Chapter 2

Sistemi di equazioni lineari e matrici

Per definire rigorosamente cosa intendiamo per equazione lineare e scrivere il generico esempio di equazione lineare, troviamo prima una notazione conveniente per denotare tali equazioni.
Infatti, per non avere limitazioni sul numero delle incognite, non possiamo continuare a indicarle con le lettere dell'alfabeto x, y, z etc., che sono in numero limitato, ma useremo sempre la stessa lettera, tradizionalmente la x, con degli indici numerici che ci dicono di quale incognita si tratta: x_1 indicherà quindi la prima incognita, x_2 la seconda, e così via in generale x_n indicherà la n-esima incognita, dove n è un numero naturale.
Possiamo allora dire che per *equazione lineare in n incognite* $x_1, x_2, \dots, x_n$ (i puntini indicano che stiamo omettendo di scrivere le incognite tra la seconda e l'ultima) intendiamo un'equazione del tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{2.1}$$

dove $b, a_1, a_2, \dots, a_n$ sono elementi di un campo (solitamente, il campo dei numeri reali o quello dei complessi) che svolgono il ruolo rispettivamente di termine noto e coefficienti delle incognite (per ogni incognita x_i, denotiamo il suo coefficiente con una lettera, a, con lo stesso indice dell'incognita).

Dare una soluzione dell'equazione (2.1) significa trovare degli elementi del campo, ovvero dei numeri, che sostituiti alle incognite rendano l'uguaglianza vera.
Ad esempio, nell'equazione lineare in due incognite $x_1 - x_2 = 1$ a coefficienti nel campo dei reali $\mathbb{R}$, ponendo $x_1 = 2$ e $x_2 = 1$ si ottiene l'uguaglianza vera $2 - 1 = 1$, mentre ad esempio ponendo $x_1 = 1$ e $x_2 = 2$ si ottiene $1 - 2 = 1$ che è falsa.

Da questo semplice esempio si vede come dare una soluzione dell'equazione $x_1 - x_2 = 1$ significa non solo dare *due* valori numerici, da sostituire alle due incognite dell'equazione, ma è necessario precisare quale valore vada sostituito alla prima incognita e quale alla seconda, ovvero specificare in quale ordine stiamo prendendo questi due elementi.
La soluzione data di tale equazione può allora essere pensata e scritta come una *coppia ordinata* di numeri, che denotiamo $(2, 1)$. La coppia $(2, 1)$ è una soluzione dell'equazione $x_1 - x_2 = 1$, mentre la coppia $(1, 2)$ non lo è.
Analogamente, per un'equazione con 3 incognite, una sua soluzione sarà data da una terna ordinata: ad esempio, se l'equazione è $x_1 - x_2 + x_3 = 2$, possiamo dire che la terna ordinata $(3, 2, 1)$ è una sua soluzione, in quanto sostituendo $x_1 = 3, x_2 = 2, x_3 = 1$ si ottiene l'uguaglianza vera $3 - 2 + 1 = 2$; la terna $(2, 1, 3)$ invece, non è una sua soluzione.
In generale, per equazioni con n incognite dovremo usare n-uple ordinate $(v_1, v_2, \dots, v_n)$: possiamo allora dare la seguente:

Definizione 2.1. Data un'equazione lineare $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ in n incognite a coefficienti in un campo $\mathbb{K}$, si dice *soluzione* dell'equazione una n-upla ordinata $(v_1, v_2, \dots, v_n)$ di elementi di $\mathbb{K}$ tale che sostituendo v_1 al posto di x_1, v_2 al posto di x_2 etc. fino a v_n al posto di x_n l'equazione risulta verificata (ovvero l'uguaglianza $a_1v_1 + a_2v_2 + \cdots + a_nv_n = b$ risulta vera).

Ora, un *sistema di equazioni lineari* è semplicemente un insieme di equazioni lineari.
Per scrivere un generico tale sistema, dobbiamo risolvere un problema di notazione simile a quello affrontato quando abbiamo scritto la generica equazione lineare, ovvero abbiamo bisogno di una notazione efficace per indicare i diversi coefficienti delle incognite nelle diverse equazioni del sistema.
A questo scopo, nell'espressione della generica equazione lineare $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ faremo precedere sia i coefficienti sia il termine noto da un ulteriore indice che ci dice di quale equazione del sistema si tratta: la prima equazione del sistema sarà cioè denotata $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$, la seconda $a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$ e così via.
Allora, il generico sistema di equazioni lineari con n incognite e m equazioni (il numero di incognite può anche essere diverso dal numero di equazioni, perciò li indichiamo con due lettere diverse) sarà

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.2}$$

(i puntini indicano che stiamo omettendo di scrivere le equazioni tra la seconda e l'ultima).
Possiamo quindi dare la seguente

Definizione 2.2. Una soluzione del sistema (2.2) è una n-upla $(v_1, v_2, \ldots, v_n) \in \mathbb{K}^n$ che è soluzione comune di tutte le equazioni del sistema.

Nel prossimo paragrafo vedremo un algoritmo che ci consentirà di determinare tutte le soluzioni di un sistema. In particolare, scopriremo che possono verificarsi solo le seguenti tre possibilità[1]:

- il sistema non ha nessuna soluzione
- il sistema ha una sola soluzione
- il sistema ha infinite soluzioni

Prima di entrare nei dettagli, vediamo un esempio di ciascuna di queste possibilità, con l'obiettivo di iniziare a capire le ragioni per cui esse possono verificarsi.

Non è difficile esibire un esempio di sistema con infinite soluzioni. Ad esempio, consideriamo il seguente sistema formato da una sola equazione in due incognite

$$\{\ x_1 + x_2 = 0.$$

Una soluzione del sistema è una coppia di numeri reali tali che la loro somma dà come risultato zero: questo significa che i numeri devono essere uno l'opposto dell'altro, e quindi scelto un qualunque $t \in \mathbb{R}$, la coppia $(t, -t)$ è una soluzione: le soluzioni sono quindi infinite, tante quanti i numeri reali.
Aggiungiamo ora alla $x_1 + x_2 = 0$ un'altra condizione, ottenendo quindi un sistema di due equazioni, ad esempio

[1] Questo è un fatto caratteristico delle equazioni lineari: per una generica equazione possono verificarsi anche altri casi, ad esempio l'equazione $x^2 = 9$ ha due soluzioni, $x = 3$ e $x = -3$.

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 0 \end{cases} \tag{2.3}$$

Le soluzioni del sistema sono quindi le coppie che soddisfano non solo la prima equazione, cioè come abbiamo detto tutte quelle del tipo $(t, -t)$, ma anche la seconda, che afferma semplicemente che $x_1 = x_2$, cioè i due elementi della coppia devono essere non solo opposti ma anche uguali tra loro. Ma l'unico numero reale uguale al suo opposto è lo zero, e quindi il sistema ha come unica soluzione la coppia $(0, 0)$.
Questo esempio suggerisce che in generale più equazioni ci sono in un sistema, maggiori sono i vincoli che imponiamo sulle incognite e quindi meno n-uple ci saranno che soddisfano tutte le condizioni espresse dalle equazioni, ovvero meno soluzioni: il sistema (2.3) sembra ad esempio suggerire che con due incognite, due condizioni siano sufficienti a ottenere una sola soluzione.
Tuttavia, è facile fare un altro esempio che mostra che questa prima impressione non è del tutto esatta: consideriamo il sistema

$$\begin{cases} x_1 + x_2 = 0 \\ 2x_1 + 2x_2 = 0 \end{cases} \tag{2.4}$$

Ora, è immediato vedere che le soluzioni $(t, -t)$ della prima equazione soddisfano tutte anche la seconda, quindi il sistema continua ad avere le infinite soluzioni $(t, -t)$. Questo accade perché la seconda equazione è in realtà del tutto equivalente alla prima (mettendo in evidenza il 2, si può riscrivere $2x_1 + 2x_2 = 0$ come $2(x_1 + x_2) = 0$, ovvero, dividendo per 2, proprio la prima equazione) e non aggiunge nessun nuovo vincolo sulle incognite: si tratta di un'equazione superflua, la cui presenza o meno non cambia l'insieme delle soluzioni.
Le equazioni superflue presenti in un sistema possono essere tuttavia molto meno evidenti che nel caso appena visto. Ad esempio, consideriamo il sistema di due equazioni in tre incognite

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \end{cases} \tag{2.5}$$

Una qualunque terna (x_1, x_2, x_3) che verifica le due equazioni soddisfa necessariamente anche l'uguaglianza che si ottiene sommandole membro a membro, ovvero

$$(x_1 + x_2 + x_3) + (2x_1 + x_2 + 3x_3) = 1 + 2$$

cioè, svolgendo i conti,

$$3x_1 + 2x_2 + 4x_3 = 3$$

Essendo tale equazione una conseguenza delle prime due, aggiungerla al sistema non modifica l'insieme delle soluzioni: in altre parole, il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ 3x_1 + 2x_2 + 4x_3 = 3 \end{cases} \tag{2.6}$$

contiene un'equazione superflua, dipendente dalle altre, certamente meno evidente a prima vista che nel caso del sistema (2.4).
Naturalmente, equazioni superflue possono essere ottenute anche con combinazioni più complicate della somma delle prime due equazioni, ad esempio sempre in riferimento al sistema (2.5), una terna che soddisfi le due equazioni necessariamente soddisfa anche l'uguaglianza

$$5(x_1 + x_2 + x_3) + (-3)(2x_1 + x_2 + 3x_3) = 5 \cdot 1 + (-3) \cdot 2 \tag{2.7}$$

cioè, svolgendo i conti,

$$-x_1 + 2x_2 - 4x_3 = -1$$

ovvero anche nel sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ -x_1 + 2x_2 - 4x_3 = -1 \end{cases} \tag{2.8}$$

la terza equazione è superflua, in un modo forse ancora meno evidente.

Per quello che riguarda i sistemi senza soluzioni, è abbastanza semplice esibirne uno. Ad esempio, il sistema di due equazioni in due incognite seguente

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 1 \end{cases}$$

è evidentemente privo di soluzioni, in quanto se la somma di due numeri è uguale a 0 non può certamente nello stesso tempo essere uguale a 1.
In altre parole, le due equazioni del sistema sono tra loro incompatibili, ovvero esprimono condizioni contraddittorie.

Per questo motivo, un sistema che non ha soluzioni si dice *incompatibile* (e per contro, si dirà *compatibile* un sistema che ha almeno una soluzione).

Analogamente a quanto fatto sopra per le equazioni superflue, si possono costruire esempi di sistemi in cui l'incompatibilità di una equazione con le altre non è così evidente come nel semplice sistema precedente.

Ad esempio, prendiamo sempre come punto di partenza il sistema (2.5). Come abbiamo visto sopra, una terna che soddisfi le due equazioni soddisfa anche l'uguaglianza $3x_1 + 2x_2 + 4x_3 = 3$ che si ottiene sommando le due equazioni membro a membro.
Ma allora, se modifichiamo solo il termine noto di quest'ultima uguaglianza, ne otteniamo una che è incompatibile con le altre due: ad esempio, il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ 3x_1 + 2x_2 + 4x_3 = 5 \end{cases} \tag{2.9}$$

non ha soluzioni, perchè per una qualunque terna che soddisfi le prime due equazioni si deve avere che $3x_1 + 2x_2 + 4x_3$ è uguale a 3, e non a 5.

Osservazione 2.3. Osserviamo che un sistema di equazioni in cui i termini noti siano tutti uguali a zero (un tale sistema si dice *omogeneo*) ha sempre almeno la soluzione $(0, 0, \dots, 0)$, in quanto ponendo tutte le incognite uguali a zero si ottengono uguaglianze vere. Quindi i sistemi omogenei sono sempre compatibili. Vedremo più avanti (Proposizione 2.16) altre importanti caratteristiche dei sistemi omogenei che li distinguono dai sistemi non omogenei.

2.1 Matrice di un sistema lineare

Ora osserviamo che, ovviamente, per conoscere un sistema abbiamo bisogno solo di sapere, equazione per equazione, quali sono i coefficienti che moltiplicano ogni singola incognita e i termini noti.
Quindi, se, dato un sistema, scriviamo una tabella di numeri disposti in righe e in colonne in modo che in ogni riga ci siano i coefficienti delle incognite di una certa equazione (ordinati secondo l'ordine scelto delle incognite) e il termine noto, tale tabella conterrà tutte le informazioni che ci servono sul sistema. Ad esempio, il sistema

$$\begin{cases} x_1 + 3x_2 = 5 \\ 2x_1 - x_2 = 4 \end{cases} \tag{2.10}$$

può essere rappresentato dalla tabella

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \end{pmatrix} \tag{2.11}$$

che chiameremo la *matrice completa del sistema*.

In generale, una *matrice* può essere definita come una tabella rettangolare di elementi di un campo $\mathbb{K}$, detti le sue *entrate*, disposti in righe e in colonne.

