij}(\alpha)$ *è la matrice elementare usata per porre* 0 *al posto* (i, j)*, oppure* 0 *se non si è eseguita alcuna operazione elementare;*
(b) *lo scalare* α^{-1} *se* $i = j$*, dove* $\mathbf{E}_i(\alpha)$ *è la matrice elementare usata per porre* 1 *al posto* (i, i)*, oppure* 1 *se non si è eseguita alcuna operazione elementare.*

DIMOSTRAZIONE. Si osservi che l'ordine con cui si eseguono le operazioni elementari nella EG fa sì che le matrici elementari che pre-moltiplicano $\mathbf{A}$ si susseguono (da sinistra verso destra) in ordine inverso rispetto a quello considerato nel Lemma 6.11. Ma la matrice $\mathbf{E}$ prodotto di tali matrici elementari va invertita (si passa da $\mathbf{EA} = \mathbf{U}$ ad $\mathbf{A} = \mathbf{E}^{-1}\mathbf{U}$). Quindi l'ordine delle matrici elementari in $\mathbf{E}^{-1} = \mathbf{L}$ viene esso pure invertito e coincide con quello del Lemma 6.11, mentre nel passaggio all'inversa i coefficienti α delle matrici $\mathbf{E}_{ij}(\alpha)$ cambiano di segno e quelli delle matrici $\mathbf{E}_i(\alpha)$ vengono invertiti, per la proprietà (I). Il Lemma 6.11 assicura allora l'asserto. $\square$

Il lettore può verificare nell'Esempio 6.10 come la matrice $\mathbf{L}$ nella decomposizione $\mathbf{PA} = \mathbf{LU}$ si ottenga proprio nel modo descritto nella Proposizione 6.13, a partire dalle quattro matrici elementari utilizzate per ottenere tale decomposizione.

La decomposizione $\mathbf{LU}$ del Teorema 6.6 non è simmetrica, nel senso che, mentre i primi elementi non nulli di ogni riga di $\mathbf{U}$ sono uguali a 1, i coefficienti l_{ii} sulla diagonale di $\mathbf{L}$ non lo sono. La cosa si rimedia nel modo seguente: si divide ogni colonna di $\mathbf{L}$ per il suo coefficiente diagonale, ottenendo la nuova matrice $\mathbf{L}_0$ che ha i primi elementi non nulli di ogni colonna uguali a 1; per l'Esempio 1.9 risulta

$$\mathbf{L} = \mathbf{L}_0\mathbf{D}$$

dove $\mathbf{D} = \mathbf{Diag}(l_{11}, \dots, l_{nn})$. Allora la decomposizione $\mathbf{A} = \mathbf{LU}$ diviene $\mathbf{A} = \mathbf{L}_0\mathbf{DU}$ che è simmetrica nel senso sopra descritto.

ESEMPIO 6.14. La decomposizione $\mathbf{LU}$ della matrice $\mathbf{B}$ dell'Esempio 6.7 è la seguente

$$\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -5 & 0 \\ -1 & 4 & 1/5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}.$$

La decomposizione $\mathbf{B} = \mathbf{L}_0\mathbf{D}\mathbf{U}$ sopra descritta è allora la seguente

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -1/2 & -4/5 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1/5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}. \qquad \square$$

Decomposizioni a rango pieno. Nel Paragrafo 3 il numero k di righe non nulle di una forma ridotta $\mathbf{U}$ della matrice $m \times n$ $\mathbf{A}$ è stato chiamato rango di $\mathbf{A}$. Abbiamo inoltre anticipato che tale numero dipende solo da $\mathbf{A}$. Se la EG su $\mathbf{A}$ non richiede scambi di righe, otteniamo, tramite il Teorema 6.6, una decomposizione $\mathbf{A} = \mathbf{L}\mathbf{U}$, dove $\mathbf{U}$ ha $m-k$ righe nulle. Queste righe vengono moltiplicate per le ultime $m-k$ colonne della matrice $\mathbf{L}$; infatti, decomposte $\mathbf{L}$ e $\mathbf{U}$ rispettivamente in blocchi colonna e blocchi riga, si ha:

$$\mathbf{A} = \mathbf{L}\mathbf{U} = [\mathbf{l}_1\ \mathbf{l}_2\ \dots\ \mathbf{l}_m] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \dots \\ \mathbf{u}_m^T \end{bmatrix} =$$
$$= \mathbf{l}_1\mathbf{u}_1^T + \dots + \mathbf{l}_k\mathbf{u}_k^T + \mathbf{l}_{k+1}\mathbf{0}^T + \dots + \mathbf{l}_m\mathbf{0}^T.$$

