Il metodo di Gauss per la risoluzione dei sistemi lineari

A. De Paris

Questi appunti sono suddivisi in tre parti: “discussione”, “teoria” e “esempi ed esercizi”. La parte che va imparata e ripassata (per capirci, quella che “serve per l’esame orale”) è chiaramente la seconda. La “discussione” serve per introdurre e per far capire le motivazioni, dunque va letta e capita bene prima di studiare la teoria perché è un aiuto per poterla imparare con meno sforzo. Naturalmente però la discussione non va ripassata per l’esame. Gli “esercizi ed esempi” sono indispensabili per fissare le nozioni apprese a teoria e sono utili in vista dell’esame scritto.

Discussione

Alle scuole superiori si imparano vari metodi per risolvere i sistemi lineari. Probabilmente il più popolare è il metodo “per sostituzione”. Tuttavia, è esperienza di molti che quando si applica il metodo per sostituzione a sistemi un po’ più complicati del normale, si commettono facilmente degli errori.
Noi siamo qui interessati ad imparare un metodo generale, per risolvere sistemi lineari con un qualsiasi numero di equazioni e con un qualsiasi numero di incognite. L’esperienza ha dimostrato che il metodo per sostituzione può essere migliorato. Il metodo più usato oggi è il metodo di Gauss, che ci proponiamo di esporre in questi appunti. Consideriamo il sistema lineare

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

Chi è appassionato del metodo per sostituzione, riuscirà facilmente a trovare la soluzione

$$x = 1, \qquad y = 3.$$

Proviamo però qui ad usare un altro metodo, che forse qualcuno avrà già incontrato con il nome di “metodo di addizione e sottrazione”. Moltiplichiamo per tre entrambi i membri della seconda equazione, e sottraiamo ordinatamente i due membri della prima. Otteniamo $-5y + 15 = 0$, da cui $y = 3$. Moltiplicando poi per due nella seconda e sommando alla prima otteniamo $5x - 5 = 0$, da cui $x = 1$.
Confrontando i due metodi, dovrebbero essere evidenti due cose:

- Sotto sotto, i conti fatti sono quasi gli stessi.
- Il secondo metodo (una volta capito) è più semplice da svolgere.

Si può pensare anche ad un metodo misto: dopo aver ricavato $y = 3$ con "addizione e sottrazione" si può passare alla sostituzione di tale valore in una delle due equazioni. Questi passaggi possono essere riassunti come segue:

$$\begin{cases} 3x + 2y - 9 = 0 \\ x - y + 2 = 0 \end{cases} \overset{e_2 \to 3e_2 - e_1}{\Longleftrightarrow} \begin{cases} 3x + 2y - 9 = 0 \\ -5y + 15 = 0 \end{cases} \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} 3x + 2 \cdot 3 - 9 = 0 \\ y = 3 \end{cases} \Longleftrightarrow \begin{cases} x = 1 \\ y = 3 \end{cases} .$$

Al primo passaggio abbiamo scritto $e_2 \to 3e_2 - e_1$ per significare: moltiplichiamo per tre entrambi i membri della seconda equazione e sottraiamo ordinatamente i membri della prima, e mettiamo l'equazione che risulta al posto della seconda. Se si dimostra (come è facile fare) che i passaggi sono dei "se e solo se" questo ci assicura che $(x, y) = (1, 3)$ è l'unica soluzione del sistema.

Facciamo un altro esempio di uso di questo metodo "misto", provandolo su un sistema un po' più grande:

$$\begin{cases} 2x + 3y + 5z + 3 = 0 \\ x - y + z = 0 \\ -x + 7y - 2z - 1 = 0 \end{cases} .$$

Cominciamo con le "addizioni e sottrazioni", in modo da semplificare:

$$\begin{matrix} e_2 \longrightarrow 2e_2 - e_1 \\ e_3 \longrightarrow 2e_3 + e_1 \end{matrix} \Longrightarrow \begin{cases} 2x + 3y + 5z + 3 = 0 \\ -5y - 3z - 3 = 0 \\ 17y + z + 1 = 0 \end{cases} ;$$

la x è rimasta solo nella prima equazione, e le altre due equazioni costituiscono un sistema più piccolo, al quale possiamo riapplicare il metodo:

$$e_3 \longrightarrow 5e_3 + 17e_2 \Longrightarrow \begin{cases} 2x + 3y + 5z + 3 = 0 \\ -5y - 3z - 3 = 0 \\ -46z - 46 = 0 \end{cases} .$$

E infine, le sostituzioni "a ritroso" (cioè partendo dall'ultima equazione e procedendo dal basso verso l'alto):

$$\begin{cases} 2x + 3 \cdot 0 + 5 \cdot (-1) + 3 = 0 \longrightarrow x = 1 \\ -5y - 3 \cdot (-1) - 3 = 0 \longrightarrow y = 0 \\ z = -1 \end{cases} .$$

Riassumendo:

$$\begin{cases} 2x + 3y + 5z + 3 = 0 \\ x - y + z = 0 \\ -x + 7y - 2z - 1 = 0 \end{cases} \Longleftrightarrow \begin{cases} 2x + 3y + 5z + 3 = 0 \\ -5y - 3z - 3 = 0 \\ 17y + z + 1 = 0 \end{cases} \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} 2x + 3y + 5z + 3 = 0 \\ -5y - 3z - 3 = 0 \\ -46z - 46 = 0 \end{cases} \Longleftrightarrow \begin{cases} x = 1 \\ y = 0 \\ z = -1 \end{cases} .$$

Quindi $(x, y, z) = (1, 0, -1)$ è l'unica soluzione del sistema. Il metodo sembra dunque funzionare abbastanza rapidamente. Inoltre, chi vuole ulteriormente risparmiare fatica (e inchiostro) può notare che quello che conta nella scelta dei calcoli da eseguire nella prima parte (quella delle addizioni e sottrazioni) sono i coefficienti e i termini noti. Quindi, stando attenti a scrivere in ordine per righe e colonne, si può evitare di scrivere le incognite. In sostanza, la prima parte consiste dunque in operazioni sulle matrici complete associate ai sistemi. Quindi, per il sistema di sopra, basterebbe scrivere

$$\begin{pmatrix} 2 & 3 & 5 & 3 \\ 1 & -1 & 1 & 0 \\ -1 & 7 & -2 & -1 \end{pmatrix} \overset{\substack{e_2 \to 2e_2 - e_1 \\ e_3 \to 2e_3 + e_1}}{\longleftrightarrow} \begin{pmatrix} 2 & 3 & 5 & 3 \\ 0 & -5 & -3 & -3 \\ 0 & 17 & 1 & 1 \end{pmatrix} \overset{e_3 \to 5e_3 + 17e_1}{\longleftrightarrow}$$

$$\longleftrightarrow \begin{pmatrix} 2 & 3 & 5 & 3 \\ 0 & -5 & -3 & -3 \\ 0 & 0 & -46 & -46 \end{pmatrix},$$

per poi procedere alle sostituzioni a ritroso nel sistema rappresentato dall'ultima matrice.