Analogamente alla notazione che abbiamo introdotto per identificare i coefficienti delle incognite di un sistema, per denotare la generica entrata di una matrice useremo due indici: il primo che ci dice in quale riga della matrice si trova, il secondo che ci dice in quale colonna. Una generica matrice sarà quindi

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{2.12}$$

Come si vede, tale matrice ha m righe e n colonne; la sua generica entrata è del tipo a_{ij}, dove il primo indice è detto *indice di riga* e va da 1 a m, mentre il secondo, detto *indice di colonna*, va da 1 a n; si dice anche che a_{ij} è *l'entrata di posto i j.*

Ad esempio, nella matrice $\begin{pmatrix} 5 & 3 \\ 2 & -1 \\ 0 & 7 \end{pmatrix}$ che ha tre righe e due colonne, 5 è la prima entrata della prima riga, e quindi $a_{11} = 5$; il numero 7 invece lo troviamo in corrispondenza della terza riga e seconda colonna, quindi $a_{32} = 7$.
Come vedremo, il concetto di matrice è di fondamentale importanza e comparirà in molti contesti in questo corso, e nei capitoli successivi ne faremo una trattazione indipendente e più approfondita. Nel contesto dei sistemi di equazioni lineari, non solo la matrice completa costituisce una "fotografia" fedele di un sistema e contiene tutte le informazioni necessarie a determinarlo, ma sarà anche l'oggetto sul quale lavoreremo per risolverlo.
Per motivi che vedremo, sarà importante prendere in considerazione anche la matrice che contiene solo i coefficienti delle incognite, senza l'ultima colonna formata dai termini noti: si ottiene così la cosiddetta *matrice dei coefficienti del sistema.* Ad esempio, la matrice dei coefficienti del sistema (2.10) è

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \tag{2.13}$$

Ora, rappresentare un sistema mediante la sua matrice completa consente di identificare ogni sua equazione con una n-upla: la generica equazione, diciamo la i-esima

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \tag{2.14}$$

è rappresentata nella matrice completa dall'i-esima riga R_i

$$\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} & b_i \end{pmatrix}$$

e questa riga può a sua volta essere pensata come la $n+1$-upla $(a_{i1}, a_{i2}, \ldots, a_{in}, b_i) \in \mathbb{K}^{n+1}$ (dove $\mathbb{K}$ è il campo dei coefficienti delle equazioni).
Questa corrispondenza è tale che qualunque combinazione delle equazioni del sistema (come quelle viste nel paragrafo precedente per costruire equazioni superflue) corrisponde a una combinazione lineare delle righe corrispondenti, viste come elementi di $\mathbb{K}^{n+1}$, e viceversa.
Infatti, se moltiplichiamo entrambi i membri dell'equazione (2.14) per un $c \in \mathbb{K}$, otteniamo

$$c(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) = cb_i$$

ovvero, svolgendo i conti a primo membro, la nuova equazione

$$ca_{i1}x_1 + ca_{i2}x_2 + \cdots + ca_{in}x_n = cb_i$$

e tale equazione corrisponde alla $(n+1)$-upla

$$cR_i = (ca_{i1}, ca_{i1_2}, \ldots, ca_{in}, cb_i)$$

ottenuta moltiplicando la $(n+1)$-upla $R_i = (a_{i1}, a_{i1_2}, \ldots, a_{in}, b_i)$ (che rappresenta l'equazione originale) per c.
Allo stesso modo, se sommiamo membro a membro l'equazione (2.14) con un'altra equazione $a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j$ del sistema, otteniamo

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_i + b_j$$

ovvero, raccogliendo gli addendi che contengono la stessa incognita, messa in evidenza,

$$(a_{i1} + a_{j1})x_1 + (a_{i2} + a_{j2})x_2 + \cdots + (a_{in} + a_{jn})x_n = b_i + b_j$$

si ottiene una nuova equazione rappresentata dalla $n+1$-upla

$$R_i + R_j = (a_{i1} + a_{j1}, a_{i2} + a_{j2}, \ldots, a_{in} + a_{jn}, b_i + b_j)$$

che si ottiene sommando le $n+1$-uple $R_i = (a_{i1}, a_{i1_2}, \ldots, a_{in}, b_i)$ e $R_j = (a_{j1}, a_{j_2}, \ldots, a_{jn}, b_j)$ che rappresentavano le due equazioni originali.
Quindi, se facciamo più in generale una combinazione di due o più equazioni di un sistema, rappresentate dalle righe $R_1, R_2, \ldots, R_m \in \mathbb{K}^{n+1}$, l'equazione ottenuta corrisponderà a una combinazione

$$c_1R_1 + c_2R_2 + \cdots + c_mR_m$$

delle righe corrispondenti.
Ad esempio, nel sistema (2.5), se come abbiamo visto in (2.7) moltiplichiamo (membro a membro) la prima equazione per 5 e la sommiamo alla seconda moltiplicata per -3, otteniamo la nuova equazione $-x_1 + 2x_2 - 4x_3 = -1$. Nella matrice completa

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 \end{pmatrix} \tag{2.15}$$

l'operazione corrispondente non è nient'altro che la combinazione lineare

$$5R_1 + (-3)R_2 = 5(1, 1, 1, 1) + (-3)(2, 1, 3, 2) = (-1, 2, -4, -1)$$

delle sue due righe (viste come elementi di $\mathbb{R}^4$).
Inoltre, come abbiamo visto, se una terna (x_1, x_2, x_3) soddisfa il sistema (2.5) essa soddisferà anche l'equazione $-x_1+2x_2-4x_3 = -1$, e in generale soddisferà tutte le possibili equazioni che si ottengono combinando le sue equazioni, ovvero tutte le equazioni che corrispondono alle combinazioni lineari delle righe R_1 e R_2 della matrice completa.
In generale, dato un sistema di m equazioni in n incognite, con matrice completa avente come righe $R_1, R_2, \ldots, R_m$, una n-upla $(x_1, x_2, \ldots, x_n)$ che soddisfa il sistema verifica anche tutte le equazioni corrispondenti alle $(n+1)$-uple del tipo $c_1R_1 + c_2R_2 + \cdots + c_mR_m$, ovvero quelle appartenenti al sottospazio

$$\langle R_1, R_2, \ldots, R_m\rangle$$

generato dalle righe $R_1, R_2, \ldots, R_m$ (viste come elementi di $\mathbb{K}^{n+1}$), in quanto per definizione (Teorema/Definizione 1.28) tale sottospazio è formato proprio da tutte le combinazioni lineari $c_1R_1 + c_2R_2 + \cdots + c_mR_m$.

Da queste osservazioni si può dedurre il seguente risultato, che ci fornisce un criterio sufficiente perché due sistemi siano *equivalenti*, ovvero abbiano le stesse soluzioni:

Proposizione 2.4. Siano dati due sistemi di equazioni lineari in n incognite, il primo con matrice completa formata dalle righe $R_1, R_2, \ldots, R_m$ e il secondo con matrice completa formata dalle righe $\tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l$. Se

$$\langle R_1, R_2, \ldots, R_m\rangle = \langle \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l\rangle$$

allora i due sistemi sono equivalenti.

Proof. Se una n-upla $x = (x_1, x_2, \ldots, x_n)$ è una soluzione del primo sistema, allora essa verifica tutte le sue equazioni, rappresentate dalle righe $R_1, R_2, \ldots, R_m$ della sua matrice completa. Come abbiamo osservato sopra, essa verifica allora anche tutte le equazioni corrispondenti alle $(n+1)$-uple contenute nel sottospazio $\langle R_1, R_2, \ldots, R_m \rangle$. Essendo tale sottospazio uguale a $\langle \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l \rangle$, come affermato nell'ipotesi, x verifica quindi tutte le equazioni corrispondenti alle $(n+1)$-uple del sottospazio $\langle \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l \rangle$, e in particolare[2] $\tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l$ stesse, che rappresentano le equazioni del secondo sistema: quindi x è soluzione anche del secondo sistema.
Viceversa[3], se una n-upla $x = (x_1, x_2, \ldots, x_n)$ è una soluzione del secondo sistema, allora essa verifica tutte le sue equazioni, rappresentate dalle righe $\tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l$ della sua matrice completa. Quindi essa verifica allora anche tutte le equazioni corrispondenti alle $(n+1)$-uple contenute nel sottospazio $\langle \tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_l \rangle$. Essendo tale sottospazio uguale a $\langle R_1, R_2, \ldots, R_m \rangle$, come affermato nell'ipotesi, x verifica quindi tutte le equazioni corrispondenti alle $(n+1)$-uple del sottospazio $\langle R_1, R_2, \ldots, R_m \rangle$, e in particolare $R_1, R_2, \ldots, R_m$ stesse, che rappresentano le equazioni del primo sistema: quindi x è soluzione anche del primo sistema. $\square$

Il metodo che useremo per risolvere un sistema, e che vedremo nel paragrafo seguente, consiste proprio nel trasformare il sistema dato in un sistema equivalente più semplice, nel quale avremo eliminato tutte le equazioni superflue (che si ottengono come combinazione delle altre).

2.2 L'algoritmo di eliminazione di Gauss-Jordan (o di riduzione a gradini)

Il metodo che vedremo ora per risolvere un qualunque sistema con m equazioni lineari in n incognite può essere spiegato come una generalizzazione dei metodi tradizionalmente usati per la risoluzione dei sistemi di due equazioni

[2] All'interno del sottospazio $\langle v_1, v_2, \ldots, v_n \rangle$ generato da un insieme di vettori e costituito come sappiamo da tutte le combinazioni lineari $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ ci sono sempre anche i vettori $v_1, v_2, \ldots, v_n$ stessi, in quanto ciascuno di loro può essere espresso come combinazione lineare: $v_1 = 1v_1 + 0v_2 + \cdots + 0v_n$, $v_2 = 0v_1 + 1v_2 + \cdots + 0v_n$, e così via.

[3] Dimostrare che i due sistemi hanno le stesse soluzioni significa dimostrare che l'insieme delle soluzioni del primo è uguale all'insieme delle soluzioni del secondo, ovvero (come previsto dalla definizione di uguaglianza di insiemi) che ogni n-upla che è soluzione del primo sistema è soluzione anche del secondo, e viceversa ogni n-upla soluzione del secondo sistema è anche soluzione del primo.

in due incognite. Per ricordare quali sono questi metodi, prendiamo ad esempio il sistema

$$\begin{cases} x_1 + x_2 = 0 \\ -x_1 + x_2 = 1 \end{cases} \tag{2.16}$$

Solitamente, per risolvere tale sistema si ricava una delle incognite in funzione dell'altra usando una delle due equazioni, ad esempio dalla prima equazione si trova $x_1 = -x_2$, e si sostituisce l'espressione così ottenuta nell'altra equazione:

$$-(-x_2) + x_2 = 1$$

ovvero

$$2x_2 = 1$$

In questo modo, abbiamo *eliminato* la prima incognita dalla seconda equazione che è diventata una semplice equazione di primo grado con una sola incognita, che ha come soluzione $x_2 = \frac{1}{2}$. A questo punto, per ricavare x_1 basta sostituire il valore ottenuto di x_2 nella prima equazione, ovvero

$$x_1 + \frac{1}{2} = 0 \ \rightarrow \ x_1 = -\frac{1}{2}.$$

Quello che ci ha permesso di risolvere il sistema è stato quindi aver ridotto il numero di incognite presenti in una delle equazioni.
Allo stesso risultato si può arrivare, equivalentemente, ad esempio sommando membro a membro le due equazioni: se $x_1 + x_2 = 0$ e $-x_1 + x_2 = 1$ allora

$$(x_1 + x_2) + (-x_1 + x_2) = 0 + 1$$

ovvero, facendo i conti, si ottiene come sopra $2x_2 = 1$.
Questo secondo metodo, apparentemente più artificioso, in realtà si rivela più semplice se si lavora sulla matrice completa del sistema invece che sulle equazioni. Infatti, la matrice completa del sistema (2.16) è

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \tag{2.17}$$

Come sappiamo, sommare membro a membro le due equazioni equivale a sommare tra loro le due righe: sostituendo poi tale somma alla seconda riga originale si ottiene, senza dover maneggiare le incognite e dover fare sostituzioni o semplificazioni,

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \tag{2.18}$$

che corrisponde proprio al sistema ridotto

$$\begin{cases} x_1 + x_2 = 0 \\ 2x_2 = 1 \end{cases}$$

risolvibile come abbiamo visto sopra risolvendo prima l'equazione con una sola incognita.

Questo stesso procedimento di eliminazione di incognite, realizzato lavorando sulle righe della matrice completa, funziona in realtà per risolvere qualunque sistema, qualunque sia il numero di equazioni e il numero di incognite.
Più precisamente, in generale ci porremo l'obiettivo di trasformare le equazioni in modo che a partire dalla prima in esse compaiano sempre meno incognite. Se, per darci un criterio, scegliamo di eliminarle seguendo l'ordine $x_1, x_2, \ldots, x_n$, questo significa che vogliamo che le righe della matrice completa inizino con un numero sempre maggiore di zeri.
Ad esempio, la seguente matrice

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 4 & 5 \end{pmatrix}$$

nella quale le righe iniziano con un numero sempre maggiore di zeri, ha come sistema corrispondente

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

che ha la proprietà desiderata che le sue equazioni presentano un numero decrescente di incognite.
Possiamo dare la seguente

Definizione 2.5. Una matrice si dice a gradini se, andando dalla prima all'ultima, ogni riga inizia con un numero sempre maggiore di zeri.
Il primo elemento non nullo in ogni riga di una matrice a gradini si chiama *pivot*.

In altre parole, una matrice è a gradini se in ogni riga il primo elemento non nullo compare con un indice di colonna sempre più grande. Ad esempio, delle matrici seguenti

$$\begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 1 & 3 & 5 \\ 0 & 2 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 0 & 3 & 5 \\ 0 & 4 & 0 & 6 \end{pmatrix}$$

la prima è a gradini perché i suoi pivot (7 nella prima riga, 4 nella seconda e 6 nella terza) si trovano, nell'ordine, sulla prima, seconda e quarta colonna (indice di colonna sempre più grande), mentre le altre due no (nella seconda, il primo elemento non nullo della terza riga sta nella stessa colonna del primo elemento non nullo della seconda riga; nella terza matrice, il primo elemento non nullo della terza riga sta in una colonna di indice più piccolo del primo elemento non nullo della seconda riga).
Un sistema si dice a gradini se la sua matrice completa è una matrice a gradini.
Il procedimento che descriveremo ora si chiama appunto *metodo di riduzione a gradini* o, dal momento che consiste nell'eliminare incognite, *metodo di eliminazione di Gauss-Jordan*.
Come stiamo per vedere, il procedimento di riduzione a gradini, oltre a semplificare il sistema, fa emergere anche le eventuali incompatibilità e le eventuali equazioni superflue presenti nel sistema.

