L'eliminazione di Gauss

3

Sommario

In questo capitolo iniziamo a studiare i sistemi lineari quadrati, descrivendo il metodo di eliminazione di Gauss, una delle procedure più efficienti per la loro risoluzione; l'esame dei sistemi lineari qualunque è rimandato al Capitolo 6, quando avremo sviluppato gli strumenti teorici necessari.

3.1 Esempi e definizioni

Studiando i sistemi lineari, i problemi principali che vogliamo risolvere sono tre: quando un sistema lineare ammette soluzioni? E se le ammette, quante sono? E come si trovano? Per avere un'idea del tipo di situazioni che si possono incontrare, cominciamo con l'esaminare un esempio semplicissimo.

Esempio 3.1

Consideriamo il caso di un'equazione di primo grado in una sola incognita: $ax = b$. Se $a \neq 0$, l'equazione ammette un'unica soluzione $x = b/a$; se $a = 0$ ma $b \neq 0$, l'equazione non ha alcuna soluzione; se $a = 0$ e $b = 0$, l'equazione ha infinite soluzioni, in quanto qualunque $x \in \mathbb{R}$ la soddisfa.

In questo caso le soluzioni o non esistono o sono infinite oppure ce n'è una sola; non capita mai che siano in numero finito maggiore di uno. È naturale chiedersi se questo accade per ogni sistema lineare. In teoria, potrebbe dipendere dall'aver considerato una sola equazione; magari ci sono sistemi di due equazioni che ammettono esattamente due soluzioni, sistemi di otto equazioni che ammettono esattamente otto soluzioni e così via. Esaminiamo qualche esempio di sistema lineare con più di un'equazione (e più di un'incognita), e vediamo che cosa succede.

Esempio 3.2

Studiamo il seguente sistema lineare di tre equazioni in tre incognite

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ x_1 - x_2 + 4x_3 = 2, \\ x_1 - 2x_2 + 2x_3 = 3. \end{cases} \tag{3.1}$$

Per risolverlo, cominciamo col notare che la prima e la terza equazione differiscono solo per il coefficiente di x_3 e il termine noto. Quindi sottraendo la prima equazione alla terza troviamo subito il valore di x_3

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ x_1 - x_2 + 4x_3 = 2, \\ x_3 = 2. \end{cases}$$

Già che ci siamo, sottraiamo la prima equazione anche alla seconda, ottenendo

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

Così la terza equazione ci fornisce il valore di x_3, che sostituito nella seconda ci dà il valore di x_2, e sostituendo entrambi nella prima otteniamo il valore di x_1

$$\begin{cases} x_1 = -11, \\ x_2 = -5, \\ x_3 = 2. \end{cases}$$

Una veloce verifica mettendo questi valori in (3.1) conferma che abbiamo risolto il sistema, che quindi ammette un'unica soluzione (e non tre).

Esempio 3.3

Ora proviamo con un sistema lievemente diverso:

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ x_1 - 2x_2 + 4x_3 = 2, \\ x_1 - 2x_2 + 2x_3 = 3, \end{cases} \tag{3.2}$$

dove è cambiato solo il coefficiente di x_2 nella seconda equazione. Procedendo come prima, cioè sottraendo la prima equazione alla seconda e alla terza, otteniamo

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ \qquad\qquad 3x_3 = 1, \\ \qquad\qquad\; x_3 = 2. \end{cases}$$

Ma allora la seconda equazione dà $x_3 = 1/3$, mentre la terza dà $x_3 = 2$, due risultati incompatibili. In altre parole, questo sistema non ha soluzione.

Esempio 3.4

Facciamo ancora una piccola modifica, cambiando stavolta il termine noto nella seconda equazione, e consideriamo il sistema

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ x_1 - 2x_2 + 4x_3 = 7, \\ x_1 - 2x_2 + 2x_3 = 3. \end{cases} \tag{3.3}$$

Sottraendo nuovamente la prima equazione alla seconda e alla terza otteniamo

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ \qquad\qquad 3x_3 = 6, \\ \qquad\qquad\; x_3 = 2. \end{cases}$$

Questa volta la seconda e la terza equazione non sono più in contrasto; entrambe ci dicono che $x_3 = 2$. Inoltre, contrariamente a quanto accadeva nell'Esempio 3.2, la seconda equazione non fornisce alcuna condizione su x_2; in altri termini, x_2 è *libera*, può assumere qualsiasi valore. Sostituendo $x_3 = 2$ nella prima equazione e ponendo $x_2 = t$ (per ricordarci che il valore di x_2 non è fissato ma è libero) si ha

$$\begin{cases} x_1 = -1 + 2t, \\ x_2 = t, \\ x_3 = 2. \end{cases}$$

Stavolta il sistema ha infinite soluzioni, una per ogni valore di $t \in \mathbb{R}$. Utilizzando le notazioni introdotte nel capitolo precedente, le soluzioni sono quindi descritte da

$$\begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} = \begin{vmatrix} -1 \\ 0 \\ 2 \end{vmatrix} + t \begin{vmatrix} 2 \\ 1 \\ 0 \end{vmatrix},$$

che è l'equazione parametrica di una retta nello spazio. Quindi in un senso molto preciso possiamo dire che il sistema (3.3) ammette una *retta* di soluzioni.