Vediamo cosa succede per un sistema che ha più di una soluzione:

$$\begin{cases} x + y + z + 3 = 0 \\ 2x - y + 3z + 4 = 0 \\ 3x + 4z + 7 = 0 \end{cases} .$$

Avviamo il metodo:

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & -1 & 3 & 4 \\ 3 & 0 & 4 & 7 \end{pmatrix} \overset{\substack{e_2 \to e_2 - 2e_1 \\ e_3 \to e_3 - 3e_1}}{\longleftrightarrow} \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & -3 & 1 & -2 \\ 0 & -3 & 1 & -2 \end{pmatrix}$$

$$\overset{e_3 \to e_3 - e_2}{\longleftrightarrow} \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & -3 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

L'ultima equazione è diventata l'identità $0 = 0$, che può essere eliminata dal sistema senza avere alcun effetto sul calcolo delle soluzioni. Otteniamo allora

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

Ora siamo in una situazione un po' diversa da quelle precedenti: non è possibile ricavare alcuna incognita per far partire le sostituzioni. Notiamo però che se assegnamo un valore a piacere a z, possiamo procedere poi come al solito. Per esempio, se $z = -4$

$$\begin{cases} x - 2 - 4 + 3 = 0 \longrightarrow x = 3 \\ -3y - 4 - 2 = 0 \longrightarrow y = -2 \end{cases}$$

e così troviamo la soluzione $(3, -2, -4)$, che però non è l'unica possibile (possiamo solo dire che è l'unica con la $z = -4$). Se scegliamo $z = -1$ otteniamo la soluzione $(-1, -1, -1)$. Capiamo dunque che ci sono infinite soluzioni. Per descriverle tutte, possiamo scrivere (sempre partendo dal basso):

$$\text{se } z = t, \text{ allora } \begin{cases} x + \frac{1}{3}t - \frac{2}{3} + t + 3 = 0 \longrightarrow x = -\frac{4}{3}t - \frac{7}{3} \\ -3y + t - 2 = 0 \longrightarrow y = \frac{1}{3}t - \frac{2}{3} \end{cases}$$

Quindi l'insieme di tutte le possibili soluzioni è

$$\left\{ \left(-\frac{4}{3}t - \frac{7}{3} \ , \ \frac{1}{3}t - \frac{2}{3} \ , \ t \ \right) : t \in \mathbb{R} \right\},$$

cioè al variare di t in $\mathbb{R}$ si ottengono tutte le terne che soddisfano il sistema:

$$\left\{ \ldots, \left(-\frac{7}{3}, -\frac{2}{3}, 0\right), \ldots, \left(-\frac{11}{3}, -\frac{1}{3}, 1\right), \ldots, \left(-\frac{4\sqrt{2}+7}{3}, \frac{\sqrt{2}-2}{3}, \sqrt{2} \right), \ldots \right\}.$$

Vediamo ora un esempio di sistema incompatibile (cioè senza soluzioni).

$$\begin{cases} 2x + y - z - 5 = 0 \\ 4x - 3y - 2z - 3 = 0 \\ 8x - y - 4z = 0 \end{cases}.$$

Risolviamo:

$$\begin{pmatrix} 2 & 1 & -1 & -5 \\ 4 & -3 & -2 & -3 \\ 8 & -1 & -4 & 0 \end{pmatrix} \overset{\substack{e_2 \to e_2 - 2e_1 \\ e_3 \to e_3 - 4e_1}}{\longleftrightarrow} \begin{pmatrix} 2 & 1 & -1 & -5 \\ 0 & -5 & 0 & 7 \\ 0 & -5 & 0 & 20 \end{pmatrix} \overset{e_3 \to e_3 - e_2}{\longleftrightarrow}$$

$$\begin{pmatrix} 2 & 1 & -1 & -5 \\ 0 & -5 & 0 & 7 \\ 0 & 0 & 0 & 13 \end{pmatrix}.$$

Abbiamo quindi il sistema

$$\begin{cases} 2x + y - z + 5 = 0 \\ -5y + 7 = 0 \\ 13 = 0 \end{cases},$$

che chiaramente non può avere soluzioni, perché l'ultima equazione non può essere soddisfatta.

Vediamo ancora un altro esempio:

$$\begin{cases} -2x + 3y + z - 2 = 0 \\ 4x - 6y - z + 3 = 0 \\ x + y + z - 3 = 0 \end{cases}.$$

Avviamo il metodo:

$$\begin{pmatrix} -2 & 3 & 1 & -2 \\ 4 & -6 & -1 & 3 \\ 1 & 1 & 1 & -3 \end{pmatrix} \overset{\substack{e_2 \to e_2 + 2e_1 \\ e_3 \to 2e_3 + e_1}}{\longleftrightarrow} \begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & 0 & 1 & -1 \\ 0 & 5 & 3 & -8 \end{pmatrix} \overset{?}{\longleftrightarrow} \dots$$

a questo punto abbiamo un dubbio su come continuare: la y è sparita dalla seconda equazione, mentre per poter fare le sostituzioni a ritroso come siamo abituati, vorremmo che nella terza ci fosse solo z, nella seconda y e z (anche se poi non è importante che ci sia z), e nella prima tutti e tre (anche se poi quella veramente importante dovrebbe essere solo la x). Inoltre, notiamo che nella terza la y c'è. Questo è un vantaggio? E poi, se la y mancasse anche nella terza, come si dovrebbe procedere? Vediamo quindi che ci sono ancora dei problemi che si possono presentare, e che il metodo che abbiamo introdotto ancora non ci dà una "regoletta meccanica" valida per tutti i sistemi (anche se, rispetto alle scuole superiori, abbiamo fatto dei bei passi avanti). Inoltre, teniamo presente che non tutti i sistemi sono di tre equazioni in x, y, z: ci sono sistemi con più incognite che equazioni, e sistemi con più equazioni che incognite. Capiamo che il metodo introdotto può aiutare anche in questi casi, ma capiamo altresì che ci sono ancora alcuni dettagli da mettere a punto. Forse la cosa più importante ancora da capire, specialmente per sistemi non "quadrati", è quale deve essere in generale il punto d'arrivo dei passaggi di addizione e sottrazione. Infine, bisogna stare ancora attenti a non procedere troppo meccanicamente, perché l'errore può essere in agguato, come mostra il seguente esempio:

$$\begin{cases} x+y+z+w=0 \\ 2x-y+3z+w-5=0 \\ 2x-y-z+2w=0 \\ \sqrt{2}x-\pi y+4,57z-88w-157121=0 \end{cases}.$$

Qualcuno, per evitare la complicazione data dalla quarta equazione, potrebbe essere tentato di procedere così:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 2 & -1 & 3 & 1 & -5 \\ 2 & -1 & -1 & 2 & 0 \\ \sqrt{2} & -\pi & 4,57 & -88 & -157121 \end{pmatrix} \overset{\substack{e_2\to e_2-2e_1 \\ e_3\to e_3-2e_1 \\ e_4\to e_3-e_2}}{\leftrightsquigarrow}$$

$$\leftrightsquigarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & -1 & -5 \\ 0 & -3 & -3 & 0 & 0 \\ 0 & 0 & -4 & 1 & 5 \end{pmatrix} \xleftrightarrow{e_3\to e_3-e_2} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & -1 & -5 \\ 0 & 0 & -4 & 1 & 5 \\ 0 & 0 & -4 & 1 & 5 \end{pmatrix} \longleftrightarrow$$

$$\xleftrightarrow{e_4\to e_4-e_3} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & -1 & -5 \\ 0 & 0 & -4 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nel sistema corrispondente all'ultima matrice possiamo porre $w = t$ e ricavare con le sostituzioni le (infinite) soluzioni. Troppo comodo! Notiamo infatti che ad esempio la soluzione $(1, -1, 1, -1)$ (corrispondente alla scelta $w = -1$) non soddisfa il sistema di partenza; c'è quindi qualche errore. Con un po' di buon senso ci si accorge che l'errore deve essere nel passaggio in cui abbiamo liquidato con troppa facilità l'equazione brutta [1]. E infatti, quel passaggio (il primo) non è un "se e solo se", in quanto è chiaro che ogni soluzione del primo sistema è anche soluzione del secondo, ma non c'è motivo per cui ogni soluzione del secondo debba essere soluzione del primo, e di fatto non è così, come mostra l'esempio $(1, -1, 1, -1)$. Se si vuole dunque stabilire una "regola meccanica" per risolvere i sistemi, bisogna capire bene quali sono i passaggi corretti (per arrivarci, potrebbe essere utile cominciare a chiedersi quale differenza c'è tra il passaggio sbagliato che abbiamo appena visto e i passaggi giusti fatti precedentemente).