Per trasformare un sistema in un sistema a gradini, trasformeremo la sua matrice completa in una matrice a gradini tramite le seguenti operazioni sulle sue righe, dette *operazioni elementari di primo, secondo e terzo tipo*:

(primo tipo): Scambiare tra loro due righe della matrice (in simboli, $R_i \leftrightarrow R_j$)

(secondo tipo): Moltiplicare una riga della matrice per un coefficiente non nullo (in simboli, $R_i \to cR_i$, con $c \neq 0$)

(terzo tipo): Sommare a una riga della matrice un'altra riga moltiplicata per un numero qualunque (in simboli, $R_i \to R_i + dR_j$)

Il fatto importante è che tali operazioni, che modificano le righe, corrispondono a modificare le equazioni del sistema *in modo però da non cambiare l'insieme delle soluzioni*, come dimostra il seguente risultato, corollario della Proposizione 2.4:

Proposizione 2.6. Se effettuiamo operazioni elementari di primo, secondo e terzo tipo sulla matrice completa di un sistema, la matrice trasformata è la matrice completa di un sistema *equivalente* a quello iniziale (ovvero avente le stesse soluzioni del sistema iniziale).

Proof. Siano $R_1, R_2, \dots, R_m$ le righe della matrice completa del sistema dato. In base alla Proposizione 2.4, ci basta verificare che se $\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_m$ sono le righe ottenute dopo una trasformazione mediante operazioni elementari, allora $\langle R_1, R_2, \dots, R_m\rangle = \langle \tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_m\rangle$, ovvero lo spazio generato non cambia.

Per quello che riguarda un'operazione elementare del primo tipo (ovvero scambiare due righe R_i e R_j tra loro) questo significa che

$$\langle R_1, \dots, R_i, \dots, R_j, \dots, R_m\rangle = \langle R_1, \dots, R_j, \dots, R_i, \dots, R_m\rangle$$

ma questo è immediato in quanto lo spazio generato da un insieme di vettori non dipende dall'ordine in cui disponiamo i vettori (dal momento che in uno spazio vettoriale vale la proprietà commutativa della somma, gli addendi di una combinazione lineare possono essere scritti in qualunque ordine).

Per quello che riguarda un'operazione elementare del secondo tipo (ovvero moltiplicare una riga R_i per uno scalare c), dobbiamo dimostrare che

$$\langle R_1, \dots, R_i, \dots, R_m\rangle = \langle R_1, \dots, cR_i, \dots, R_m\rangle \tag{2.19}$$

In effetti, se un vettore v appartiene a $\langle R_1, \dots, R_i, \dots, R_m\rangle$, esso si scrive come combinazione lineare $v = c_1R_1 + \cdots + c_jR_j + \cdots + c_mR_m$. Ma chiaramente vale l'uguaglianza

$$v = c_1R_1 + \cdots + c_jR_j + \cdots + c_mR_m = c_1R_1 + \cdots + \frac{c_j}{c}(cR_j) + \cdots + c_mR_m$$

che ci dice che v è anche combinazione lineare di $R_1, \dots, cR_i, \dots, R_m$, ovvero appartiene a $\langle R_1, \dots, cR_i, \dots, R_m\rangle$. Questo dimostra l'inclusione $\langle R_1, \dots, R_i, \dots, R_m\rangle \subseteq \langle R_1, \dots, cR_i, \dots, R_m\rangle$. Per dimostrare anche l'inclusione opposta, e quindi l'uguaglianza (2.19), sia v un vettore che appartiene a $\langle R_1, \dots, cR_i, \dots, R_m\rangle$: allora esso si scrive come combinazione lineare $v = c_1R_1 + \cdots + c_jcR_j + \cdots + c_mR_m$. Ma questa può essere chiaramente vista anche come una combinazione lineare di $R_1, \dots, R_i, \dots, R_m$, ovvero v appartiene anche a $\langle R_1, \dots, R_i, \dots, R_m\rangle$, il che conclude la dimostrazione dell'uguaglianza (2.19).

Infine, per quello che riguarda un'operazione elementare del terzo tipo (ovvero sostituire una riga R_i con $R_i + dR_j$), dobbiamo dimostrare che

$$\langle R_1, \dots, R_i, \dots, R_j, \dots, R_m\rangle = \langle R_1, \dots, R_i + dR_j, \dots, R_j, \dots, R_m\rangle \tag{2.20}$$

In effetti, se un vettore v appartiene a $\langle R_1, \dots, R_i, \dots, R_j, \dots, R_m\rangle$, esso si scrive come combinazione lineare $v = c_1R_1 + \cdots + c_iR_i + \cdots + c_jR_j + \cdots + c_mR_m$. Ma, aggiungendo e sottraendo l'addendo c_idR_j, vediamo che chiaramente vale l'uguaglianza

$$v = c_1R_1 + \cdots + c_iR_i + \cdots + c_jR_j + \cdots + c_mR_m =$$

$$= c_1R_1 + \cdots + c_i(R_i + dR_j) + \cdots + (c_j - c_id)R_j + \cdots + c_mR_m$$

che ci dice che v è anche combinazione lineare di $R_1, \ldots, R_i+dR_j, \ldots, R_j, \ldots, R_m$, ovvero appartiene a $\langle R_1, \ldots, R_i + dR_j, \ldots, R_j, \ldots, R_m\rangle$. Questo dimostra l'inclusione $\langle R_1, \ldots, R_i, \ldots, R_j, \ldots, R_m\rangle \subseteq \langle R_1, \ldots, R_i+dR_j, \ldots, R_j, \ldots, R_m\rangle$. Per dimostrare anche l'inclusione opposta, e quindi l'uguaglianza (2.20), sia v un vettore che appartiene a $\langle R_1, \ldots, R_i + dR_j, \ldots, R_j, \ldots, R_m\rangle$: allora esso si scrive come combinazione lineare

$$v = c_1R_1 + \cdots + c_i(R_i + dR_j) + \cdots + c_jR_j + \cdots c_mR_m$$

Ma, svolgendo i conti, si ha allora

$$v = c_1R_1 + \cdots + c_iR_i + \cdots + (c_j + c_id)R_j + \cdots c_mR_m$$

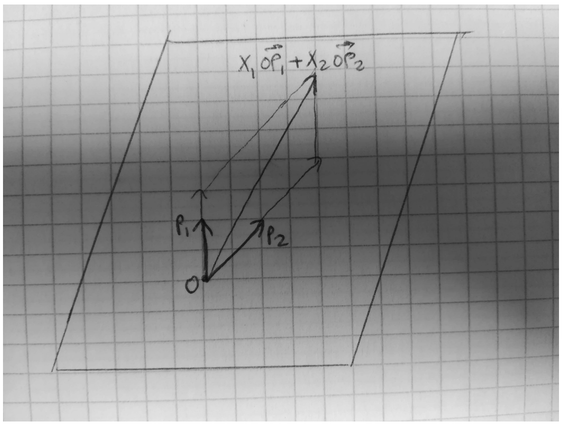
il che mostra che v appartiene anche a $\langle R_1, \ldots, R_i, \ldots, R_j, \ldots, R_m\rangle$ e conclude la dimostrazione dell'uguaglianza (2.20).

$\square$

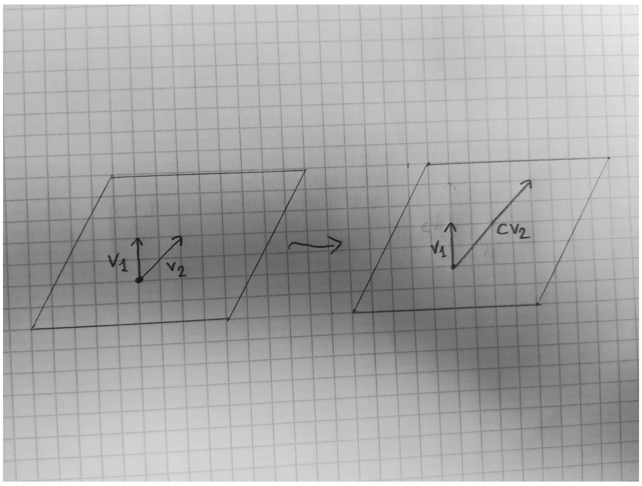
Osservazione 2.7. Si noti che nella dimostrazione appena vista in effetti non abbiamo mai usato il fatto che $R_1, R_2, \ldots, R_m$ sono righe della matrice completa di un sistema di equazioni, e neanche che siano n-uple: potrebbero essere vettori di un qualunque spazio vettoriale e la dimostrazione funzionerebbe lo stesso, in quanto nei passaggi fatti abbiamo usato solamente proprietà di somma e prodotto per scalari che sono valide in qualunque spazio vettoriale. Quindi, possiamo affermare che in un qualunque spazio vettoriale V su un campo $\mathbb{K}$, dati m vettori $v_1, \ldots, v_m$, valgono le seguenti

- $\langle v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m\rangle = \langle v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m\rangle$
- $\langle v_1, \ldots, v_i, \ldots, v_m\rangle = \langle v_1, \ldots, cv_i, \ldots, v_m\rangle \quad (c \neq 0)$
- $\langle v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m\rangle = \langle v_1, \ldots, v_i + dv_j, \ldots, v_j, \ldots, v_m\rangle, \quad d \in \mathbb{K}$

ovvero il sottospazio $\langle v_1, \ldots, v_m\rangle$ generato da un insieme di vettori non cambia se scambiamo il loro ordine, oppure se moltiplichiamo uno dei vettori generatori per uno scalare $c \neq 0$, oppure se sommiamo a uno dei generatori un multiplo di un altro generatore.
Ad esempio, se V è lo spazio tridimensionale V^3_O dei vettori geometrici applicati, e prendiamo due vettori $v_1 = \vec{OP_1}$, $v_2 = \vec{OP_2}$ indipendenti, il sottospazio $\langle v_1, v_2\rangle$ da loro generato coincide con l'insieme di tutti i vettori geometrici che giacciono sul piano a cui appartengono O, P_1, P_2

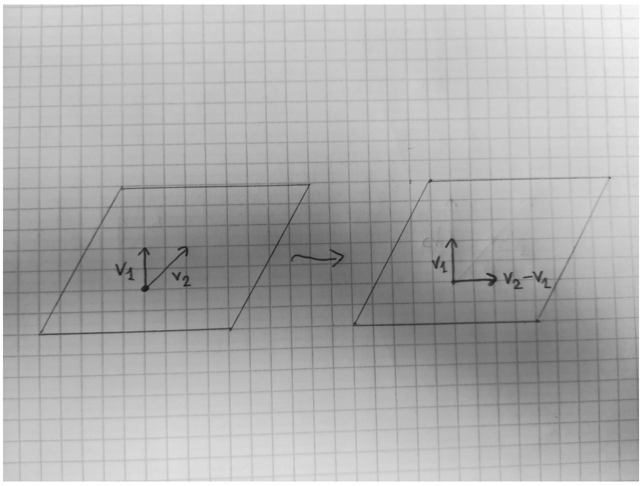


Allora, è chiaro che anche se moltiplicassimo uno dei due vettori per uno scalare c non nullo[4] il sottospazio generato non cambierebbe



oppure, come nel disegno seguente, possiamo sommare a uno dei due vettori un multiplo dell'altro e avremo sempre due generatori dello stesso spazio.

[4]In effetti, moltiplicare uno dei due vettori, ad esempio v_2, per $c = 0$ avrebbe l'effetto di trasformarlo nel vettore nullo $\bar{0}$, e il sottospazio generato cambierebbe drasticamente, in quanto in tal caso rimarrebbe solo v_1, che genera un sottospazio unidimensionale (la retta su cui si trova) e non tutto il piano.



Questo risultato può essere quindi anche usato per ottenere nuovi sistemi di generatori a partire da un sistema di generatori dato per un sottospazio (questo è utile se vogliamo costruire generatori con proprietà particolari, ad esempio ortogonali, come vedremo in un capitolo successivo).

Mostriamo ora come mediante l'uso delle tre operazioni elementari, ogni sistema possa essere trasformato in un sistema a gradini: sia

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ a_{31} & a_{32} & \dots & a_{3n} & b_3 \\ & \dots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{2.21}$$

la matrice completa del generico sistema (2.2).
L'algoritmo inizia come segue: se la prima entrata a_{11} della prima riga è uguale a zero, scambiamo tra loro due righe (applicando quindi un'operazione elementare del primo tipo) in modo da garantire che la nuova entrata a_{11} sia diversa da zero. Fatto ciò, applichiamo alla matrice (2.21) le seguenti operazioni elementari (del terzo tipo) sulle righe $R_2, \dots, R_m$ dalla seconda all'ultima:

$$R_2 \to R_2 - \frac{a_{21}}{a_{11}} R_1$$

$$R_3 \to R_3 - \frac{a_{31}}{a_{11}} R_1$$

$$\vdots$$

$$R_m \to R_m - \frac{a_{m1}}{a_{11}} R_1$$

(si noti che le operazioni si possono applicare proprio perché $a_{11} \neq 0$). Queste trasformazioni rendono sicuramente uguale a zero la prima entrata di ogni riga dalla seconda in poi, e eventualmente potrebbero aver annullato anche altre entrate, ovvero trasformano la matrice (2.21) in una matrice del tipo

$$\begin{pmatrix} a_{11} & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\ 0 & \dots & 0 & a'_{2k} & \dots & a'_{2n} & b'_2 \\ 0 & \dots & 0 & a'_{3k} & \dots & a'_{3n} & b'_3 \\ & \dots & & & & & \\ 0 & \dots & 0 & a'_{mk} & \dots & a'_{mn} & {}'b_m \end{pmatrix} \tag{2.22}$$

in cui possiamo supporre che la seconda riga sia quella che inizia con il minor numero di zeri (se così non fosse possiamo garantirlo con uno scambio di righe, ovvero con un'operazione elementare del primo tipo), con $a'_{2k} \neq 0$.
A questo punto, ripetiamo quanto fatto nella prima parte della trasformazione, applicando stavolta alle righe dalla terza in poi le trasformazioni elementari del terzo tipo

$$R_3 \to R_3 - \frac{a'_{3k}}{a'_{2k}} R_2$$

$$\vdots$$

$$R_m \to R_m - \frac{a'_{mk}}{a'_{2k}} R_2$$

che sono tali da annullare la prima entrata non nulla dalla terza riga in poi, ovvero da trasformare la matrice in una matrice del tipo

$$\begin{pmatrix} a_{11} & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\ 0 & \dots & 0 & a'_{2k} & \dots & a'_{2n} & b'_2 \\ 0 & \dots & 0 & 0 & \dots & a''_{3n} & b''_3 \\ & \dots & & & & & \\ 0 & \dots & 0 & 0 & \dots & a''_{mn} & b''_m \end{pmatrix} \tag{2.23}$$

In questo modo possiamo trasformare la matrice del nostro sistema in una matrice in cui ogni riga inizia con un numero sempre maggiore di zeri, ovvero nella matrice a gradini voluta, e il sistema corrispondente sarà equivalente al sistema iniziale in quanto la trasformazione è stata effettuata con operazioni elementari. Illustriamo subito tale procedimento con un esempio.