Quindi $\mathbf{A} = \mathbf{B}\mathbf{C}$, dove

$$\mathbf{B} = [\mathbf{l}_1\ \mathbf{l}_2\ \dots\ \mathbf{l}_k), \qquad \mathbf{C} = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \dots \\ \mathbf{u}_k^T \end{bmatrix}.$$

Se invece la EG su $\mathbf{A}$ richiede scambi di righe, allora si ha la decomposizione $\mathbf{P}\mathbf{A} = \mathbf{L}\mathbf{U}$, con $\mathbf{P}$ opportuna matrice di permutazione, quindi $\mathbf{A} = \mathbf{P}^{-1}\mathbf{L}\mathbf{U}$. Si ha allora, analogamente a prima, una decomposizione $\mathbf{A} = \mathbf{B}\mathbf{C}$, dove ora $\mathbf{B}$ è la sottomatrice di $\mathbf{P}^{-1}\mathbf{L}$ formata dalle prime k colonne. Tali decomposizioni di $\mathbf{A}$ si chiamano *decomposizioni a rango pieno*, e sono caratterizzate dal fatto che il numero di colonne del primo fattore, che coincide col numero di righe del secondo fattore, è uguale al rango di $\mathbf{A}$.

Osserviamo che la decomposizione ottenuta in questo modo mostra che ogni matrice $\mathbf{A}$ soddisfa alle ipotesi della Proposizione 4.19, che assicura l'esistenza della matrice pseudo-inversa: infatti è immediato verificare che $\mathbf{B}$ ha inversa sinistra, essendo formata dalle prime k colonne di una matrice invertibile (vedi l'Esercizio ??), e che $\mathbf{C}$ ha inversa destra, essendo matrice in forma ridotta senza righe nulle.

Esempio 6.15. Data la matrice

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 4 \\ 2 & -2/3 & 5/3 \\ -6 & 2 & -8 \end{bmatrix}$$

la EG, non essendo necessari scambi di righe, produce:

$$\mathbf{E}_2(-1)\mathbf{E}_{31}(6)\mathbf{E}_{21}(-2)\mathbf{E}_1(1/3)\mathbf{A} = \begin{bmatrix} 1 & -1/3 & 4/3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{U}.$$

Quindi si ha la decomposizione

$$\mathbf{A} = \mathbf{L}\mathbf{U} = \mathbf{E}_1(3)\mathbf{E}_{21}(2)\mathbf{E}_{31}(-6)\mathbf{E}_2(-1)\mathbf{U}$$
$$= \begin{bmatrix} 3 & 0 & 0 \\ 2 & -1 & 0 \\ -6 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/3 & 4/3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ne consegue che una decomposizione a rango pieno di $\mathbf{A}$ è data da

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

Esercizi

Paragrafo 1

1.1. Sia $\mathbf{A}$ una matrice $m \times n$ e $\mathbf{D} = \mathbf{Diag}(d, d, \dots, d)$ una matrice scalare $m \times m$. Allora $\mathbf{DA} = d\mathbf{A}$. Analogamente se si moltiplica $\mathbf{A}$ per una matrice scalare.

1.2. Provare quanto asserito nell'Esempio 1.7.

1.3. Si trovino tutte le matrici complesse 2×2 che commutano con tutte le matrici complesse 2×2 triangolari superiori.

1.4. Si trovino due matrici complesse 2×2 non nulle $\mathbf{X}$ e $\mathbf{Y}$ tali che $\mathbf{X}^2 + \mathbf{Y}^2 = \mathbf{O}$.