Facendo il punto della situazione, possiamo dire di aver delineato un metodo generale molto conveniente per risolvere qualsiasi sistema lineare, ma non siamo ancora riusciti a mettere a punto con chiarezza tutti i dettagli. Come abbiamo messo in luce poc'anzi, dobbiamo ancora rispondere con precisione ad alcune domande in sospeso, come ad esempio:

- Come si procede quando un'incognita "sparisce" inaspettatamente?
- Qual è in generale la forma finale da raggiungere con i passaggi di addizione e sottrazione, in modo da far partire in modo efficace le sostituzioni?
- Quali sono i passaggi corretti (per i quali vale il "se e solo se")?

Per fortuna, possiamo risparmiarci la fatica di pensare a tutti questi problemi: questo è già stato fatto da altri, e abbiamo a disposizione una teoria bella e

[1] Nel presente caso basta il buon senso per capire il passaggio sbagliato, altre volte potrebbe essere più nascosto: approfittiamo quindi di questa nota per dare un consiglio su come trovare gli errori, anche in ambiti diversi da quello dei sistemi lineari (può essere utile nello svolgimento degli esami scritti).
Se non comporta molto tempo, è buona regola fare sempre qualche verifica sui risultati ottenuti; e se ci si accorge che un risultato ottenuto con una serie di passaggi è sbagliato, si provi a sostituire tale risultato sbagliato nei passaggi intermedi. Quando succede che tale risultato va bene ad un certo punto, ma non nel punto precedente, allora c'è per forza un errore nel passaggio tra questi due punti.

pronta che ci fornisce la voluta “regola meccanica” (il metodo di Gauss) per risolvere molto efficacemente qualsiasi sistema lineare.

Per imparare ad usare questo metodo (soprattutto nei casi speciali che abbiamo lasciato in sospeso in questa discussione) è utile dunque studiare ora la parte di teoria, allenandosi poi ad applicarla agli esempi proposti e ad altri ancora che si possono inventare o trovare sui libri di testo.

Prima però di partire con la teoria, vogliamo qui dire qualcosa di informale riguardo la seconda domanda della lista riportata poco più sopra, cioè quella riguardante la “forma finale” da raggiungere con i passaggi di addizione e sottrazione. Si capisce bene che lo scopo principale è di far sparire molte incognite in molte equazioni, il che equivale a far comparire molti zeri nelle matrici. Tuttavia gli zeri devono essere disposti in una certa maniera, perché come abbiamo visto, se presenti in certe posizioni, possono creare problemi. Riportiamo qui sotto le matrici finali dei vari sistemi risolti finora, e cerchiamo di individuare la logica con cui devono essere disposti gli zeri:

$$\begin{pmatrix} 3 & 2 & -9 \\ 0 & -5 & 15 \end{pmatrix}; \qquad \begin{pmatrix} 2 & 3 & 5 & 3 \\ 0 & -5 & -3 & -3 \\ 0 & 0 & -46 & -46 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & -3 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \qquad \begin{pmatrix} 2 & 1 & -1 & -5 \\ 0 & -5 & 0 & 7 \\ 0 & 0 & 0 & 13 \end{pmatrix}.$$

Abbiamo evidenziato che gli zeri “utili” possono essere disposti sotto una figura “a scalini” (anche se possono essere presenti zeri anche al di fuori di tale figura; per esempio, nell’ultima matrice, c’è uno zero fuori dalla figura, che pur facendo comodo, non è strettamente necessario per poter avviare le sostituzioni a ritroso). Questo ci dà l’indicazione per capire la definizione, che sarà data nella parte teorica, di *matrice a scalini*. Naturalmente, anche qui ci sono alcune cose da precisare: per esempio, la scala deve partire in alto a sinistra e scendere terminando nel lato in basso a destra; gli scalini devono essere tutti “alti uno”, ma possono essere di varie lunghezze; gli elementi “sotto la scala” devono essere uguali a zero; gli elementi negli “angoletti degli scalini” (detti *pivot*) devono essere per forza diversi da zero (mentre negli altri posti al di sopra della scala possono esserci degli zeri).

Questo modo di descrivere è abbastanza facile da comprendere, ma non è molto preciso: si corre il rischio di fraintendere. Per esempio, la seguente matrice è a scalini?

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

La risposta è sì:

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

Una maniera più precisa e concisa per dare la definizione di matrice a scalini è la seguente:

- si definisce *pivot* di una riga il suo primo elemento non nullo;
- si definisce una matrice *a scalini* se ha la proprietà che i pivot delle righe non nulle sono sempre più a sinistra dei pivot delle righe non nulle ad esse successive, e che inoltre le eventuali righe nulle sono tutte alla fine.

Naturalmente anche questa definizione non è una definizione formale (i termini "a sinistra", "righe successive", "primo elemento", fanno riferimento a come una matrice viene usualmente rappresentata, e non alla sua definizione formale come applicazione $\{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{R}$). Nella parte teorica daremo una definizione più rigorosa (anche se meno espressiva delle precedenti), che non sarà altro che una traduzione in termini formali della definizione appena data qui.

A questo punto possiamo finalmente partire con la sezione "teoria". Ribadiamo di nuovo che quanto detto fin qui non dovrà essere ripassato per l'esame (sarebbe molto dannoso), ma, se è stato letto con un po' di attenzione, aiuterà senz'altro a studiare e capire la teoria molto più rapidamente e profondamente, e renderà più agevole il ripasso della stessa, questo sì necessario per l'esame.

Teoria

Il metodo di Gauss consiste essenzialmente in una serie di trasformazioni sulla matrice del sistema; cominciamo quindi a definire quali sono le trasformazioni utilizzate. Formalmente, una trasformazione su matrici di tipo $[m, n]$

è un'operazione 1-aria definita nell'insieme $\mathcal{M}_{mn}(\mathbb{R})$, il che significa semplicemente che ad ogni matrice $A \in \mathcal{M}_{mn}(K)$ la nostra trasformazione associa un'altra matrice $B \in \mathcal{M}_{mn}(K)$; dunque le nostre trasformazioni non sono altro che applicazioni $\tau : \mathcal{M}_{mn}(K) \to \mathcal{M}_{mn}(K)$. Diamo dunque la definizione "ufficiale":

Definizione 1 *Una* trasformazione elementare *(detta anche* operazione elementare*) per righe su matrici di tipo $[m, n]$ è un'applicazione $\tau : \mathcal{M}_{mn}(\mathbb{R}) \to \mathcal{M}_{mn}(K)$ che soddisfa una almeno delle seguenti condizioni.*
(I) Esiste un $\bar{i} \in \{1, \ldots, m\}$ e un $\lambda \in K$ <u>non nullo</u> tali che per ogni $A \in \mathcal{M}_{mn}(K)$, posto $B = \tau(A)$, si ha $\underline{b}_{\bar{i}} = \lambda \underline{a}_{\bar{i}}$ e $\underline{b}_i = \underline{a}_i$ $\forall i \neq \bar{i}$. In tal caso la trasformazione si dice di primo tipo.
(II) Esistono $i', i'' \in \{1, \ldots, m\}$ tali che per ogni $A \in \mathcal{M}_{mn}(K)$, posto $B = \tau(A)$, si ha $\underline{b}_{i'} = \underline{a}_{i''}$, $\underline{b}_{i''} = \underline{a}_{i'}$ e $\underline{b}_i = \underline{a}_i \forall i \notin \{i', i''\}$. In tal caso la trasformazione si dice di secondo tipo.
(III) Esistono $\bar{i}, i' \in \{1, \ldots, m\}$ <u>distinti</u> ed esiste $\lambda \in K$ tali che per ogni $A \in \mathcal{M}_{mn}(K)$, posto $B = \tau(A)$, si ha $\underline{b}_{\bar{i}} = \underline{a}_{\bar{i}} + \lambda \underline{a}_{i'}$ e $\underline{b}_i = \underline{a}_i$ $\forall i \neq \bar{i}$. In tal caso la trasformazione si dice di terzo tipo.
In tale situazione diremo anche che B è ottenuta da A tramite la trasformazione elementare τ.