Sia

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ -x_1 + x_2 - 3x_3 = 0 \\ -x_1 + x_2 + x_3 = -3 \end{cases} \tag{2.24}$$

il sistema con matrice completa[5]

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -3 & 0 \\ -1 & 1 & 1 & -3 \end{array}\right). \tag{2.25}$$

Ora, vogliamo trasformare tale matrice in una matrice a gradini usando le operazioni elementari, in modo da ottenere un sistema a gradini equivalente al sistema (2.24).

Ricordiamo che, in base alla definizione di matrice a gradini, visto che il primo elemento a_{11} della prima riga è diverso da zero, e sta nella prima colonna, i primi elementi diversi da zero della seconda e della terza riga non possono stare anche loro nella prima colonna: in altre parole, dobbiamo trasformare la matrice in modo che a_{21} e a_{31} siano uguali a zero.

Otteniamo sicuramente questo scopo se applichiamo le operazioni elementari del terzo tipo $R_2 \to R_2 + R_1$ e $R_3 \to R_3 + R_1$: infatti,

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -3 & 0 \\ -1 & 1 & 1 & -3 \end{array}\right) \xrightarrow[\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 + R_1}]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right)$$

La matrice trasformata non è ancora una matrice a gradini in quanto il primo elemento non nullo della terza riga si trova in corrispondenza della stessa colonna (la seconda) del primo elemento non nullo nella seconda riga: dobbiamo far sì che $a_{32} = 0$. A questo scopo, basta applicare l'operazione elementare $R_3 \to R_3 - R_2$: così facendo si ottiene

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right) \xrightarrow[R_3 \to R_3 - R_2]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 4 & -3 \end{array}\right)$$

Abbiamo ottenuto quindi lo scopo desiderato di trasformare la matrice in una matrice a gradini.

Il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_2 - 2x_3 = 1 \\ 4x_3 = -3 \end{cases} \tag{2.26}$$

[5] D'ora in poi, nella matrice completa tracceremo a volte una linea per separare la matrice dei coefficienti dalla colonna dei termini noti.

corrispondente alla matrice trasformata è, come sappiamo, equivalente al sistema originale (2.24), quindi trovando le sue soluzioni avremo risolto il sistema (2.24).
Ora, il primo vantaggio di un sistema a gradini consiste nel fatto che nelle equazioni compaiono sempre meno incognite (leggendole dalla prima all'ultima): per risolverlo, basta quindi iniziare a risolvere le equazioni da quella che contiene meno incognite (l'ultima) e risalire mediante sostituzioni fino alla prima. Più precisamente, dall'ultima equazione $4x_3 = -3$ ricaviamo subito $x_3 = -\frac{3}{4}$; sostituendo il valore così trovato nella seconda equazione troviamo

$$2x_2 - 2x_3 = 1 \to 2x_2 = 1 + 2x_3 = 1 + 2(-\frac{3}{4}) = 1 - \frac{3}{2} = -\frac{1}{2} \to x_2 = -\frac{1}{4}$$

e analogamente, sostituendo i valori di x_2 e x_3 così ottenuti nella prima equazione troviamo

$$x_1 + x_2 + x_3 = 1 \to x_1 = 1 - x_2 - x_3 = 1 - (-\frac{1}{4}) - (-\frac{3}{4}) = 2$$

Riassumendo, la terna $(2, -\frac{1}{4}, -\frac{3}{4})$ è l'unica soluzione del sistema (2.26), ovvero del sistema iniziale (2.24).

La riduzione a gradini non solo semplifica il sistema grazie alla eliminazione di incognite, ma mette anche in evidenza eventuali "equazioni superflue" e incompatibilità tra le equazioni. Illustriamo ciò con due esempi. Consideriamo il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 1 \end{cases} \tag{2.27}$$

che ha come matrice completa

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 3 & 3 & 1 \end{array}\right). \tag{2.28}$$

Come fatto per il sistema precedente, trasformiamo tale matrice in una matrice a gradini mediante operazioni elementari.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 3 & 3 & 1 \end{array}\right) \xrightarrow[\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 2 & 2 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 2 & 2 & 0 \end{array}\right) \xrightarrow[R_3 \to R_3 + R_2]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 0 & 0 & -1 \end{array}\right)$$

Notiamo che la terza riga della matrice trasformata corrisponde all'equazione $0x_1 + 0x_2 + 0x_3 = -1$, ovvero $0 = -1$: poichè questa uguaglianza è falsa, non esiste nessuna terna che soddisfi le tre condizioni del sistema ridotto corrispondente, ovvero tale sistema non ha soluzioni. Questo, in virtù dell'equivalenza tra il sistema originale e quello ridotto, ci dice che il sistema di partenza non ha soluzioni, ovvero è incompatibile.
Evidentemente tra le equazioni del sistema di partenza vi era una incompatibilità non evidente che il procedimento di riduzione a gradini ha fatto emergere: infatti, se moltiplichiamo membro a membro la prima equazione per 2 e le sottraiamo la seconda equazione otteniamo

$$2(x_1 + x_2 + x_3) - (x_1 - x_2 - x_3) = 2 \cdot 1 - 0$$

ovvero, svolgendo i calcoli,

$$x_1 + 3x_2 + 3x_3 = 2$$

Questa condizione, che è conseguenza delle prime due equazioni ed è quindi soddisfatta da qualunque terna le soddisfi, è chiaramente incompatibile con la terza equazione: il procedimento di riduzione a gradini ha messo alla luce questa incompatibilitá trasformandola nell'incompatibilità evidente $0 = -1$.

Consideriamo ora come ultimo esempio il sistema

$$\begin{cases} x_1 + x_2 + 3x_3 = 1 \\ x_1 - 2x_2 + x_3 = 0 \\ x_1 - 5x_2 - x_3 = -1 \end{cases} \tag{2.29}$$

che ha come matrice completa

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & -5 & -1 & -1 \end{array}\right). \tag{2.30}$$

Applicando operazioni elementari per ridurre a gradini,

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & -5 & -1 & -1 \end{array}\right) \xrightarrow[\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}]{} \left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & -6 & -4 & -2 \end{array}\right) \tag{2.31}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & -6 & -4 & -2 \end{array}\right) \xrightarrow[R_3 \to R_3 - 2R_2]{} \left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{2.32}$$

Notiamo che la terza riga della matrice trasformata corrisponde all'equazione $0x_1 + 0x_2 + 0x_3 = 0$, ovvero $0 = 0$.
Quest'ultima condizione è un'identità vera indipendentemente dal valore che diamo alle incognite, quindi essa può essere cancellata dal sistema senza influire sulle sue soluzioni. In altre parole, il sistema iniziale di tre equazioni si è trasformato nel sistema di due equazioni equivalente

$$\begin{cases} x_1 + x_2 + 3x_3 = 1 \\ -3x_2 - 2x_3 = -1 \end{cases} \tag{2.33}$$

Benché non sia rimasta un'equazione con una sola incognita come nel primo sistema che abbiamo risolto, possiamo comunque procedere nel modo seguente: ricaviamo x_2 dalla seconda equazione:

$$-3x_2 - 2x_3 = -1 \to -3x_2 = 2x_3 - 1 \to x_2 = -\frac{2}{3}x_3 + \frac{1}{3} \tag{2.34}$$

e sostituiamo l'espressione ottenuta nella prima equazione per ricavare x_1:

$$x_1 + x_2 + 3x_3 = 1 \to x_1 = 1 - x_2 - 3x_3 = 1 - \left(-\frac{2}{3}x_3 + \frac{1}{3}\right) - 3x_3 = \frac{2}{3} - \frac{7}{3}x_3. \tag{2.35}$$

Ora, qualunque valore $t \in \mathbb{R}$ assegniamo a x_3, la (2.34) e la (2.35) ci dicono che se poniamo $x_2 = -\frac{2}{3}t + \frac{1}{3}$ e $x_1 = \frac{2}{3} - \frac{7}{3}t$, le equazioni del sistema saranno soddisfatte, ovvero otterremo una soluzione.
In altre parole, le soluzioni del sistema sono esattamente tutte le terne del tipo $(\frac{2}{3} - \frac{7}{3}t, -\frac{2}{3}t + \frac{1}{3}, t)$ al variare di $t \in \mathbb{R}$: il sistema ha quindi infinite soluzioni.

Più precisamente, dal momento che le infinite soluzioni del sistema dipendono da un solo parametro libero t, si dice che il sistema ha "infinito alla uno" (si scrive ∞^1) soluzioni.

In generale, possiamo dare la seguente

Definizione 2.8. Un sistema di equazioni lineari ha ∞^k *soluzioni* se l'espressione generale della sua soluzione dipende da k parametri liberi.

Nell'ultimo esempio visto, la riduzione ha eliminato una delle tre equazioni del sistema riducendola all'identità $0 = 0$. In effetti, non è difficile vedere che la terza equazione $x_1 - 5x_2 - x_3 = -1$ era un'equazione "superflua", o più precisamente dipendente dalle altre due: come si vede nella matrice completa (2.30), la terza riga, che la rappresenta, è combinazione delle altre due:

$$(1, -5, -1, -1) = -(1, 1, 3, 1) + 2(1, -2, 1, 0)$$

In effetti, non è difficile vedere che una matrice a gradini, escluse le righe nulle, non ha più righe dipendenti (e quindi le equazioni non nulle di un sistema ridotto a gradini sono sicuramente indipendenti):

Proposizione 2.9. Le righe non nulle di una matrice ridotta a gradini sono linearmente indipendenti

Proof. Per definizione di matrice a gradini le sue righe saranno del tipo

$$\begin{gathered} R_1 = (a_{11}, \dots), \quad a_{11} \neq 0 \\ R_2 = (0, \dots, 0, a_{2k}, \dots,), \quad a_{2k} \neq 0 \\ R_3 = (0, \dots, 0, 0, \dots, 0, a_{3j}, \dots,), \quad a_{3j} \neq 0 \\ \vdots \end{gathered}$$

con $k > 1$, $j > k$ etc. ovvero in ogni riga il primo elemento non nullo compare via via con secondo indice sempre più grande.
Ora, come sappiamo dall'Osservazione 1.20, per dimostrare che tali righe sono indipendenti basta supporre che

$$c_1(a_{11}, \dots) + c_2(0, \dots, 0, a_{2k}, \dots,) + c_3(0, \dots, 0, 0, \dots, 0, a_{3j}, \dots,) + \cdots = (0, \dots, 0) \tag{2.36}$$

e dimostrare che i coefficienti c_1, c_2, c_3 etc. devono essere necessariamente tutti nulli.
Andiamo a guardare cosa significa l'uguaglianza (2.36) entrata per entrata: poichè la prima riga è l'unica con prima entrata diversa da zero, facendo i calcoli a primo membro della (2.36) vediamo che nella prima entrata rimane solo $c_1 a_{11} = 0$: ma, essendo per ipotesi $a_{11} \neq 0$, necessariamente deve essere $c_1 = 0$. Quindi, la (2.36) si riduce a

$$c_2(0, \dots, 0, a_{2k}, \dots,) + c_3(0, \dots, 0, 0, \dots, 0, a_{3j}, \dots,) + \cdots = (0, \dots, 0) \tag{2.37}$$

Ora, guardiamo la k-esima entrata di questa relazione (cioè la prima diversa da zero nella seconda riga): dal momento che tutte le righe successive alle seconda hanno la prima entrata diversa da zero con indice più alto, otteniamo $c_2 a_{2k} = 0$, che, essendo $a_{2k} \neq 0$, ci dice che $c_2 = 0$.
Dunque la (2.37) si riduce a

$$c_3(0, \dots, 0, 0, \dots, 0, a_{3j}, \dots,) + \dots = (0, \dots, 0)$$

e, continuando a ragionare in questo modo, vedremo che tutti i coefficienti c_i si devono annullare, e quindi non può esistere una combinazione lineare delle righe uguale al vettore nullo e con coefficienti non tutti nulli, ovvero le righe sono indipendenti. □

Osservazione 2.10. Quando effettuiamo operazioni elementari sulle righe di una matrice, siamo autorizzati anche a considerare trasformazioni del tipo $R_i \to cR_i + dR_j$, purché il coefficiente c per cui moltiplichiamo la riga R_i che stiamo sostituendo non sia zero: infatti, benché tale trasformazione non sia una delle tre operazioni elementari, essa può essere pensata come il risultato dell'applicazione di due operazioni elementari una dopo l'altra: prima applicando alla riga R_i l'operazione elementare del secondo tipo $R_i \to cR_i$ (con $c \neq 0$ come previsto) e poi applicando alla nuova riga cR_i così ottenuta l'operazione elementare del terzo tipo $cR_i \to cR_i + dR_j$.