1.5. Si trovino tutte le matrici reali simmetriche 2×2 $\mathbf{A}$ tali che $\mathbf{A}^2 = \mathbf{I}_2$.

1.6. Si provi che non esiste alcuna matrice complessa 2×2 $\mathbf{A}$ tale che

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

1.7. Si trovino tutte le matrici complesse 2×2 $\mathbf{A}$ tali che $\mathbf{A}^2 = -\mathbf{I}_2$.

1.8. Siano $\mathbf{A}$ una matrice con la i-esima riga nulla e $\mathbf{B}$ una matrice a essa conforme per il prodotto a destra. Si provi che $\mathbf{AB}$ ha la i-esima riga nulla.

1.9. Sia $\mathbf{B}$ una matrice con la i-esima colonna nulla e $\mathbf{A}$ una matrice a essa conforme per il prodotto a sinistra. Si provi che $\mathbf{AB}$ ha la i-esima colonna nulla.

Paragrafo 2

1.10. Si provi che, dato un qualunque vettore complesso non nullo $\mathbf{x}$, $\mathbf{x}^H\mathbf{x}$ è un numero reale positivo.

1.11. Una matrice complessa $n \times n$ $\mathbf{U}$ si dice *unitaria* se $\mathbf{U}^H\mathbf{U} = \mathbf{I}_n = \mathbf{U}\mathbf{U}^H$. Si provi che, se $\mathbf{U}$ è unitaria ed hermitiana, allora la matrice $\mathbf{P} = \frac{1}{2}(\mathbf{I}_n - \mathbf{U})$ soddisfa alle uguaglianze $\mathbf{P} = \mathbf{P}^H$ e $\mathbf{P}^2 = \mathbf{P}$.

1.12. Sia $\mathbf{P}$ una matrice che soddisfa alle uguaglianze $\mathbf{P} = \mathbf{P}^H$ e $\mathbf{P}^2 = \mathbf{P}$. Si provi che la matrice $\mathbf{U} = \mathbf{I}_n - 2\mathbf{P}$ è unitaria ed hermitiana.

1.13. Si verifichino le proprietà (a)–(f) enunciate dopo l'Esempio 2.7.

1.14. Si verifichi che le matrici hermitiane, quelle anti-hermitiane e quelle unitarie sono matrici normali.

1.15. Trovare una matrice normale 2×2 che non è né hermitiana, né anti-hermitiana, né unitaria.

1.16. Quante sono le matrici diagonali $n \times n$ che sono contemporaneamente hermitiane e unitarie?

1.17. Data una matrice $n \times n$ $\mathbf{A} = [a_{ij}]$, si chiama *traccia* di $\mathbf{A}$ lo scalare $\mathrm{Tr}(\mathbf{A}) = \sum_{1 \leq i \leq n} a_{ii}$, cioè la somma dei suoi coefficienti diagonali. Si provi che, se $\mathbf{A}$ e $\mathbf{B}$ sono matrici conformi per la somma e α è uno scalare, si ha $\mathrm{Tr}(\mathbf{A}+\mathbf{B}) = \mathrm{Tr}(\mathbf{A}) + \mathrm{Tr}(\mathbf{B})$ e $\mathrm{Tr}(\alpha\mathbf{A}) = \alpha \,\mathrm{Tr}(\mathbf{A}$; se invece $\mathbf{A}$ e $\mathbf{B}$ sono matrici conformi per il prodotto, si ha $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$.

Paragrafo 3

1.18. Sia $\mathbf{A}$ una matrice $m \times n$ di rango 1. Si provi che $\mathbf{A} = \mathbf{v}\mathbf{u}^T$, dove $\mathbf{v} \in \mathbb{C}^m$ e $\mathbf{u}^T \in \mathbb{C}_n$.

1.19. Si spieghi perché un sistema lineare con una matrice dei coefficienti $\mathbf{A}$ $m \times n$ con $m < n$ non può avere esattamente una soluzione.

1.20. Una matrice $\mathbf{A}$ si dice *equivalente per righe* alla matrice $\mathbf{B}$, e si scrive $\mathbf{A} \sim \mathbf{B}$, se $\mathbf{B}$ si ottiene da $\mathbf{A}$ con un numero finito di operazioni elementari. Si provi che
- (i) $\mathbf{A} \sim \mathbf{A}$;
- (ii) se $\mathbf{A} \sim \mathbf{B}$, allora $\mathbf{B} \sim \mathbf{A}$;
- (iii) se $\mathbf{A} \sim \mathbf{B}$ e $\mathbf{B} \sim \mathbf{C}$, allora $\mathbf{A} \sim \mathbf{C}$.