La definizione data sopra può essere riassunta in maniera meno precisa, ma più espressiva, dicendo che le trasformazioni del primo tipo consistono nel moltiplicare una riga per uno scalare non nullo (ed è importante che sia non nullo, perciò lo abbiamo sottolineato nella definizione), quelle del secondo tipo consistono nello scambiare di posto due righe, quelle di terzo tipo consistono nel sommare ad una riga un multiplo di un'altra riga (ed è importante che questa riga sia davvero "un'altra", perciò nella definizione abbiamo sottolineato che i due indici devono essere distinti). Inoltre le tre operazioni elementari possono essere indicate efficacemente come segue:

$$\underline{a}_{\bar{i}} \to \lambda \underline{a}_{\bar{i}} \text{ (tipo I)}, \qquad \underline{a}_{i'} \longleftrightarrow \underline{a}_{i''} \text{ (tipo II)}, \qquad \underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} + \lambda \underline{a}_{i'} \text{ (tipo III)}.$$

Proposizione 1 *L'inversa di una trasformazione elementare è ancora una trasformazione elementare dello stesso tipo.*

Dimostrazione. L'inversa di $\underline{a}_{\bar{i}} \to \lambda \underline{a}_{\bar{i}}$ è $\underline{a}_{\bar{i}} \to \frac{1}{\lambda} \underline{a}_{\bar{i}}$, l'inversa di $\underline{a}_{i'} \longleftrightarrow \underline{a}_{i''}$ è sé stessa, l'inversa di $\underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} + \lambda \underline{a}_{i'}$ è $\underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} - \lambda \underline{a}_{i'}$. C.V.D.

Le affermazioni fatte in questa dimostrazione sono abbastanza evidenti, quindi le abbiamo ritenute sufficienti. Tuttavia, qualcuno potrebbe chiedere "e perché l'inversa di $\underline{a}_{\bar{i}} \rightarrow \lambda \underline{a}_{\bar{i}}$ è $\underline{a}_{\bar{i}} \rightarrow \frac{1}{\lambda}\underline{a}_{\bar{i}}$, l'inversa di $\tau = \underline{a}_{i'} \longleftrightarrow \underline{a}_{i''}$ è sé stessa, ecc. ecc.?". La tentazione sarebbe quella di rispondere "tu sei troppo pignolo!". Tuttavia, riflettiamo un attimo. Abbiamo sottolineato che lo scalare coinvolto nella definizione delle trasformazioni del primo tipo deve essere non nullo, e infatti senza questa condizione non si potrebbe considerare $\underline{a}_{\bar{i}} \rightarrow \frac{1}{\lambda}\underline{a}_{\bar{i}}$. Nella definizione del terzo tipo abbiamo sottolineato che $\bar{i}$ deve essere diverso da i': questa condizione è necessaria per poter avere sempre l'inversa? La risposta è sì, ma bisogna ammettere che a prima vista può sfuggire questo fatto. Capiamo allora che è bene avere sempre la risposta pronta alle domande dei pignoli, allenandosi talvolta a giustificare in maniera precisa anche le cose che possono sembrare più evidenti (questa però non deve diventare una mania, che sarebbe estremamente dannosa). Per questi motivi, proponiamo il seguente esercizio.

Esercizio 1 *Giustificare in dettaglio le affermazioni fatte nella dimostrazione precedente, sottolineando in quale punto interviene la condizione $\bar{i} \neq i'$ per le trasformazioni di terzo tipo.*

Naturalmente le soluzioni dell'esercizio possono essere tante (ciascuno può svolgerlo secondo il suo stile personale). Qui sotto ne proponiamo una.

Soluzione

Tipo I

:

L'inversa di $\tau = \underline{a}_{\bar{i}} \rightarrow \lambda \underline{a}_{\bar{i}}$ è $\sigma = \underline{a}_{\bar{i}} \rightarrow \frac{1}{\lambda}\underline{a}_{\bar{i}}$. Infatti, sia A una qualsiasi matrice di tipo $[m, n]$, sia B la matrice ottenuta da A tramite τ e C la matrice ottenuta da B tramite σ. Abbiamo:

$$\left.\begin{array}{cccc} \text{definizione di } \tau & \text{definizione di } \sigma & & \\ \Downarrow & \Downarrow & & \\ \underline{b}_{\bar{i}} = \lambda \underline{a}_{\bar{i}} & \underline{c}_{\bar{i}} = \frac{1}{\lambda}\underline{b}_{\bar{i}} & \Longrightarrow & \underline{c}_{\bar{i}} = \frac{1}{\lambda}\underline{b}_{\bar{i}} = \frac{1}{\lambda}(\lambda \underline{a}_{\bar{i}}) = \underline{a}_{\bar{i}} \\ \text{e} & \text{e} & & \text{e} \\ \forall i \neq \bar{i},\ \underline{b}_i = \underline{a}_i & \forall i \neq \bar{i},\ \underline{c}_i = \underline{b}_i & \Longrightarrow & \forall i \neq \bar{i},\ \underline{c}_i = \underline{b}_i = \underline{a}_i \end{array}\right\} \Rightarrow$$

$$\Rightarrow \forall i,\ \underline{c}_i = \underline{a}_i.$$

Quindi $C = A$, e dunque $\sigma \circ \tau$ è l'identità di $\mathcal{M}_{mn}(K)$ (perché $(\sigma \circ \tau)(A) = \sigma(\tau(A)) = \sigma(B) = C = A$, per ogni $A \in \mathcal{M}_{mn}(K)$). In maniera perfettamente analoga si prova che $\tau \circ \sigma$ è ancora l'identità di $\mathcal{M}_{mn}(K)$, e dunque σ è l'inversa di τ, come volevamo.
Tipo II:

L'inversa di $\tau = \underline{a}_{i'} \longleftrightarrow \underline{a}_{i''}$ è sé stessa. Infatti sia A una qualsiasi matrice di tipo $[m, n]$, sia B la matrice ottenuta da A tramite τ e C la matrice ottenuta da B ancora tramite τ. Per ogni $i \notin \{i', i''\}$ si ha $\underline{c}_i = \underline{b}_i = \underline{a}_i$, per i' si ha $\underline{c}_{i'} = \underline{b}_{i''} = \underline{a}_{i'}$ e per i'' si ha $\underline{c}_{i''} = \underline{b}_{i'} = \underline{a}_{i''}$. Quindi $C = A$, il che implica che $\tau \circ \tau$ è l'identità di $\mathcal{M}_{mn}(K)$, e dunque τ è inversa di sé stessa, come volevamo.
Tipo III
:

L'inversa di $\tau = \underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} + \lambda \underline{a}_{i'}$ è $\sigma = \underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} - \lambda \underline{a}_{i'}$. Infatti, sia A una qualsiasi matrice di tipo $[m, n]$, sia B la matrice ottenuta da A tramite τ e C la matrice ottenuta da B tramite σ. Per ogni $i \neq \bar{i}$ si ha $\underline{c}_i = \underline{b}_i = \underline{a}_i$, per $\bar{i}$ si ha

$$\underline{c}_{\bar{i}} = \underline{b}_{\bar{i}} - \lambda \underline{b}_{i'} = \underline{(\underline{a}_{\bar{i}} + \lambda \underline{a}_{i'}) - \lambda \underline{b}_{i'} = (\underline{a}_{\bar{i}} + \lambda \underline{a}_{i'}) - \lambda \underline{a}_{i'}} = \underline{a}_{\bar{i}}.$$

La condizione $\bar{i} \neq i'$ è stata usata nel terzo passaggio (quello sottolineato). Quindi $C = A$, da cui ricaviamo che $\sigma \circ \tau$ è l'identità di $\mathcal{M}_{mn}(K)$. In maniera perfettamente analoga si procede per la composizione $\tau \circ \sigma$, e dunque σ è l'inversa di τ, come volevamo.