Il procedimento di riduzione a gradini, che noi abbiamo usato come strumento di risoluzione di un sistema, in realtà è sostanzialmente un metodo che stabilisce se dei vettori sono indipendenti. Infatti, più precisamente, noi abbiamo visto i seguenti due fatti:

(1) Il procedimento non modifica lo spazio generato dalle righe, ovvero se $R_1, R_2, \dots, R_m$ sono le righe della matrice iniziale, e $\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_m$ sono le righe della matrice trasformata, allora $\langle R_1, R_2, \dots, R_m \rangle = \langle \tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_m \rangle$;

(2) le righe non nulle alla fine del procedimento formano un insieme di vettori indipendenti

(nel contesto della risoluzione dei sistemi, il primo fatto è stato usato per garantire che il sistema trasformato è equivalente a quello iniziale; il secondo fatto giustifica l'affermazione che alla fine del procedimento non vi sono più equazioni "superflue").
Quindi, se le righe non nulle dopo la riduzione sono le prime l, ovvero $\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_l$, dal fatto che queste sono indipendenti deduciamo che costituiscono una base del sottospazio da loro generato, e quindi $\dim\langle \tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_l \rangle =$

l; ma poiché il sottospazio generato non cambia, concludiamo che $\dim\langle R_1, R_2, \ldots, R_m\rangle = l$.
Il numero di righe non nulle rimaste dopo la riduzione a gradini ci dà quindi la dimensione dello spazio generato dalle righe iniziali: dal momento che, come abbiamo visto a pagina 45, in uno spazio di dimensione l ci sono al massimo l vettori indipendenti, deduciamo che tra le righe $R_1, R_2, \ldots, R_m$ ve n'erano l indipendenti, e le restanti scrivibili come combinazione di quelle l.
In poche parole, il numero di righe non nulle rimaste dopo la riduzione ci dice quante righe indipendenti aveva la matrice prima della riduzione.
Questa informazione, importantissima, giustifica la seguente

Definizione 2.11. Il massimo numero di righe indipendenti di una matrice A si chiama il *rango per righe* di A.

Quindi la riduzione a gradini ci dà un modo per calcolare la dimensione di uno spazio e per verificare se delle n-uple date siano indipendenti (che rappresentino equazioni di un sistema o meno).
Ad esempio, consideriamo le seguenti 4-uple

$$(1,1,2,1),\ (1,2,1,0),\ (1,-1,4,3),\ (2,1,1,0) \tag{2.38}$$

e supponiamo di voler determinare se esse sono indipendenti. Disponiamo le 4-uple a formare le righe di una matrice

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

e riduciamola a gradini seguendo il procedimento di eliminazione:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & -1 & 4 & 3 \\ 2 & 1 & 1 & 0 \end{pmatrix} \underset{\substack{R_2\to R_2-R_1 \\ R_3\to R_3-R_1 \\ R_4\to R_4-2R_1}}{\longrightarrow} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & -2 & 2 & 2 \\ 0 & -1 & -3 & -2 \end{pmatrix} \underset{\substack{R_3\to R_3+2R_2 \\ R_4\to R_4+R_2}}{\longrightarrow} \tag{2.39}$$
$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & -3 \end{pmatrix} \underset{R_4\leftrightarrow R_3}{\longrightarrow} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -4 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(l'ultimo scambio di righe è stato necessario per portare la matrice nella forma a gradini).

Vediamo quindi che dopo la riduzione sono rimaste 3 righe non nulle, ovvero il rango della matrice è 3: nell'insieme iniziale vi erano allora 3 righe indipendenti e una quarta dipendente dalle altre (quindi le 4-uple non erano linearmente indipendenti).
In particolare, possiamo anche affermare che il vettore da escludere se vogliamo estrarre un insieme di vettori indipendenti dai quattro vettori dati era il terzo, *corrispondente alla riga annullata dalla riduzione.* Infatti, per arrivare all'annullamento di tale riga abbiamo dapprima eseguito $R_2 \to R_2 - R_1$ e $R_3 \to R_3 - R_1$, e poi abbiamo sommato alla (nuova) terza riga $R_3 - R_1$ la (nuova) seconda riga $R_2 - R_1$ moltiplicata per 2, ovvero

$$(R_3 - R_1) + 2(R_2 - R_1) = 0$$

Svolgendo i conti, questa uguaglianza ci dice che $R_3 - 3R_1 + 2R_2 = 0$, ovvero

$$R_3 = 3R_1 - 2R_2$$

che ci conferma che la terza riga è scrivibile come combinazione delle altre.
Il fatto appena illustrato in questo esempio è vero in generale: se una riga si annulla in seguito all'algoritmo di riduzione a gradini allora essa era esprimibile come combinazione delle altre. Infatti, nel corso della riduzione a gradini trasformiamo dapprima tutte le righe dalla seconda in poi combinandole con la prima (eventualmente dopo uno scambio di righe)

$$R_2 \to R_2 + c_2R_1,\ R_3 \to R_3 + c_3R_1, \dots, R_m \to R_m + c_mR_1$$

Nel secondo passaggio, ogni riga così trasformata viene combinata con la seconda riga (anche lei trasformata):

$$R_k \to (R_k + c_kR_1) + c'_k(R_2 + c_2R_1) = R_k + (c_k + c'_kc_2)R_1 + c'_kR_2$$

e così via: a ogni passaggio la riga R_k viene combinata con una in più delle righe precedenti, fino a che o non la tocchiamo più perché iniziamo a usarla per ridurre le successive, oppure essa si annulla: in tal caso arriviamo quindi a una relazione del tipo

$$R_k + d_1R_1 + d_2R_2 + \cdots + d_jR_j = 0$$

ovvero $R_k = -d_1R_1 - d_2R_2 - \cdots - d_jR_j$, che ci dice che la riga R_k che si è annullata era combinazione lineare delle altre righe.

Osservazione 2.12. Si presti attenzione al fatto che l'affermazione secondo cui le righe che si annullano sono combinazione lineare delle altre è vera quando si segue l'algoritmo di riduzione a gradini, ma in generale se una riga si annulla dopo una serie qualunque di operazioni elementari non è detto che sia combinazione lineare delle altre. Ad esempio, si consideri la seguente sequenza di operazioni elementari (che non segue i passi dell'algoritmo di riduzione a gradini)

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \underset{R_3 \to R_3 + R_1}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} \underset{R_2 \to R_2 + R_3}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 3 & 1 \\ 1 & 1 \end{pmatrix} \underset{R_3 \to R_3 - R_2}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 3 & 1 \\ -2 & 0 \end{pmatrix} \underset{R_3 \to R_3 + 2R_1}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 3 & 1 \\ 0 & 0 \end{pmatrix}$$

a seguito della quale la terza riga si annulla, pur non essendo combinazione lineare delle altre: non c'è nessun modo di esprimere $(0, 1)$ come combinazione di $(1, 0)$ e $(2, 0)$ (si noti che non sono stati fatti scambi di riga).

La Definizione 2.11 suggerisce che si può definire anche il *rango per colonne* di una matrice come il numero massimo di colonne linearmente indipendenti (ovvero la dimensione dello spazio generato dalle colonne). Tuttavia, vale la seguente importante

Proposizione 2.13. Per una qualunque matrice, il rango per righe coincide con il rango per colonne.

Alla luce di questo risultato il rango per righe (o per colonne) si chiama allora semplicemente *rango* della matrice.

Non dimostriamo la Proposizione 2.13, ma illustriamola con un semplice esempio: nella matrice $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{pmatrix}$ la seconda riga R_2 è evidentemente dipendente dalle altre, in quanto $R_2 = 2R_1$ (se volessimo far apparire anche la terza riga in questa relazione di dipendenza, potremmo equivalentemente scrivere $R_2 = 2R_1 + 0R_3$).

Per il risultato appena citato, allora anche una delle colonne della matrice deve essere dipendente dalle altre: in effetti, si ha $\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, che era molto meno evidente della relazione di dipendenza esistente tra le righe.

Come ulteriore esempio, si considerino le stesse 4-uple viste sopra in (2.38): grazie all'uguaglianza del rango per righe e per colonne, avremmo potuto anche disporre i vettori in colonna anzichè in riga:

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

Il rango per righe di questa matrice, che ci apprestiamo a calcolare con il procedimento di riduzione a gradini, è quindi uguale al rango per colonne della precedente, ovvero deve sempre essere uguale a 3. Infatti

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 2 & 1 & 4 & 1 \\ 1 & 0 & 3 & 0 \end{pmatrix} \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1}}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & -1 & 2 & -3 \\ 0 & -1 & 2 & -2 \end{pmatrix} \underset{\substack{R_3 \to R_3 + R_2 \\ R_4 \to R_4 + R_2}}{\longrightarrow} \tag{2.40}$$

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & -3 \end{pmatrix} \underset{R_4 \leftrightarrow 4R_4 - 3R_3}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Come previsto dall'uguaglianza del rango per righe o per colonne, abbiamo nuovamente ottenuto che il rango della matrice è 3.
Si noti che, in questo caso, a dirci quale vettore è combinazione degli altri (e quindi quale escludere nel caso volessimo estrarre una base dai vettori visti come insieme di generatori di un sottospazio) è la posizione dei pivot nella matrice ridotta: poiché questi si trovano in prima, seconda e quarta colonna, i vettori da tenere sono il primo, il secondo e il quarto, mentre il terzo è da escludere (come già sapevamo dalla riduzione fatta sui vettori disposti in riga).
Infatti, se seguiamo la riduzione guardando solo le prime due colonne, dove si trovano i primi due pivot, vediamo che il rango della matrice è 2, il che ci dice che i primi due vettori sono indipendenti tra loro; se guardiamo cosa succede solo alle prime tre colonne, vediamo che il rango è ancora 2 (in quanto la terza colonna non contiene pivot) e questo ci dice che il terzo vettore era allora combinazione dei primi due: è solo aggiungendo la quarta colonna, dove si trova il terzo pivot, che otteniamo una matrice di rango 3, il che significa che è il quarto vettore, contrariamente al terzo, ad essere indipendente dai primi due.

Facendo uso della nozione di rango, possiamo riassumere tutto quello che sappiamo sulla risolubilità di un sistema e sul numero delle sue soluzioni nel seguente risultato, detto *teorema di Rouché-Capelli*.

Teorema 2.14. Un sistema di m equazioni lineari in n incognite è compatibile se e solo se il rango della matrice dei coefficienti è uguale al rango della matrice completa, e in tal caso il sistema ha ∞^{n-r} soluzioni (dove r denota il rango della matrice). In particolare, il sistema ha un'unica soluzione se e solo se $r = n$.

Proof. Come sappiamo, un sistema è incompatibile se e solo se in seguito alla riduzione a gradini compaiono righe del tipo $\left(\begin{array}{ccc|c} 0 & \dots & 0 & b \end{array} \right)$ con $b \neq 0$. Ma questo equivale a dire che nella matrice dei coefficienti si è annullata una riga in più che nella matrice completa, ovvero il rango (che, ricordiamo, è il numero di righe non nulle dopo la riduzione) della matrice dei coefficienti è diverso (in particolare, minore) del rango della matrice completa. Questo dimostra la prima affermazione del teorema.

Per quello che riguarda la seconda affermazione, sappiamo che una volta ridotto il sistema si ricava la prima incognita che compare in ogni equazione non nulla in funzione delle rimanenti. Se il rango della matrice è r, ci sono proprio r righe non nulle e quindi si ricavano r incognite in funzione delle rimanenti, che sono $n - r$ e fungono da parametri liberi. Quindi la soluzione generale si scrive in funzione di $n - r$ parametri, ovvero il sistema ha ∞^{n-r} soluzioni.

L'ultima affermazione del teorema discende dal fatto che la soluzione è unica quando non dipende da nessun parametro libero, ovvero $n - r = 0$. □

Osservazione 2.15. Si noti che il teorema afferma che hanno un'unica soluzione i sistemi (compatibili) in cui il numero di incognite è uguale al numero di equazioni a patto che queste ultime siano indipendenti (si veda anche il semplice esempio del sistema (2.4) che ha due equazioni, due incognite ma infinite soluzioni).

Prima di vedere alcune applicazioni geometriche di tutta la teoria dei sistemi e della riduzione vista, concludiamo questa parte con due importanti risultati che mostrano come i sistemi omogenei (ovvero in cui tutti i termini noti sono uguali a zero, si veda l'Osservazione 2.3) hanno delle importanti caratteristiche che li distinguono dai sistemi in generale:

Proposizione 2.16. Dato un sistema omogeneo di m equazioni in n incognite a coefficienti in un campo $\mathbb{K}$, valgono le seguenti:

(i) se $s = (s_1, \dots, s_n)$ e $s' = (s'_1, \dots, s'_n)$ sono soluzioni del sistema, lo è anche $s + s' = (s_1 + s'_1, \dots, s_n + s'_n)$

(ii) se $s = (s_1, \dots, s_n)$ è una soluzione del sistema e $c \in \mathbb{K}$, allora lo è anche $cs = (cs_1, \dots, cs_n)$

In altre parole, l'insieme delle soluzioni di un sistema omogeneo in n incognite è un sottospazio vettoriale di $\mathbb{K}^n$.

Proof. Sia $a_{j1}x_1 + \cdots + a_{jn}x_n = 0$ la generica equazione del sistema. Il fatto che $s = (s_1, \ldots, s_n)$ e $s' = (s'_1, \ldots, s'_n)$ sono soluzioni del sistema significa che $a_{j1}s_1 + \cdots + a_{jn}s_n = 0$ e $a_{j1}s'_1 + \cdots + a_{jn}s'_n = 0$. Ma allora

$$a_{j1}(s_1 + s'_1) + \cdots + a_{jn}(s_n + s'_n) = a_{j1}s_1 + a_{j1}s'_1 + \cdots + a_{jn}s_n + a_{jn}s'_n =$$

$$= a_{j1}s_1 + \cdots + a_{jn}s_n + a_{j1}s'_1 + \cdots + a_{jn}s'_n = 0 + 0 = 0 \tag{2.41}$$

ovvero anche $s + s' = (s_1 + s'_1, \ldots, s_n + s'_n)$ è soluzione: questo mostra (i).