1.21. Siano $\mathbf{A}$ una matrice $m \times n$ e $\mathbf{B}$ una matrice $n \times p$. Si considerino i due sistemi lineari
- (i) $\mathbf{ABx} = \mathbf{Av}$
- (ii) $\mathbf{Bx} = \mathbf{v}$.

Si provi che: ogni soluzione di (ii) è anche soluzione di (i) e che $\mathbf{x}_0$ è soluzione di (i) se e solo se $\mathbf{B}\mathbf{x}_0 = \mathbf{v} + \mathbf{n}$, dove $\mathbf{An} = \mathbf{0}$.

Paragrafo 4

1.22. Si verifichino le proprietà (a)–(c) delle matrici invertibili enunciate prima dell'Esempio 4.16.

1.23. Dimostrare quanto enunciato nell'Esempio 4.16 sulle matrici triangolari invertibili (si faccia induzione sull'ordine).

1.24. Si provi direttamente che, se la matrice $\mathbf{R}$ è una inversa destra della matrice $\mathbf{A}$ tale che $\mathbf{RA}$ è hermitiana, allora $\mathbf{R} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$ (utilizzare il prodotto $\mathbf{R}\mathbf{A}\mathbf{A}^H$).

1.25. Sia $\mathbf{B}$ una sottomatrice di una matrice invertibile $n \times n$ $\mathbf{A}$, formata dalle prime k colonne di $\mathbf{A}$. Si provi che $\mathbf{B}$ ha inversa sinistra.

1.26. Si provi che una matrice $\mathbf{A}$ che ha una inversa destra simmetrica $\mathbf{R}$ è invertibile (provare che $\mathbf{R}\mathbf{A}^T = \mathbf{I}$).

1.27. Si provi direttamente che una matrice quadrata $\mathbf{A}$ che ha una unica inversa destra $\mathbf{R}$ è invertibile (si consideri la matrice $\mathbf{RA} - \mathbf{I} + \mathbf{R}$).

1.28. Si trovino una matrice 2×2 $\mathbf{A}$ e una sua fattorizzazione $\mathbf{A} = \mathbf{BC}$ tali che $\mathbf{A}^+ \neq \mathbf{C}^+\mathbf{B}^+$.

1.29. Si verifichino le proprietà (pi_1)–(pi_6) della matrice pseudo-inversa.

1.30. Sia $\mathbf{A}$ una matrice $m \times n$ di rango n. Si provi che $\mathbf{A}^+ = \mathbf{A}^H$ se e solo se $\mathbf{A}\mathbf{A}^H$ è una matrice idempotente, cioè coincide col suo quadrato.

1.31. Sia $\mathbf{A}$ una matrice di rango 1. Si provi che $\mathbf{A}^+ = (\mathbf{u}^H\mathbf{u})^{-1}(\mathbf{v}^H\mathbf{v})^{-1}\mathbf{A}^H$ per opportuni vettori $\mathbf{u}$ e $\mathbf{v}$.

1.32. Si provi che le due matrici $\mathbf{A}\mathbf{A}^+$ e $\mathbf{A}^+\mathbf{A}$ sono idempotenti e che $(\mathbf{A}\mathbf{A}^H)^+ = (\mathbf{A}^+)^H\mathbf{A}^+$.

Paragrafo 5

1.33. Trovare tutte le inverse destre $\mathbf{R}$ di una matrice $m \times n$ $\mathbf{A}$ di rango $m < n$, post-moltiplicando $\mathbf{A}$ per una matrice di permutazione $\mathbf{P}$ in modo che $\mathbf{AP}$ soddisfi alle ipotesi della Proposizione 5.3 e ricavando $\mathbf{R}$ dalle inverse destre di $\mathbf{AP}$ (si veda il Paragrafo 6 per le matrici di permutazione).

1.34. Calcolare con l'algoritmo di inversione di Gauss-Jordan la inversa della matrice

$$\mathbf{A} = \begin{bmatrix} a & 1 & 0 \\ 0 & b & 1 \\ 0 & 0 & c \end{bmatrix}$$

dove a, b e c sono scalari non nulli.