Ora cominciamo ad applicare queste nozioni ai sistemi lineari.

Proposizione 2 *Se la matrice completa di un sistema lineare è ottenuta dalla matrice completa di un altro sistema lineare tramite una trasformazione elementare, allora i due sistemi sono equivalenti.*

Dimostrazione. Per trasformazioni di tipo II, i due sistemi differiscono solo per l'ordine in cui sono scritte le equazioni e dunque sono evidentemente equivalenti. Per definizione di trasformazione di I o di III tipo, si ha subito che le equazioni del secondo sistema sono combinazioni lineari delle equazioni del primo; ma siccome l'inversa della trasformazione elementare usata è ancora una trasformazione elementare, allo stesso modo si ha che le equazioni del primo sistema sono combinazioni lineari di quelle del secondo. Dunque, per una nota proposizione sui sistemi lineari, i nostri due sistemi sono equivalenti. C.V.D.

Corollario 1 *Se la matrice completa di un sistema lineare è ottenuta dalla matrice completa di un altro sistema lineare tramite una sequenza (finita) di trasformazioni elementari, allora i due sistemi sono equivalenti*[2].

Dimostrazione. Ovvia conseguenza della proposizione precedente, visto che la relazione di equivalenza tra sistemi lineari, essendo appunto una relazione di equivalenza, ha la proprietà transitiva. C.V.D.

L'idea fondamentale per risolvere qualsiasi sistema lineare in maniera efficiente è quella di trasformare la sua matrice con operazioni elementari, in modo da ottenere un sistema equivalente più semplice da risolvere. La seguente definizione introduce appunto un tipo di matrice che dà luogo a sistemi che si risolvono molto agevolmente, e quindi può essere presa come punto di arrivo per le trasformazioni.

Definizione 2 *Sia A una matrice di tipo $[m, n]$ su un campo K. La matrice A si dice* a scalini *se valgono le seguenti implicazioni:*
(a) $\quad a_{\bar{\imath}j} = 0 \ \forall j \implies a_{ij} = 0 \ \forall i > \bar{\imath}, \ \forall j$ *;*
(b) $\quad (a_{\bar{\imath}j} = 0 \ \forall j < \bar{\jmath}$ *e* $a_{\bar{\imath}\bar{\jmath}} \neq 0) \implies a_{ij} = 0 \ \forall i > \bar{\imath}, \ \forall j \leq \bar{\jmath}$ *.*
Un elemento che soddisfi l'ipotesi dell'implicazione (b) si dice pivot *della riga* $\underline{a}_{\bar{\imath}}$.

Vediamo ora come trasformare con operazioni elementari una qualsiasi matrice in una matrice a scalini: cerchiamo perciò di descrivere il procedimento generale detto appunto *di riduzione a scalini*. Useremo qui un linguaggio non troppo formale, perché il nostro obiettivo principale è di consentire una rapida acquisizione di questo metodo e dunque può essere troppo pesante approfondire i dettagli algoritmici in maniera strettamente formale, cosa che comunque può costituire un ottimo esercizio di Calcolo e Programmazione. Anzi, per i più interessati a questo tipo di questioni, può essere divertente implementare questo algoritmo, cioè creare un programma (software) che riduca a scalini le matrici: la descrizione che daremo ora può essere considerata come un primo passo nella schematizzazione dell'algoritmo.

[2] Questo corollario spiega tra l'altro il motivo per cui, in alcuni testi, due matrici sono dette *equivalenti* quando possono essere ottenute una dall'altra tramite una sequenza di trasformazioni elementari. Così si ha che sistemi con matrici equivalenti sono equivalenti. Il viceversa però non vale, se non nell'ipotesi aggiuntiva in cui i sistemi sono compatibili.

Osservazione 1 *Può essere a volte comodo effettuare su una matrice una trasformazione del tipo* $\underline{a}_{\bar{i}} \longrightarrow \lambda\underline{a}_{\bar{i}} + \mu\underline{a}_{i'}$. *Questa non è in generale una trasformazione elementare, tuttavia, se* $\lambda \neq 0$*, si può ottenere componendo le due operazioni elementari* $\underline{a}_{\bar{i}} \longrightarrow \underline{a}_{\bar{i}} + \frac{\mu}{\lambda}\underline{a}_{i'}$ *(tipo III) e* $\underline{a}_{\bar{i}} \longrightarrow \lambda\underline{a}_{\bar{i}}$ *(tipo I).*

Definizione 3 *Sia A una matrice di tipo $[m, n]$. Il procedimento* di riduzione a scalini *consiste nei seguenti passi:*

- *Si parta ponendo* $\bar{i} = 1$.
- *Fintanto che esiste una riga* $\underline{a}_i \neq \underline{0}$ *con* $i > \bar{i}$*, si ripeta il seguente ciclo di operazioni*[3]*:*
  - *INIZIO CICLO*
  - *Si ponga*
    $$\begin{array}{l} \bar{j} = \min\left\{j : \exists i \geq \bar{i}\ a_{ij} \neq 0\right\} \\ i' = \min\left\{i : i \geq \bar{i} \text{ e } \ a_{i\bar{j}} \neq 0\right\}. \end{array}$$
  - *Si effettui la trasformazione elementare* $\underline{a}_{\bar{i}} \longleftrightarrow \underline{a}_{i'}$.
  - *Per ogni* $i > \bar{i}$ *si effettui la trasformazione* $\underline{a}_i \longrightarrow \lambda_i\underline{a}_i + \mu_i\underline{a}_{\bar{i}}$ *con* λ_i *e* μ_i *tali che* $\frac{\mu_i}{\lambda_i} = -\frac{a_{i\bar{j}}}{a_{\bar{i}\bar{j}}}$ $(\lambda_i \neq 0)$.
  - *Si aumenti* $\bar{i}$ *di una unità.*
  - *FINE CICLO*

Osservazione 2 *La trasformazione nel terzo passo del ciclo può essere sempre effettuata, perché* $a_{\bar{i}\bar{j}} \neq 0$ *(basta porre per esempio* $\lambda_i = 1$ *e* $\mu_i = -\frac{a_{i\bar{j}}}{a_{\bar{i}\bar{j}}}$*, oppure* $\lambda_i = a_{\bar{i}\bar{j}}$ *e* $\mu_i = -a_{i\bar{j}}$*: la scelta può essere fatta in funzione della comodità dei calcoli) ed è una composizione di due trasformazioni elementari (cfr. osservazione 1). Si tenga presente che mano a mano che il procedimento descritto sopra viene svolto, la matrice trasformata (e i suoi elementi) viene sempre indicata con la lettera A, così come avviene nei linguaggi di programmazione. Se avessimo seguito l'uso comune nelle definizioni matematiche, avremmo dovuto cambiare nome ogni volta che si otteneva (con una trasformazione elementare) una nuova matrice. Lo stesso vale per altri simboli usati (variabili), come* $\bar{i}$*,* i'*,* $\bar{j}$*, ecc.*

[3] In alcuni linguaggi di programmazione istruzioni simili si chiamano WHILE…DO. Nel nostro caso, il significato è che il ciclo si deve interrompere quando tutte le righe dopo la riga $\bar{i}$ sono tutte nulle (o non ci sono proprio). All'inizio la riga $\bar{i}$ è la prima, ma durante il ciclo cambierà via via, aumentando di uno ad ogni ripetizione (è scritto nell'ultimo passo del ciclo).

Per chi preferisce un liguaggio ancora meno formale, riportiamo qui sotto una traduzione "più umana" dell'algoritmo di riduzione a scalini:

— Si individui il primo elemento non nullo della prima colonna non nulla e si indichino con i' e $\overline{j}$ i suoi indici di riga e colonna (in A): questo elemento sarà il nostro primo pivot.