Per mostrare (ii), supponiamo sempre che $s = (s_1, \ldots, s_n)$ sia soluzione del sistema, ovvero $a_{j1}s_1 + \cdots + a_{jn}s_n = 0$ per la generica equazione, e osserviamo che allora

$$a_{j1}cs_1 + \cdots + a_{jn}cs_n = c(a_{j1}s_1 + \cdots + a_{jn}s_n) = c \cdot 0 = 0 \tag{2.42}$$

ovvero anche $cs = (cs_1, \ldots, cs_n)$ è soluzione, come affermato in (ii). □

Osserviamo che, ovviamente, nel caso di sistemi non omogenei, ovvero quelli che hanno almeno un'equazione $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$ con $b_j \neq 0$, i passaggi visti sopra non sono più veri: ad esempio, il calcolo in (2.41) diventerebbe

$$a_{j1}(s_1 + s'_1) + \cdots + a_{jn}(s_n + s'_n) = a_{j1}s_1 + a_{j1}s'_1 + \cdots + a_{jn}s_n + a_{jn}s'_n =$$

$$= a_{j1}s_1 + \cdots + a_{jn}s_n + a_{j1}s'_1 + \cdots + a_{jn}s'_n = b_j + b_j = 2b_j$$

e quindi $s+s' = (s_1+s'_1, \ldots, s_n+s'_n)$ non risolve più l'equazione $a_{j1}x_1+\cdots+a_{jn}x_n = b_j$ ma l'equazione $a_{j1}x_1 + \cdots + a_{jn}x_n = 2b_j$ (chiaramente diversa se $b_j \neq 0$).
Analogamente, il calcolo in (2.42) diventerebbe

$$a_{j1}cs_1 + \cdots + a_{jn}cs_n = c(a_{j1}s_1 + \cdots + a_{jn}s_n) = cb_j$$

e, di nuovo, vedremmo che $cs = (cs_1, \ldots, cs_n)$ non risolve più l'equazione $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$ ma l'equazione $a_{j1}x_1 + \cdots + a_{jn}x_n = cb_j$ (di nuovo, chiaramente diversa se $b_j \neq 0$).
Quindi per i sistemi non omogenei non possiamo più dire che l'insieme delle soluzioni è un sottospazio vettoriale; tuttavia, vale il seguente risultato che descrive comunque la struttura dell'insieme delle sue soluzioni:

Proposizione 2.17. Dato un sistema non omogeneo di m equazioni in n incognite a coefficienti in un campo $\mathbb{K}$, l'insieme delle sue soluzioni si ottiene sommando a una sua soluzione particolare le soluzioni del *sistema omogeneo associato* (ovvero il sistema che si ottiene ponendo tutti i termini noti uguali a zero).

Proof. Sia $s = (s_1, \dots, s_n)$ una soluzione particolare del sistema non omogeneo e $(w_1, \dots, w_n)$ una soluzione del sistema omogeneo associato. Questo significa che $a_{j1}s_1 + \dots + a_{jn}s_n = b_j$ e che $a_{j1}w_1 + \dots + a_{jn}w_n = 0$. Ma allora

$$a_{j1}(s_1 + w_1) + \dots + a_{jn}(s_n + w_n) = a_{j1}s_1 + a_{j1}w_1 + \dots + a_{jn}s_n + a_{jn}w_n =$$

$$= a_{j1}s_1 + \dots + a_{jn}s_n + a_{j1}w_1 + \dots + a_{jn}w_n = b_j + 0 = b_j$$

ovvero anche $s + w = (s_1 + w_1, \dots, s_n + w_n)$ è soluzione del sistema non omogeneo: questo mostra che sicuramente sommando a una soluzione particolare del sistema una soluzione del suo sistema omogeneo associato si ottiene ancora una soluzione del sistema (non omogeneo): per concludere la dimostrazione, dobbiamo mostrare che in questo modo si ottengono *tutte* le soluzioni del sistema non omogeneo.
A questo scopo consideriamo un'altra qualunque soluzione $s' = (s'_1, \dots, s'_n)$ del sistema non omogeneo: allora, dal momento che si ha sia $a_{j1}s_1 + \dots + a_{jn}s_n = b_j$ che $a_{j1}s'_1 + \dots + a_{jn}s'_n = b_j$, si ha

$$a_{j1}(s'_1 - s_1) + \dots + a_{jn}(s'_n - s_n) = a_{j1}s'_1 - a_{j1}s_1 + \dots + a_{jn}s'_n - a_{jn}s_n =$$

$$= a_{j1}s'_1 + \dots + a_{jn}s'_n - (a_{j1}s_1 + \dots + a_{jn}s_n) = b_j - b_j = 0$$

Questo mostra che $s' - s = (s'_1 - s_1, \dots, s'_n - s_n)$ è una soluzione del sistema omogeneo associato: poiché chiaramente $s' = s + (s' - s)$, questo ci dice che qualunque soluzione s' del sistema (non omogeneo) si può scrivere come somma della soluzione particolare s fissata fin dall'inizio più una soluzione del sistema omogeneo associato, come volevamo. □

Esempio 2.18. Si consideri il sistema omogeneo seguente:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ 2x_1 + x_2 + x_3 - x_4 = 0 \end{cases} \tag{2.43}$$

Con un solo passaggio

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & -1 \end{pmatrix} \underset{R_2 \to R_2 - 2R_1}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -3 \end{pmatrix}$$

riduciamo a gradini la sua matrice[6] e lo riduciamo al sistema a gradini equivalente $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ -x_2 - x_3 - 3x_4 = 0 \end{cases}$
che possiamo risolvere dal basso: posto $x_3 = t$ e $x_4 = s$, la seconda equazione ci dà $x_2 = -t-3s$, e sostituendo nella prima otteniamo $x_1 = -x_2-x_3-x_4 = t + 3s - t - s = 2s$.
Le soluzioni del sistema sono quindi tutte e sole le 4-uple del tipo $(2s, -t - 3s, t, s)$, al variare di $t, s \in \mathbb{R}$.
Che l'insieme di tali 4-uple costituisca un sottospazio vettoriale (di $\mathbb{R}^4$) si può mettere in evidenza riscrivendo la soluzione generale nella forma seguente:

$$(2s, -t - 3s, t, s) = t(0, -1, 1, 0) + s(2, -3, 0, 1)$$

uguaglianza dalla quale si vede che la generica soluzione è combinazione lineare dei vettori $(0, -1, 1, 0)$ e $(2, -3, 0, 1)$: in altre parole, l'insieme delle soluzioni coincide con il sottospazio $\langle(0, -1, 1, 0), (2, -3, 0, 1)\rangle$ generato da tali vettori.

Per un esempio di sistema non omogeneo, si consideri ad esempio il sistema

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 3 \\ 2x_1 + x_2 + x_3 - x_4 = 5 \end{cases}$$

che ha come sistema omogeneo associato proprio il (2.43) appena risolto.
Anche qui, con un solo passaggio

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & -1 & 5 \end{pmatrix} \underset{R_2 \to R_2 - 2R_1}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 & 3 \\ 0 & -1 & -1 & -3 & -1 \end{pmatrix}$$

riduciamo a gradini la sua matrice completa trasformandolo nel sistema a gradini equivalente $\begin{cases} x_1 + x_2 + x_3 + x_4 = 3 \\ -x_2 - x_3 - 3x_4 = -1 \end{cases}$
che possiamo risolvere dal basso: posto $x_3 = t$ e $x_4 = s$, la seconda equazione ci dà $x_2 = 1-t-3s$, e sostituendo nella prima otteniamo $x_1 = 3-x_2-x_3-x_4 = 3 - 1 + t + 3s - t - s = 2 + 2s$.
Le soluzioni del sistema sono quindi tutte e sole le 4-uple del tipo $(2+2s, 1 - t - 3s, t, s)$, al variare di $t, s \in \mathbb{R}$.
Allora, si vede che tale soluzione generale può essere decomposta come

$$(2 + 2s, 1 - t - 3s, t, s) = (2, 1, 0, 0) + (2s, -t - 3s, t, s)$$

[6] Si noti che non abbiamo riportato i termini noti: infatti, essendo questi tutti uguali a zero, qualunque operazione elementare applichiamo alle righe rimarranno uguali a zero, quindi non è necessario scriverli.

ovvero, come previsto dalla Proposizione 2.17, come somma della soluzione particolare $(2, 1, 0, 0)$ (quella che si ottiene ponendo $t = s = 0$) più le 4-uple del tipo $(2s, -t - 3s, t, s)$, che sono proprio le soluzioni del suo sistema omogeneo associato (2.43).

Osservazione 2.19. I teoremi appena visti hanno un analogo con quello che succede nel caso delle equazioni differenziali lineari, ovvero le equazioni del tipo

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = f(t)$$

Infatti, anche per tali equazioni si dimostra che la soluzione generale si trova sommando una soluzione particolare a una soluzione dell'equazione omogenea associata $y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = 0$, il cui insieme delle soluzioni è, esattamente come accade per i sistemi omogenei, un sottospazio vettoriale (in questo caso dello spazio delle funzioni).

Ad esempio, l'equazione $y'' + y = t^2$ ha come soluzione generale $y(t) = c_1 \cos t + c_2 \sin t + t^2 - 2$, dove è facile vedere che $t^2 - 2$ è una sua soluzione particolare, mentre $c_1 \cos t + c_2 \sin t$ è, al variare di $c_1, c_2 \in \mathbb{R}$, la soluzione generale dell'equazione omogenea associata $y'' + y = 0$ (quindi, l'insieme delle soluzioni di tale equazione omogenea è il sottospazio vettoriale generato dalle funzioni $\cos t$ e $\sin t$).

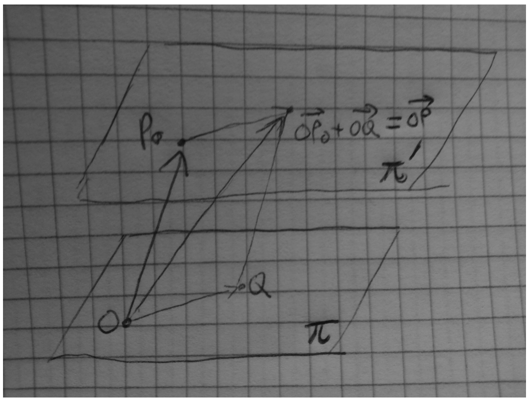
Esattamente come nel caso dei sistemi, quindi, riusciamo a esprimere tutte le (infinite) soluzioni di un'equazione semplicemente usando una soluzione particolare e due soluzioni (indipendenti) dell'equazione omogenea.

In generale, in uno spazio vettoriale V, un sottoinsieme $S \subseteq V$ ottenuto sommando un vettore fissato v^0 a tutti i vettori di un sottospazio vettoriale W dato (scriveremo $S = v^0 + W$) si chiama anche *sottospazio affine*.

Quindi, abbiamo dimostrato che l'insieme delle soluzioni di un sistema non omogeneo compatibile di m equazioni in n incognite e coefficienti in $\mathbb{K}$ è un sottospazio affine di $\mathbb{K}^n$.

Un altro esempio di sottospazio affine, che abbiamo già incontrato, è quello dato dall'insieme S dei vettori $\vec{OP} \in V_O^3$ il cui secondo estremo P appartiene a un piano π' fissato *non* passante per O (nell'Esempio 1.27 abbiamo visto che S *non* è un sottospazio vettoriale di V_O^3).

Infatti, osserviamo (si veda il disegno seguente) che, fissato un vettore $\vec{OP_0} \in S$, qualunque altro $\vec{OP} \in S$ si può scrivere come somma di $\vec{OP_0}$ e di un vettore $\vec{OQ}$ che giace sul piano π parallelo a π' e passante per O, ovvero $S = \vec{OP_0} + W$, dove W è l'insieme dei vettori che giacciono su π (che, come sappiamo, è un sottospazio vettoriale).



Con un ragionamento e un disegno analogo, si può vedere che l'insieme dei vettori $\vec{OP}$ il cui secondo estremo P sta su una retta non passante per O è un sottospazio affine.
Ispirati da questi esempi, si dice spesso che un sottospazio affine è *il traslato di un sottospazio vettoriale*.

2.3 Qualche applicazione geometrica

Vediamo ora alcuni problemi geometrici che possono essere risolti grazie al procedimento di riduzione a gradini e al concetto di rango.

Ad esempio, consideriamo le due rette r e r' di equazioni cartesiane

$$r : \begin{cases} x + y + z = 2 \\ 2x - y + z = 5 \end{cases} , \quad r' : \begin{cases} x - y - 2z = -2 \\ x + 3y + 2z = 2 \end{cases} \tag{2.44}$$

e supponiamo di voler determinare se esse hanno punti in comune.
Dal momento che i punti di una retta espressa in equazioni cartesiane sono proprio le soluzioni del sistema formato dalle due equazioni, i punti comuni alle due rette sono dati dalle soluzioni comuni a tutte e quattro le equazioni delle due rette, ovvero le soluzioni del sistema

$$r : \begin{cases} x + y + z = 2 \\ 2x - y + z = 5 \\ x - y - 2z = -2 \\ x + 3y + 2z = 2 \end{cases} \tag{2.45}$$

Riducendo la matrice completa otteniamo

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 5 \\ 1 & -1 & -2 & -2 \\ 1 & 3 & 2 & 2 \end{array}\right) \xrightarrow[\substack{R_2\to R_2-2R_1 \\ R_3\to R_3-R_1 \\ R_4\to R_4-R_1}]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & -2 & -3 & -4 \\ 0 & 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[\substack{R_3\to 3R_3-2R_2 \\ R_4\to 3R_4+2R_2}]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & 0 & -7 & -14 \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow[R_4\to 7R_4+R_3]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & 0 & -7 & -14 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Vediamo quindi che il sistema è compatibile e, essendosi annullata una riga, la matrice ha rango 3, quindi avendo 3 incognite in base a quanto detto nel Teorema 2.14 abbiamo una sola soluzione, che troviamo risolvendo il sistema ridotto corrispondente

$$\begin{cases} x+y+z=2 \\ -3y-z=1 \\ -7z=-14 \end{cases}$$

Dall'ultima equazione si ottiene $z=2$, che sostituito nella seconda dà

$$-3y = z+1 = 2+1 = 3$$

ovvero $y=-1$. Sostituendo nella prima equazione per ottenere x si ha

$$x = 2-y-z = 2-(-1)-2 = 1$$

Quindi l'unica equazione del sistema è data dalla terna $(1,-1,2)$, che sono le coordinate del punto in cui si incontrano le due rette.