Paragrafo 6

1.35. Una matrice in forma ridotta $\mathbf{S}$ di rango k si dice *in forma a scala per righe ridotta* se le k colonne dominanti coincidono con i vettori cooordinati $\mathbf{e}_1, \dots, \mathbf{e}_k$. Si provi che ogni matrice $\mathbf{A}$ si trasforma con operazioni elementari in una matrice in forma a scala per righe ridotta $\mathbf{S}$ e che $\mathbf{S} = \mathbf{EA}$, dove $\mathbf{E}$ è una matrice invertibile.

1.36. Siano $\mathbf{S}$ ed $\mathbf{S}'$ due forme a scala per righe ridotte della stessa matrice $\mathbf{A}$. Si provi che $\mathbf{S} = \mathbf{S}'$ (si provi che $\mathbf{VS} = \mathbf{S}'$, dove $\mathbf{V}$ è una matrice invertibile, e si sfrutti la forma speciale di $\mathbf{S}$ ed $\mathbf{S}'$ per provare che $\mathbf{V}\mathbf{e}_i = \mathbf{e}_i$ per ogni i).

1.37. Una matrice $\mathbf{A}$ è invertibile se e solo se è prodotto di matrici elementari.

1.38. Data una matrice $m \times n$ $\mathbf{A}$ di rango k, esistono matrici invertibili $\mathbf{U}$ e $\mathbf{V}$ di dimensioni rispettivamente m ed n, tali che

$$\mathbf{UAV} = \begin{bmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

1.39. Si verifichi che $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij} = \mathbf{E}_{ij}^T$ e $\mathbf{E}_i(\alpha)^{-1} = \mathbf{E}_i(\alpha^{-1}) = \mathbf{E}_i(\alpha)^T$ $(\alpha \neq 0)$.

1.40. Provare la Proposizione 6.5 utilizzando la Proposizione 6.3 e passando alle matrici trasposte.

1.41. Si provi che, se $\mathbf{P}$ è una matrice di permutazione, allora $\mathbf{P}^{-1} = \mathbf{P}^T$.

1.42. La permutazione π dell'insieme $\{1, 2, \dots, n\}$ che manda 1 in i_1, 2 in i_2, ..., n in i_n, si denota usualmente col simbolo

$$\pi = \begin{bmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{bmatrix}.$$

Sia P la matrice di permutazione ottenuta dalla matrice identica I_n permutandone le righe tramite la permutazione π. Si provi che P^T si ottiene da I_n permutandone le colonne con la permutazione π (si scriva P come prodotto di matrici di trasposizione).

1.43. Sia $\mathbf{A}$ una matrice $n \times n$ e sia $\mathbf{P}$ una matrice di permutazione dello stesso ordine ottenuta dalla matrice identica permutandone le righe con la permutazione π. Si provi che $\mathbf{PA}$ si ottiene da $\mathbf{A}$ permutandone le righe con la permutazione π, mentre $\mathbf{AP}^T$ si ottiene da $\mathbf{A}$ permutandone le colonne con la permutazione π.

1.44. Nelle notazioni dell'Esercizio 1.43, si provi che la matrice $\mathbf{B} = \mathbf{PAP}^T$ (che si dice *cogrediente* alla matrice $\mathbf{A}$) si ottiene da $\mathbf{A}$ permutandone le righe e le colonne con la permutazione π. Si provi inoltre che la diagonale di $\mathbf{B}$ si ottiene da quella di $\mathbf{A}$ tramite la medesima permutazione.

1.45. Si provi che, se $s < i < j$, risulta $\mathbf{E}_{ij}\mathbf{E}_{is}(\alpha) = \mathbf{E}_{js}(\alpha)\mathbf{E}_{ij}$, $\mathbf{E}_{ij}\mathbf{E}_{js}(\alpha) = \mathbf{E}_{is}(\alpha)\mathbf{E}_{ij}$, e se $r \neq i, j$, allora $\mathbf{E}_{ij}\mathbf{E}_{rs}(\alpha) = \mathbf{E}_{rs}(\alpha)\mathbf{E}_{ij}$.