— Si effettui la trasformazione elementare $\underline{a}_1 \longleftrightarrow \underline{a}_{i'}$. Così facendo il pivot è stato portato più in alto possibile.

— Per ogni $i > 1$ si effettui la trasformazione $\underline{a}_i \longrightarrow \lambda_i \underline{a}_i + \mu_i \underline{a}_{\overline{i}}$ con λ_i e μ_i tali che $\frac{\mu_i}{\lambda_i} = -\frac{a_{i\overline{j}}}{a_{\overline{ij}}}$ ($\lambda_i \neq 0$). Così facendo tutti gli elementi al di sotto del pivot sono stati trasformati in zero, e il pivot sarà più a sinistra dei pivot successivi, come richiesto per una matrice a scalini.

— Si ripeta tutto dall'inizio, applicandolo però solo alla sottomatrice ottenuta cancellando la prima riga (la prima riga rimarrà d'ora in poi sempre invariata)

La nuova prima riga da considerare nel secondo ciclo sarà quindi quella che in effetti è la seconda riga di A. Alla fine del secondo ciclo la seconda riga sarà messa da parte, e si ripeterà considerando solo le righe dalla terza in poi, e così via, finché è possibile.

Non ci addentreremo nella dimostrazione che l'algoritmo ora esposto funziona, cioè che davvero applicando l'algoritmo ad una qualsiasi matrice si ottiene alla fine una matrice a scalini. In effetti, dalla descrizione appena fatta si intuisce abbastanza bene che l'algoritmo mira a disporre i pivot nell'ordine richiesto in una matrice a scalini (quelli più in basso, sono più a destra), l'unico dubbio serio è se, durante lo svolgimento dell'algoritmo, qualcuno degli zeri ottenuti in passi precedenti possa ridiventare non nullo. Ma questo non accade perché, ad ogni passo, l'azione che le trasformazioni vengono ad avere sulla parte di matrice già sistemata coinvolgono solo zeri, e dunque non può ricomparire nulla di diverso da zero.

La maniera migliore per capire l'algoritmo, è applicarlo a molti esempi. Si consiglia quindi di studiare bene la terza parte "esempi ed esercizi". In tale sezione saranno descritte anche alcune varianti del procedimento di riduzione che, sebbene non siano strettamente richieste per l'esame, è utile conoscere (ad esempio possono aiutare a svolgere più rapidamente gli esercizi del compito).

Per la parte teorica, rimane solo da spiegare perché i sistemi con matrice completa a scalini sono facili da risolvere. Lo facciamo nella seguente definizione, ancora in maniera informale. Per chi ha letto attentamente la prima parte di questi appunti, non c'è in realtà niente di nuovo.

Definizione 4 *Consideriamo un sistema lineare con matrice completa A a scalini. Il* procedimento di sostituzione a ritroso *per la risoluzione di tale sistema consiste nei seguenti passi.*

- *Sia $\underline{a}_{\bar{i}}$ l'ultima riga non nulla.*
- *Se $\underline{a}_{\bar{i}}$ ha tutti gli elementi nulli tranne l'ultimo, allora il sistema è incompatibile e il procedimento termina.*
- *Sia $a_{\bar{i}\bar{j}}$ il pivot della riga $\underline{a}_{\bar{i}}$.*
- *Si ricavi l'incognita $x_{\bar{j}}$ rispetto alle rimanenti nell'equazione $\bar{i}$.*
- *In tutte le equazioni precedenti l'equazione $\bar{i}$ si sostituisca $x_{\bar{j}}$ con l'espressione trovata.*
- *Se $\bar{i} = 1$ si termina, altrimenti si diminuisca $\bar{i}$ di una unità, e si ripeta a partire dal terzo passo.*

Proposizione 3 *Dopo il procedimento di sostituzione a ritroso, le soluzioni sono tutte e sole quelle ottenibili ricavando le incognite $x_{\bar{j}}$ (corrispondenti alle colonne contenenti i pivot) secondo le espressioni calcolate durante il procedimento, dopo aver assegnato dei valori arbitrari alle incognite rimanenti.*

Tralasciamo la dimostrazione formale di quest'ultima proposizione (che concettualmente è assolutamente banale), visto che il procedimento di sostituzione a ritroso è stato descritto abbastanza informalmente. Il punto saliente da sottolineare è che i passaggi di sostituzione sono tutti dei "se e solo se".

A questo punto abbiamo tutti gli ingredienti per dare la definizione "ufficiale" del metodo di Gauss.

Definizione 5 *Il metodo di Gauss per la risoluzione di un qualunque sistema lineare consiste nelle seguenti due fasi:*
— Riduzione a scalini della matrice completa del sistema.
— Risoluzione tramite sostituzioni a ritroso del sistema corrispondente alla matrice a scalini ottenuta nella fase precedente.

Osservazione 3 *Per la proposizione 3 le soluzioni ottenute nella seconda fase del procedimento sono quelle del sistema a scalini. Per la proposizione 2 tale sistema a scalini è equivalente a quello assegnato. Dunque le soluzioni trovate col metodo di Gauss sono tutte e sole quelle del sistema assegnato, come volevamo.*

La cosa più importante ora è imparare ad utilizzare con padronanza tale metodo, in quanto esso è fondamentale soprattutto dal punto di vista computazionale. Passiamo quindi senz'altro alla parte esercitativa.

Esempi ed esercizi

Incominciamo a verificare il procedimento su matrici già considerate nella prima parte di questi appunti.

Esempio 1 *Riduciamo a scalini la matrice*

$$\begin{pmatrix} 3 & 2 & -9 \\ 1 & -1 & 2 \end{pmatrix}$$

— *Poniamo* $\bar{i} = 1$.
— *Poiché la riga* $\underline{a}_2$ *è non nulla* ($2 > \bar{i} = 1$), *dobbiamo far partire il primo ciclo:*
— *Si ha* $\bar{j} = 1$, $i' = 1$ *(infatti il primo elemento non nullo della prima colonna non nulla è* a_{11}*).*
— *Si effettua* $\underline{a}_1 \longleftrightarrow \underline{a}_1$ *(naturalmente questa trasformazione non cambia proprio niente; questo perché il pivot selezionato è già il più in alto possibile).*
— *Si annullano tutti gli elementi sotto il pivot (quindi, in questo caso, si agisce solo sulla seconda riga); si ha, per* $i = 2$, $a_{\bar{i}\bar{j}} = a_{11} = 3$ *e* $a_{i\bar{j}} = a_{21} = 1$ *quindi possiamo scegliere ad esempio* $\lambda_2 = 3$ *e* $\mu_2 = -1$:

$$\begin{pmatrix} 3 & 2 & -9 \\ 1 & -1 & 2 \end{pmatrix} \overset{\underline{a}_2 \to 3\underline{a}_2 - \underline{a}_1}{} \begin{pmatrix} 3 & 2 & -9 \\ 0 & -5 & 15 \end{pmatrix}$$

— *Aumentiamo* $\bar{i}$, *quindi* $\bar{i} = 2$, *e ritorniamo all'inizio del ciclo.*
— *Poiché non esistono righe oltre la seconda, non dobbiamo far partire il secondo ciclo, dunque l'algoritmo termina.*

Abbiamo ottenuto una matrice a scalini: l'algoritmo ha funzionato.