Supponiamo invece che le rette siano

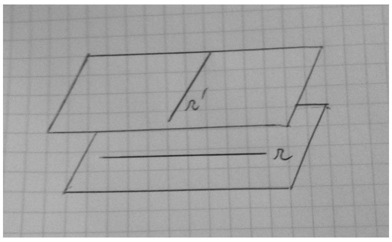
$$r: \begin{cases} x+y+z=0 \\ 2x+y-z=1 \end{cases}, \quad r': \begin{cases} 2x-y=3 \\ x+y-z=1 \end{cases} \tag{2.46}$$

e supponiamo ancora di voler determinare se esse hanno punti in comune. Come sopra, mettiamo insieme le quattro equazioni e riduciamo la matrice completa del sistema così ottenuto:

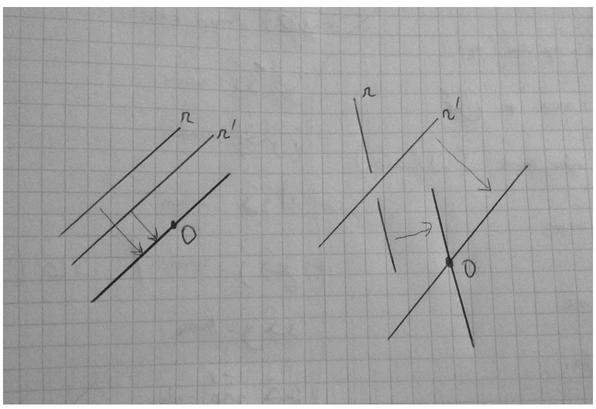
$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 1 \\ 2 & -1 & 0 & 3 \\ 1 & 1 & -1 & 1 \end{array}\right) \xrightarrow[\substack{R_2\to R_2-2R_1 \\ R_3\to R_3-2R_1 \\ R_4\to R_4-R_1}]{} \tag{2.47}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & -3 & -2 & 3 \\ 0 & 0 & -2 & 1 \end{array}\right) \underset{R_3 \to R_3 - 3R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right) \underset{R_4 \to 7R_4 + 2R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{array}\right)$$

Essendo il sistema incompatibile (l'ultima riga corrisponde all'uguaglianza falsa $0 = 7$) deduciamo che le due rette non hanno punti in comune.
Ora, mentre nel piano due rette che non hanno punti in comune sono necessariamente parallele, nello spazio tridimensionale questo non è più vero: come si vede nel seguente disegno, due rette nello spazio, grazie alla dimensione in più presente rispetto al piano, possono trovarsi su piani paralleli e quindi non incontrarsi pur non avendo la stessa direzione



In tal caso si dice che le rette sono *sghembe*.
Vediamo ora come sia possibile determinare se le rette date sono sghembe o parallele senza fare ulteriori conti, ma sfruttando la riduzione già fatta.
Infatti, due rette che non hanno punti in comune sono parallele se e solo se quando le trasliamo parallelamente a se stesse sull'origine esse risultano coincidere (ovvero hanno infiniti punti in comune), mentre sono sghembe se e solo se quando le trasliamo parallelamente a se stesse sull'origine esse hanno in comune un solo punto, l'origine stessa:



Possiamo così tradurre il problema di capire se le due rette hanno la stessa direzione nel problema di determinare un'intersezione (tra le rette traslate). Ora, per traslare una retta espressa in equazioni cartesiane, parallelamente a se stessa, basta modificare i termini noti delle equazioni lasciando invariati i primi membri[7]: in particolare, otteniamo la traslazione sull'origine se poniamo i termini noti uguali a zero (in quanto in tal caso le equazioni risultato soddisfatte dalla terna $x = 0$, $y = 0$, $z = 0$, che sono le coordinate dell'origine, il che significa che la retta traslata è proprio quella che passa per l'origine). Nel caso delle equazioni delle due rette r e r' date da (2.46), le rette traslate sull'origine sono rappresentate dalle equazioni

$$\begin{cases} x + y + z = 0 \\ 2x + y - z = 0 \end{cases}, \quad \begin{cases} 2x - y = 0 \\ x + y - z = 0 \end{cases} \tag{2.48}$$

Per determinare se tali rette traslate hanno infiniti punti in comune o uno solo, ovvero come abbiamo detto sopra se le rette di partenza erano rispettivamente parallele o sghembe, dobbiamo risolvere il sistema che si ottiene mettendo insieme le 4 equazioni, che ha come matrice associata

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right)$$

Ora, per ridurre a gradini questa matrice dovremmo applicare esattamente le stesse operazioni usate per ridurre la (2.47), che differisce da essa solo per il fatto di avere tutti i termini noti uguali a zero: l'unica differenza sarà che i termini noti rimarranno sempre nulli qualunque operazione elementare applichiamo, e quindi, senza dover rifare i conti, sappiamo che arriveremmo alla stessa matrice ridotta ma con l'ultima colonna (quella dei termini noti) tutta nulla, ovvero

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Guardando questa matrice, che rappresenta ora un sistema compatibile con 3 incognite e rango 3, concludiamo che traslando le rette sull'origine avremmo

[7]Infatti, modificando solo il termine noto dell'equazione $ax + by + cz = d$ di un piano si ottiene un piano parallelo in quanto non avremo cambiato la normale al piano, data dalla terna (a, b, c). Poiché una retta è intersezione di due piani, modificando i termini noti delle due equazioni stiamo muovendo parallelamente a se stessi i piani e quindi muovendo parallelamente a se stessa la retta.

una sola soluzione (l'origine stessa) e quindi le rette di partenza non erano parallele.
Per vedere invece cosa succederebbe se le rette fossero parallele, vediamo l'ulteriore esempio seguente:

$$r: \begin{cases} x+y+z=1 \\ x-y+2z=0 \end{cases}, \quad r': \begin{cases} 2x+3z=3 \\ x+3y=-3 \end{cases} \tag{2.49}$$

Allo scopo di controllare se le rette hanno punti in comune, mettiamo insieme le quattro equazioni e riduciamo la matrice completa del sistema così ottenuto:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ 2 & 0 & 3 & 3 \\ 1 & 3 & 0 & -3 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1}}{\longrightarrow} \tag{2.50}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 1 & 1 \\ 0 & 2 & -1 & -4 \end{array}\right) \underset{\substack{R_3 \to R_3 - R_2 \\ R_4 \to R_4 + R_2}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & -5 \end{array}\right) \underset{R_4 \to 2R_4 + 5R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Come si vede, da una parte le rette non hanno punti in comune in quanto la terza riga corrisponde all'uguaglianza falsa $0 = 2$; dall'altra, il sistema formato dalle due rette traslate sull'origine, ovvero con termini noti nulli, avrebbe come matrice ridotta la stessa matrice ottenuta ora ma con ultima colonna di zeri, ovvero

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

che rappresenta la matrice di un sistema ridotto compatibile con 3 incognite e rango 2, quindi infinite soluzioni: questo significa che le due rette, traslate sull'origine, hanno infiniti punti in comune, ovvero coincidono, e quindi le due rette di partenza, prima della traslazione, erano parallele.

Osservazione 2.20. Per rette date in equazioni parametriche verificare se esse sono parallele o meno è immediato, in quanto come sappiamo dal capitolo precedente nelle equazioni parametriche $\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}$ un vettore che rappresenta la direzione della retta è dato dalla terna (l, m, n) dei coefficienti di t: basta quindi confontare i due vettori così ottenuti per ognuna delle due rette, che avranno la stessa direzione se tali vettori sono proporzionali. Nel caso in cui le rette siano date in equazioni cartesiane, si può passare alle parametriche semplicemente risolvendo i sistemi dati dalle cartesiane stesse. Ad esempio, per le rette viste sopra in (2.49), riducendo la matrice completa delle cartesiane di r otteniamo

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \end{array}\right) \underset{R_2 \to R_2 - R_1}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \end{array}\right)$$

ovvero il sistema ridotto $\begin{cases} x + y + z = 1 \\ -2y + z = -1 \end{cases}$, da cui ponendo $z = t$ si ricava $-2y = -1 - t$ (ovvero $y = \frac{1}{2} + \frac{1}{2}t$) e $x = 1 - y - z = 1 - (\frac{1}{2} + \frac{1}{2}t) - t = \frac{1}{2} - \frac{3}{2}t$: quindi r ha equazioni parametriche

$$\begin{cases} x = \frac{1}{2} - \frac{3}{2}t \\ y = \frac{1}{2} + \frac{1}{2}t \\ z = t \end{cases} \tag{2.51}$$

Analogamente, riducendo la matrice completa delle cartesiane di r' otteniamo

$$\left(\begin{array}{ccc|c} 2 & 0 & 3 & 3 \\ 1 & 3 & 0 & -3 \end{array}\right) \underset{R_2 \to 2R_2 - R_1}{\longrightarrow} \left(\begin{array}{ccc|c} 2 & 0 & 3 & 1 \\ 0 & 6 & -3 & -9 \end{array}\right)$$

ovvero il sistema ridotto $\begin{cases} 2x + 3z = 3 \\ 6y - 3z = -9 \end{cases}$, da cui ponendo $z = t$ si ricava $6y = 9 + 3t$ (ovvero $y = \frac{3}{2} + \frac{1}{2}t$) e $2x = 3 - 3z = 3 - 3t$, ovvero $x = \frac{3}{2} - \frac{3}{2}t$: quindi r ha equazioni parametriche

$$\begin{cases} x = \frac{3}{2} - \frac{3}{2}t \\ y = \frac{3}{2} + \frac{1}{2}t \\ z = t \end{cases} \tag{2.52}$$

Confrontando i coefficienti di t nelle parametriche (2.51) e (2.52) vediamo che le rette hanno entrambe direzione rappresentata dal vettore di coordinate $(-\frac{3}{2}, \frac{1}{2}, 1)$, e quindi sono parallele[8].

Osservazione 2.21. Nel caso in cui si voglia fare il passaggio inverso, ovvero passare da parametriche a cartesiane, nel caso della retta basta ricavare il parametro t da una delle espressioni che compongono le parametriche e

[8]Si noti che sappiamo che le rette hanno la stessa direzione, che non esclude il caso in cui esse siano parallele coincidenti, ovvero che le equazioni cartesiane date rappresentassero in realtà la stessa retta.

sostituirlo nelle altre. Ad esempio, se la retta ha equazioni parametriche $\begin{cases} x = 1 - t \\ y = 2 + t \\ z = 1 - 3t \end{cases}$ si può ricavare $t = 1 - x$ dalla prima che, sostituita nelle altre, ci dà $y = 2 + (1 - x) = 3 - x$ e $z = 1 - 3(1 - x) = -2 + 3x$. Possiamo allora affermare che la retta ha equazioni cartesiane $\begin{cases} x + y = 3 \\ -3x + z = -2 \end{cases}$
Nel caso del piano, si può seguire un procedimento analogo ricavando prima uno dei due parametri e poi il secondo: se ad esempio il piano ha equazioni parametriche $\begin{cases} x = 1 + t + s \\ y = 2 + t - s \\ z = 3 - 2t - s \end{cases}$ si può ricavare $t = x - 1 - s$ dalla prima che, sostituita nelle altre, ci dà $y = 2 + (x - 1 - s) - s = 1 + x - 2s$ e $z = 3 - 2(x - 1 - s) - s = 5 - 2x + s$; ricavano poi s da questa seconda espressione, ovvero $s = z + 2x - 5$, possiamo sostituirlo nell'altra ottenendo $y = 1 + x - 2(z + 2x - 5) = 11 - 3x - 2z$, ovvero $3x + y + 2z = 11$, che è l'equazione cartesiana del piano. Nel capitolo successivo vedremo comunque un metodo più elegante per arrivare allo stesso risultato.

Un altro problema geometrico che può essere risolto con l'aiuto delle tecniche viste a proposito della risoluzione dei sistemi è il seguente: supponiamo di avere una retta data in equazioni cartesiane

$$\begin{cases} Ax + By + Cz = D \\ A'x + B'y + C'z = D' \end{cases} \tag{2.53}$$

Come abbiamo visto nel capitolo precedente quando abbiamo ricavato tali equazioni, la (2.53) ci sta semplicemente dicendo che la retta data è intersezione del piano dato dall'equazione cartesiana $Ax + By + Cz = D$ e dal piano di equazione cartesiana $A'x + B'y + C'z = D'$: le due equazioni che compongono le cartesiane sono quindi le equazioni di due particolari piani che contengono la retta.
Ora, vogliamo determinare *tutti* i piani che contengono la retta.
Si ha la seguente

Proposizione 2.22. La generica equazione cartesiana del piano che contiene la retta (2.53) è data da

$$\alpha(Ax + By + Cz - D) + \beta(A'x + B'y + C'z - D') = 0 \tag{2.54}$$

al variare di $\alpha, \beta \in \mathbb{R}$.