Osservazione 4 *Quanto fatto ora coincide con quanto fatto nella parte di "discussione", solo che lì lavoravamo direttamente sul sistema, invece che con la matrice. Naturalmente si poteva anche scegliere un'altra trasformazione, per esempio* $\underline{a}_2 \to \underline{a}_2 - \frac{1}{3}\underline{a}_1$ *(che è un unica trasformazione elementare) invece di* $\underline{a}_2 \to 3\underline{a}_2 - \underline{a}_1$*(che è la composizione successiva di* $\underline{a}_2 \to \underline{a}_2 - \frac{1}{3}\underline{a}_1$ *con* $\underline{a}_2 \to 3\underline{a}_2$*) o anche* $\underline{a}_2 \to 903\underline{a}_2 - 301\underline{a}_1$*. Si capisce che vanno bene tutte, visto che l'obiettivo è quello di annullare* a_{21}*, infatti si ottengono matrici diverse, ma tutte a scalini e tutte danno luogo a sistemi equivalenti. L'operazione* $\underline{a}_2 \to 3\underline{a}_2 - \underline{a}_1$ *ha il vantaggio di evitare le frazioni (e di non essere pesante come* $\underline{a}_2 \to 903\underline{a}_2 - 301\underline{a}_1$*).*

Esercizio 2 *Ridurre a scalini le seguenti matrici*

$$\begin{pmatrix} 2 & 3 & 5 & 3 \\ 1 & -1 & 1 & 0 \\ -1 & 7 & -2 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & -1 & 3 & 4 \\ 3 & 0 & 4 & 7 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & -1 & -5 \\ 4 & -3 & -2 & -3 \\ 8 & -1 & -4 & 0 \end{pmatrix},$$

e verificare che i procedimenti usati nella parte di "discussione" (pagg. 3–5) rispondono a questo scopo.

Gli esempi predendenti mostrano che il procedimento funziona bene nei casi già descritti. Nel seguente esempio facciamo vedere come funzioni nei casi che erano rimasti dubbi nella prima parte (vedi discussione alla fine di pag. 6).

Esempio 2 *Applichiamo il procedimento di riduzione a scalini alla matrice*

$$\begin{pmatrix} -2 & 3 & 1 & -2 \\ 4 & -6 & -1 & 3 \\ 1 & 1 & 1 & -3 \end{pmatrix}.$$

- *Poniamo* $\bar{i} = 1$.
- *Poiché la riga* $\underline{a}_2$ *è non nulla dobbiamo far partire il primo ciclo (infatti* $2 > \bar{i} = 1$*, e anzi anche la* $\underline{a}_3$ *è non nulla, comunque ne basta una) :*
- *Si ha* $\bar{j} = 1$*,* $i' = 1$ *(infatti il primo elemento non nullo della prima colonna non nulla è* a_{11}*).*
- *Lo scambio è inutile (sarebbe* $\underline{a}_1 \longleftrightarrow \underline{a}_1$*).*

- *Si annullano tutti gli elementi sotto il pivot: per* $i = 2$ *si ha* $a_{\overline{i}\overline{j}} = a_{11} = -2$ *e* $a_{i\overline{j}} = a_{21} = 4$*, quindi possiamo scegliere* $\lambda_2 = 1$ *e* $\mu_2 = 2$ *(perché* $\frac{2}{1} = -\frac{4}{-2}$*); mentre per* $i = 3$*, si ha* $a_{\overline{i}\overline{j}} = a_{11} = -2$ *e* $a_{i\overline{j}} = a_{31} = 1$*, quindi possiamo scegliere* $\lambda_3 = 2$ *e* $\mu_3 = 1$

$$\begin{pmatrix} -2 & 3 & 1 & -2 \\ 4 & -6 & -1 & 3 \\ 1 & 1 & 1 & -3 \end{pmatrix} \underset{\underline{a}_3 \to 2\underline{a}_3 + \underline{a}_1}{\underline{a}_2 \to \underline{a}_2 + 2\underline{a}_1} \begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & 0 & 1 & -1 \\ 0 & 5 & 3 & -8 \end{pmatrix}$$

- *Aumentiamo* $\overline{i}$*, quindi* $\overline{i} = 2$*, e ritorniamo all'inizio del ciclo.*

- *Poiché la riga* $\underline{a}_3$ *è non nulla dobbiamo far partire il secondo ciclo:*

- *Si ha* $\overline{j} = 2$*,* $i' = 3$ *(infatti, se cancelliamo la prima riga, nella parte rimanente della matrice, il primo elemento non nullo della prima colonna non nulla è* $a_{32} = -5$*).*

- *Si effettua* $\underline{a}_2 \longleftrightarrow \underline{a}_3$*:*

$$\begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & 0 & -2 & 2 \\ 0 & -5 & -3 & 8 \end{pmatrix} \underset{\underline{a}_2 \longleftrightarrow \underline{a}_3}{} \begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & -5 & -3 & 8 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

  (ecco il punto fondamentale, che risolve il dubbio di cui abbiamo parlato a pag. 5 e pag. 7! Bastava scambiare la seconda e la terza riga)

- *Si annullano tutti gli elementi sotto il pivot: per* $i = 3$ *si ha* $a_{\overline{i}\overline{j}} = a_{22} = -5$ *e* $a_{i\overline{j}} = a_{32} = 0$*, quindi la trasformazione da fare è* $\underline{a}_3 \to -5\underline{a}_3 - 0\underline{a}_2$ *(naturalmente questo passaggio è inutile, perché la matrice è già a scalini; lo mettiamo solo come ulteriore esempio per illustrare come si applicano meccanicamente i passi del procedimento):*

$$\begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & -5 & -3 & 8 \\ 0 & 0 & -2 & 2 \end{pmatrix} \underset{\underline{a}_3 \longrightarrow -5\underline{a}_3}{} \begin{pmatrix} -2 & 3 & 1 & -2 \\ 0 & -5 & -3 & 8 \\ 0 & 0 & 10 & -10 \end{pmatrix}$$

- *Aumentiamo* $\overline{i}$*, quindi* $\overline{i} = 3$*, e ritorniamo all'inizio del ciclo.*

- *Poiché non esistono righe oltre la terza, non dobbiamo far partire il terzo ciclo, dunque l'algoritmo termina.*

Naturalmente, dopo un po' di allenamento, ci si accorge facilmente che talvolta alcuni passaggi possono essere saltati, o che altre volte è più comodo usare altre trasformazioni che "si vedono ad occhio". Sottolineiamo però che effettivamente i metodi trattati consentono di affrontare ogni situazione in maniera meccanica, ed evitano errori come quelli di cui abbiamo parlato a pag. 7 (infatti lì l'errore era nella trasformazione $\underline{a}_4 \to \underline{a}_3 - \underline{a}_2$, che non è una trasformazione elementare).

Passiamo ora ad esercizi completi, cioè non solo sulla riduzione a scalini, ma sulla risoluzione di sistemi tramite il metodo di Gauss.

Esempio 3 *Risolviamo col metodo di Gauss il sistema*

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

La matrice completa è

$$\begin{pmatrix} 3 & -1 & 1 & -2 \\ 6 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

Riduciamo a scalini:

$$\begin{pmatrix} 3 & -1 & 1 & -2 \\ 6 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix} \underset{\underline{a}_3 \to 3\underline{a}_3 - \underline{a}_1}{\underline{a}_2 \to \underline{a}_2 - 2\underline{a}_1} \begin{pmatrix} 3 & -1 & 1 & -2 \\ 0 & 0 & 0 & 4 \\ 0 & 1 & -4 & 2 \end{pmatrix} \underline{a}_2 \longleftrightarrow \underline{a}_3$$

$$\begin{pmatrix} 3 & -1 & 1 & -2 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Applichiamo ora le sostituzioni a ritroso, seguendo la definizione 4.