Proof. Iniziamo con l'osservare che se un piano ha equazione della forma (2.54), allora esso contiene la retta. Infatti, dire che una retta è contenuta in un piano significa che se un punto appartiene alla retta allora esso appartiene anche al piano: ma se un punto appartiene alla retta, allora le sue coordinate (x, y, z) soddisfano entrambe le equazioni $Ax + By + Cz = D$ e $A'x + B'y + C'z = D'$ della retta, e quindi

$$\alpha(Ax+By+Cz-D)+\beta(A'x+B'y+C'z-D') = \alpha(D-D)+\beta(D'-D') = \alpha\cdot 0+\beta\cdot 0 = 0$$

ovvero (x, y, z) soddisfa anche l'equazione (2.54), cioè il punto appartiene al piano rappresentato da tale equazione. Questo dimostra che, per ogni $\alpha, \beta \in \mathbb{R}$, l'equazione (2.54) rappresenta un piano che contiene la retta.

Viceversa, dobbiamo ora essere sicuri che qualunque piano che contenga la retta può essere rappresentato nella forma (2.54). Per vederlo, osserviamo che un generico piano di equazione $A''x+B''y+C''z = D''$ contiene tutta la retta di equazioni (2.53) solo se ogni punto che sta sulla retta sta anche sul piano, ovvero se e solo se ogni terna che soddisfa le equazioni $Ax + By + Cz = D$ e $A'x + B'y + C'z = D'$ soddisfa automaticamente anche l'equazione $A''x + B''y + C''z = D''$ del piano.
In altre parole, nel sistema

$$\begin{cases} Ax + By + Cz = D \\ A'x + B'y + C'z = D' \\ A''x + B''y + C''z = D'' \end{cases}$$

che si ottiene mettendo insieme tutte le cartesiane, la terza equazione è superflua ovvero dipendente dalle altre. A livello della matrice completa

$$\begin{pmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{pmatrix}$$

questo si traduce nel fatto che la terza riga deve essere combinazione lineare delle altre due, ovvero devono esistere $\alpha, \beta \in \mathbb{R}$ tali che

$$\begin{pmatrix} A'' & B'' & C'' & D'' \end{pmatrix} = \alpha \begin{pmatrix} A & B & C & D \end{pmatrix} + \beta \begin{pmatrix} A' & B' & C' & D' \end{pmatrix}$$

ovvero

$$A'' = \alpha A + \beta A',\ B'' = \alpha B + \beta B',\ C'' = \alpha C + \beta C',\ D'' = \alpha D + \beta D'$$

Quindi l'equazione $A''x + B''y + C''z = D''$ si riscrive

$$(\alpha A + \beta A')x + (\alpha B + \beta B')y + (\alpha C + \beta C')z = \alpha D + \beta D'$$

che, si vede facilmente svolgendo i conti e confrontando, equivale proprio alla (2.54).

$\square$

Esempio 2.23. Data la retta

$$r : \begin{cases} x + y + z = 1 \\ x - y + 2z = 0 \end{cases}$$

si determini il piano π che contiene r e passa per il punto P_0 di coordinate $(1, 1, 1)$.
Determiniamo prima tutti i piani che contengono r, che secondo la Proposizione 2.22 sono dati al variare di $\alpha, \beta \in \mathbb{R}$ dall'equazione

$$\alpha(x + y + z - 1) + \beta(x - y + 2z) = 0. \tag{2.55}$$

Poiché tra tutti questi piani cerchiamo quello che contiene il punto $(1, 1, 1)$, vogliamo che tale equazione sia soddisfatta quando poniamo $x = 1, y = 1, z = 1$, ovvero, sostituendo,

$$\alpha(1 + 1 + 1 - 1) + \beta(1 - 1 + 2) = 2\alpha + 2\beta = 0$$

da cui $\alpha = -\beta$. Sostituendo questa relazione nella (2.55), si ottiene

$$-\beta(x + y + z - 1) + \beta(x - y + 2z) = 0$$

ovvero, svolgendo i calcoli,

$$-2\beta y + \beta z + \beta = 0.$$

Al variare del parametro β, queste equazioni rappresentano tutte lo stesso piano (il piano π cercato) in quanto si tratta di equazioni proporzionali, tutte equivalenti: dividendo per il parametro β (o, equivalentemente, scegliendo per esempio $\beta = 1$), possiamo allora scrivere che π ha equazione cartesiana

$$-2y + z + 1 = 0.$$

Se, invece del passaggio per il punto, volessimo imporre per esempio che il piano, oltre a contenere r, fosse parallelo a un altro piano, ad esempio quello di equazione cartesiana $x + 2y + 3z = -1$, dovremmo procedere come segue.

Come abbiamo visto nel capitolo precedente, due piani $Ax+By+Cz=D$ e $A'x+B'y+C'z=D'$ sono paralleli se e solo se le terne (A,B,C) e (A',B',C') sono proporzionali, in quanto rappresentano le coordinate di vettori normali (perpendicolari) ai piani. In realtà, poiché abbiamo la libertà di moltiplicare l'equazione di un piano per qualunque coefficiente senza che il piano venga modificato, possiamo sempre far sì che sia $(A,B,C)=(A',B',C')$. Allora, poiché svolgendo i calcoli nella (2.55), vediamo che il generico piano che contiene r è della forma

$$(\alpha+\beta)x+(\alpha-\beta)y+(\alpha+2\beta)z-\alpha=0, \tag{2.56}$$

la condizione di parallelismo tra questo piano e il piano di equazione $x+2y+3z=-1$ è

$$(\alpha+\beta,\alpha-\beta,\alpha+2\beta)=(1,2,3)$$

ovvero

$$\begin{cases} \alpha+\beta=1 \\ \alpha-\beta=2 \\ \alpha+2\beta=3 \end{cases}$$

Per trovare il piano dato, basta quindi risolvere tale sistema e sostituire i valori di α e β trovati nella (2.56). In questo caso, si vede riducendo a gradini la sua matrice completa

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 2 & 3 \end{array}\right) \underset{\substack{R_2\to R_2-R_1 \\ R_3\to R_3-R_1}}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 1 & 2 \end{array}\right) \underset{R_3\to 2R_3+R_2}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 5 \end{array}\right)$$

il sistema è incompatibile e quindi la condizione di parallelismo non può essere soddisfatta: tra i piani che contengono la retta r, non ne esiste nessuno che è parallelo al piano dato.

Esempio 2.24. Date le due rette

$$r_1:\begin{cases} x+y+z=3 \\ x-2y+z=0 \end{cases} \quad r_2:\begin{cases} 2x+y-z=2 \\ x-y-z=-1 \end{cases}$$

determiniamo, se esiste, il piano che le contiene.
Come abbiamo visto sopra, il generico piano che contiene r_1 ha equazione

$$\alpha_1(x+y+z-3)+\beta_1(x-2y+z)=0$$

ovvero

$$(\alpha_1+\beta_1)x+(\alpha_1-2\beta_1)y+(\alpha_1+\beta_1)z-3\alpha_1=0 \tag{2.57}$$

mentre il generico piano che contiene r_2 ha equazione

$$\alpha_2(2x + y - z - 2) + \beta_2(x - y - z + 1) = 0$$

ovvero

$$(2\alpha_2 + \beta_2)x + (\alpha_2 - \beta_2)y + (-\alpha_2 - \beta_2)z + (-2\alpha_2 + \beta_2) = 0. \qquad (2.58)$$

Per trovare, se esiste, il piano che contiene entrambe le rette basta vedere se esistono valori di $\alpha_1, \beta_1, \alpha_2, \beta_2$ tali che la (2.57) e la (2.58) sono uguali: uguagliando i coefficienti di x, y, z e il termine noto in tali equazioni si ottiene

$$\begin{cases} \alpha_1 + \beta_1 = 2\alpha_2 + \beta_2 \\ \alpha_1 - 2\beta_1 = \alpha_2 - \beta_2 \\ \alpha_1 + \beta_1 = -\alpha_2 - \beta_2 \\ -3\alpha_1 = -2\alpha_2 + \beta_2 \end{cases}$$

ovvero il sistema omogeneo di 4 equazioni in 4 incognite

$$\begin{cases} \alpha_1 + \beta_1 - 2\alpha_2 - \beta_2 = 0 \\ \alpha_1 - 2\beta_1 - \alpha_2 + \beta_2 = 0 \\ \alpha_1 + \beta_1 + \alpha_2 + \beta_2 = 0 \\ -3\alpha_1 + 2\alpha_2 - \beta_2 = 0 \end{cases}$$

Tale sistema ha sicuramente sempre la soluzione $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0$, ma se sostituissimo tali valori nelle (2.57) e la (2.58) otterremmo $0 = 0$, che non è l'equazione di un piano: quindi per l'esistenza del piano che contiene entrambe le rette deve esistere una soluzione non nulla di tale sistema.
Riducendo a gradini la sua matrice dei coefficienti troviamo

$$\begin{pmatrix} 1 & 1 & -2 & -1 \\ 1 & -2 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -3 & 0 & 2 & -1 \end{pmatrix} \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \\ R_4 \to R_4 + 3R_1}}{\longrightarrow} \begin{pmatrix} 1 & 1 & -2 & -1 \\ 0 & -3 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 3 & -4 & -4 \end{pmatrix}$$

$$\underset{R_4 \to R_4 + R_2}{\longrightarrow} \begin{pmatrix} 1 & 1 & -2 & -1 \\ 0 & -3 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & -3 & -2 \end{pmatrix} \underset{R_4 \to R_4 + R_3}{\longrightarrow} \begin{pmatrix} 1 & 1 & -2 & -1 \\ 0 & -3 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Poichè la matrice ha rango 3 il sistema ha sicuramente altre soluzioni oltre alla 4-upla nulla $(0, 0, 0, 0)$, quindi esiste il piano che contiene le rette.
Per trovarlo, basta risolvere il sistema, che abbiamo ridotto alla forma equivalente

$$\begin{cases} \alpha_1 + \beta_1 - 2\alpha_2 - \beta_2 = 0 \\ -3\beta_1 + \alpha_2 + 2\beta_2 = 0 \\ 3\alpha_2 + 2\beta_2 = 0 \end{cases}$$

Vediamo però che non è necessario determinare completamente la soluzione del sistema: infatti, l'ultima equazione non nulla ci dà il valore di α_2 (in funzione di β_2), mentre le prime due i valori di α_1 e β_1 (sempre in funzione di β_2): se sostituiamo i valori di α_1 e β_1 così trovati nella (2.57) o quello di α_2 nella (2.58) otteniamo lo stesso piano (il sistema esprime proprio la condizione che i due piani siano uguali), quindi per trovarlo basta determinare solo α_2 dall'ultima equazione non nulla $3\alpha_2 + 2\beta_2 = 0$ senza dover risolvere le altre due.
Questa equazione ci dà $\alpha_2 = -\frac{2}{3}\beta_2$ che sostituita nella (2.58) dà

$$(-\frac{4}{3}\beta_2 + \beta_2)x + (-\frac{2}{3}\beta_2 - \beta_2)y + (\frac{2}{3}\beta_2 - \beta_2)z + (\frac{4}{3}\beta_2 + \beta_2) = 0$$

ovvero

$$-\frac{1}{3}\beta_2 x - \frac{5}{3}\beta_2 y - \frac{1}{3}\beta_2 z + \frac{7}{3}\beta_2 = 0.$$

Dividendo per β_2 e moltiplicando per -3 (abbiamo la libertà di moltiplicare l'equazione di un piano per qualunque coefficiente non nullo) si ottiene infine

$$x + 5y + z - 7 = 0$$

che è l'equazione del piano cercato.

Come abbiamo detto, il procedimento di riduzione a gradini è uno strumento non solo per risolvere un sistema ma più in generale per scoprire se delle n-uple sono linearmente indipendenti e,in caso negativo, per individuare le n-uple che sono combinazione delle altre. A sua volta, questo ci consente di determinare se sono indipendenti i vettori in un qualunque spazio vettoriale: infatti, in virtù della Proposizione 1.23, basta fissare una base, rappresentare i vettori mediante le n-uple delle loro coordinate e poi controllare se sono indipendenti tali n-uple mediante una semplice riduzione a gradini.

Esempio 2.25. Fissata una base $O\vec{P}_1, O\vec{P}_2, O\vec{P}_3$ dello spazio tridimensionale dei vettori applicati V_O^3, si dica se i tre vettori applicati

$$v_1 = O\vec{P}_1 + 2O\vec{P}_2 + 3O\vec{P}_3$$

$$v_2 = 3\vec{OP_1} + 4\vec{OP_2} + 5\vec{OP_3}$$
$$v_3 = -\vec{OP_1} + 2\vec{OP_2} + 5\vec{OP_3}$$

formano ancora una base.

Come sappiamo, tre vettori in uno spazio di dimensione 3 formano una base se e solo se sono indipendenti: per determinare l'indipendenza di v_1, v_2, v_3 basta considerare le terne $(1, 2, 3), (3, 4, 5), (-1, 2, 5)$ delle loro coordinate rispetto alla base di partenza $\vec{OP_1}, \vec{OP_2}, \vec{OP_3}$ e determinare equivalentemente se esse sono indipendenti in $\mathbb{R}^3$.
Costruiamo la matrice che ha tali terne come righe

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ -1 & 2 & 5 \end{pmatrix}$$

e effettuiamo una riduzione a gradini:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ -1 & 2 & 5 \end{pmatrix} \underset{\substack{R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 + R_1}}{\longrightarrow} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 4 & 8 \end{pmatrix} \underset{R_3 \to R_3 + 2R_2}{\longrightarrow} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & 0 \end{pmatrix}$$

Essendosi annullata la terza riga, significa che questa era combinazione delle prime due: quindi le tre terne non sono indipendenti e non lo sono neanche i vettori applicati v_1, v_2, v_3 da esse rappresentate (tali vettori giacevano quindi su uno stesso piano).