- *L'ultima riga non nulla è la terza* ($\bar{i} = 3$)*.*
- *In tale riga sono nulli tutti gli elementi tranne l'ultimo. Dunque il sistema è incompatibile.*

Osservazione 5 *Anche nell'esempio di sopra, svolto applicando meccanicamente tutti i passi dei procedimenti definiti nella parte teorica, ci si poteva*

fermare prima. Infatti, dopo il primo ciclo della riduzione a scalini, è comparsa la riga $(0,0,0,4)$*: già a questo punto si poteva concludere che il sistema è incompatibile, dato che anche i sistemi corrispondenti alle matrici intermedie della riduzione sono equivalenti al sistema assegnato, e la presenza della riga in questione implica la presenza dell'equazione* $0 = 4$*, che non ha soluzioni.*
Ribadiamo dunque ancora una volta che sebbene i procedimenti indicati siano perfettamente efficienti, con un po' di pratica e buon senso, in casi particolari si può pervenire alla soluzione anche più rapidamente.

Esempio 4 *Risolvere il sistema lineare*

$$\begin{cases} x_1 + x_2 - 3x_3 + 4x_4 = 0 \\ x_1 + x_2 - 4x_3 + 5x_4 + 2 = 0 \\ 2x_1 + 5x_2 - 11x_3 + 9x_4 - 5 = 0 \\ -x_1 - 4x_2 + 10x_3 - 5x_4 + 1 = 0 \end{cases}.$$

Riduciamo a scalini la matrice completa:

$$\begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 1 & 1 & -4 & 5 & 2 \\ 2 & 5 & -11 & 9 & -5 \\ -1 & -4 & 10 & -5 & 1 \end{pmatrix} \underset{\substack{\underline{a}_2 \to \underline{a}_2 - \underline{a}_1 \\ \underline{a}_3 \to \underline{a}_3 - 2\underline{a}_1 \\ \underline{a}_4 \to \underline{a}_4 + \underline{a}_1}}{} \begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & -3 & 7 & -1 & 1 \end{pmatrix}$$

$$\underset{\underline{a}_2 \longleftrightarrow \underline{a}_3}{} \begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & -3 & 7 & -1 & 1 \end{pmatrix} \underset{\underline{a}_4 \to \underline{a}_4 + \underline{a}_2}{} \begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & 0 & 2 & 0 & -4 \end{pmatrix}$$

$$\underset{\underline{a}_4 \to \underline{a}_4 + 2\underline{a}_2}{} \begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix}.$$

Risolviamo a ritroso:

$$\begin{cases} x_1 + x_2 - 3x_3 + 4x_4 \qquad = 0 \\ \qquad 3x_2 - 5x_3 + x_4 - 5 = 0 \\ \qquad\qquad -x_3 + x_4 + 2 = 0 \\ \qquad\qquad\qquad 2x_4 \qquad = 0 \end{cases} \leftrightarrow \begin{cases} x_1 + 5 - 3 \cdot 2 + 4 \cdot 0 = 0 \to x_1 = 1 \\ 3x_2 - 5 \cdot 2 + 0 - 5 = 0 \to x_2 = 5 \\ -x_3 + 0 + 2 = 0 \to x_3 = 2 \\ x_4 = 0 \end{cases}.$$

Dunque il sistema ha un'unica soluzione: $(1,5,2,0)$.

Osservazione 6 *Se si volessero evitare del tutto le sostituzioni, si potrebbe semplificare ancora la matrice a scalini. Per esempio, a partire dalla matrice di sopra, si potrebbe continuare con le seguenti trasformazioni elementari:*

$$
\begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix}
\xrightarrow{\underline{a}_4 \to \frac{1}{2}\underline{a}_4}
\begin{pmatrix} 1 & 1 & -3 & 4 & 0 \\ 0 & 3 & -5 & 1 & -5 \\ 0 & 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}
\xrightarrow{\substack{\underline{a}_1 \to \underline{a}_1 - 4\underline{a}_4 \\ \underline{a}_2 \to \underline{a}_2 - \underline{a}_4 \\ \underline{a}_3 \to \underline{a}_3 - \underline{a}_4}}
$$

$$
\begin{pmatrix} 1 & 1 & -3 & 0 & 0 \\ 0 & 3 & -5 & 0 & -5 \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}
\xrightarrow{\underline{a}_3 \to -\underline{a}_3}
\begin{pmatrix} 1 & 1 & -3 & 0 & 0 \\ 0 & 3 & -5 & 0 & -5 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}
\xrightarrow{\substack{\underline{a}_1 \to \underline{a}_1 + 3\underline{a}_3 \\ \underline{a}_2 \to \underline{a}_2 + 5\underline{a}_3}}
$$

$$
\begin{pmatrix} 1 & 1 & 0 & 0 & -6 \\ 0 & 3 & 0 & 0 & -15 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}
\xrightarrow{\underline{a}_2 \to \frac{1}{3}\underline{a}_2}
\begin{pmatrix} 1 & 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 0 & -5 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}
\xrightarrow{\underline{a}_1 \to \underline{a}_1 - \underline{a}_2}
$$

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

La matrice ottenuta corrisponde al sistema:

$$
\begin{cases} x_1 - 1 = 0 \\ x_2 - 5 = 0 \\ x_3 - 2 = 0 \\ \quad x_4 = 0 \end{cases},
$$

che in pratica è già bell'e risolto. Il metodo che abbiamo applicato qui si chiama metodo di Gauss-Jordan, e può essere definito in generale. Tale metodo però comporta più fatica del metodo di Gauss, cioè, in termini più precisi, ha un maggiore costo computazionale. *I costi computazionali degli algoritmi sono uno strumento utile in ambito informatico: servono per capire quale programma è più veloce per raggiungere un determinato scopo. Chi è interessato ad un'esposizione dettagliata del metodo di Gauss-Jordan o ai costi computazionali, può consultare ad esempio [F. Orecchia,* Elementi di geometria e algebra lineare - Vol. II: Matrici, determinanti e sistemi lineari, *Liguori editore].*

Esempio 5 *Risolviamo col metodo di Gauss il sistema lineare*

$$\begin{cases} 2x_1 - 3x_2 + x_3 + 4x_4 - 7 = 0 \\ -4x_1 + 6x_2 - 2x_3 - 8x_4 + 14 = 0 \\ -2x_1 + 3x_2 + x_3 - 3 = 0 \end{cases} .$$

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 \\ -4 & 6 & -2 & -8 & 14 \\ -2 & 3 & 1 & 0 & -3 \end{pmatrix} \underset{\underline{a}_3 \to \underline{a}_3 + \underline{a}_1}{\overset{\underline{a}_2 \to \underline{a}_2 + 2\underline{a}_1}{}} \begin{pmatrix} 2 & -3 & 1 & 4 & -7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 4 & -10 \end{pmatrix} \underset{\underline{a}_2 \longleftrightarrow \underline{a}_3}{} $$

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 \\ 0 & 0 & 2 & 4 & -10 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

$$\begin{cases} 2x_1 - 3x_2 + x_3 + 4x_4 - 7 = 0 \\ \qquad 2x_3 + 4x_4 - 10 = 0 \\ \qquad\qquad 0 = 0 \end{cases} \rightarrow$$

$$\begin{cases} 2x_1 - 3x_2 + (5 - 2x_4) + 4x_4 - 7 = 0 \rightarrow x_1 = 1 + \frac{3}{2}x_2 - x_4 \\ \qquad x_3 = 5 - 2x_4 \\ \end{cases} .$$

Le soluzioni allora si ottengono tutte assegnando valori arbitrari alle incognite x_2 *e* x_4*, e ricavando* x_1 *e* x_3 *secondo le espressioni appena trovate. Dunque l'insieme delle soluzioni è*

$$\left\{ \left(\quad 1 + \frac{3}{2}s - t \quad , \quad s \quad , \quad 5 - 2t \quad , \quad t \quad \right) : s, t \in \mathbb{R} \right\} .$$

A questo punto il lettore può esercitarsi da solo, provando a risolvere col metodo di Gauss sistemi lineari proposti nei vari testi di algebra lineare (o anche in testi di altre materie; per esempio, il bilanciamento di reazioni chimiche è essenzialmente un esercizio di risoluzione di sistemi lineari).