Osservazione 3.1

Nota che per passare da un sistema con soluzione unica a uno senza soluzione (o con infinite soluzioni) abbiamo modificato i coefficienti, mentre per passare da un sistema senza soluzioni a uno con infinite soluzioni abbiamo toccato i termini noti.

In questi esempi le soluzioni o non esistono o sono infinite oppure ce n'è una sola. Per vedere che cosa accade in generale ci serve un po' di terminologia.

Osservazione 3.2

Quanto faremo da qui in poi (a meno che non venga espressamente detto il contrario) dipende solo dalle proprietà formali della somma e del prodotto di numeri reali descritte nel Paragrafo 1.3, e quindi vale in qualunque campo, anche se, spesso, per semplicità lavoreremo con $\mathbb{R}$.

Definizione 3.1

La forma generale di un *sistema lineare* di *m equazioni* in *n incognite* è

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \qquad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{3.4}$$

Il numero n delle incognite è anche detto *ordine* del sistema. I numeri $a_{11}, \ldots, a_{mn}$ sono i *coefficienti* del sistema, e solitamente sono raccolti in una tabella di numeri, la *matrice dei coefficienti*

$$A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix}. \tag{3.5}$$

I numeri $b_1, \ldots, b_m$ sono i *termini noti*, e usualmente sono raccolti in una colonna b (che, per motivi che vedremo nel Paragrafo 4.1, si chiama *vettore dei termini noti*). Analogamente abbiamo il *vettore x delle incognite*

$$b = \begin{vmatrix} b_1 \\ \vdots \\ b_m \end{vmatrix}, \qquad x = \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}.$$

La tabella di numeri

$$A' = \begin{vmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{vmatrix}$$

è detta *matrice completa* del sistema. Per indicare il sistema (3.4) per brevità (e per motivi più profondi che discuteremo nel Paragrafo 5.1) spesso scriveremo $Ax = b$.

Definizione 3.2

Una *soluzione* del sistema (3.4) è una n-upla $(v_1, \ldots, v_n)$ di numeri che sostituiti ordinatamente alle incognite $x_1, \ldots, x_n$ soddisfano le equazioni del sistema. Il sistema (3.4) è *compatibile* se ammette almeno una soluzione.

Definizione 3.3

In generale, una *matrice* con m righe e n colonne è una tabella rettangolare di numeri con m righe e n colonne, come quella in (3.5). L'insieme delle matrici con m righe e n colonne (dette anche matrici $m \times n$) a coefficienti reali sarà indicato con $M_{m,n}(\mathbb{R})$. Se i coefficienti sono invece in un campo $\mathbb{K}$, scriveremo $M_{m,n}(\mathbb{K})$. A volte la matrice (3.5) sarà indicata con $A = (a_{ij})$; il numero a_{ij} è l'*elemento* (o *coefficiente*) di *posto* (i, j) della matrice A, dove il primo indice i è l'*indice di riga*, mentre il secondo indice j è l'*indice di colonna*.

Definizione 3.4

Sia $A \in M_{m,n}(\mathbb{R})$ una matrice $m \times n$. Indicheremo con A_i la sua *riga* i-esima (per $i = 1, \dots, m$), e con A^j la sua j-esima *colonna* (per $j = 1, \dots, n$)

$$A_i = \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{vmatrix}, \qquad A^j = \begin{vmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{vmatrix}.$$

Una *matrice quadrata* è una matrice con tante righe quante colonne. Una matrice quadrata $n \times n$ è detta di *ordine* n. La *diagonale principale* di una matrice quadrata è la diagonale che va dall'angolo in alto a sinistra a quello in basso a destra; è composta dagli elementi a_{ii}, in cui l'indice di riga è uguale all'indice di colonna. Una matrice quadrata A è detta: *diagonale* se tutti gli elementi al di fuori della diagonale principale sono nulli; *triangolare superiore* se gli elementi al di sotto (cioè a sinistra) della diagonale principale sono tutti nulli; *triangolare inferiore* se invece sono zero gli elementi sopra la diagonale principale.

Esempio 3.5

La matrice

$$A = \begin{vmatrix} 2 & 3 & 1 \\ \sqrt{2} & 0 & 6/7 \end{vmatrix}$$

ha 2 righe e 3 colonne ed è a coefficienti reali, cioè $A \in M_{2,3}(\mathbb{R})$. I suoi elementi sono $a_{11} = 2$, $a_{12} = 3$, $a_{13} = 1$, $a_{21} = \sqrt{2}$, $a_{22} = 0$, e $a_{23} = 6/7$, le sue colonne sono

$$A^1 = \begin{vmatrix} 2 \\ \sqrt{2} \end{vmatrix}, \qquad A^2 = \begin{vmatrix} 3 \\ 0 \end{vmatrix}, \qquad \text{e} \qquad A^3 = \begin{vmatrix} 1 \\ 6/7 \end{vmatrix}$$

e le sue righe sono

$$A_1 = \begin{vmatrix} 2 & 3 & 1 \end{vmatrix} \qquad \text{e} \qquad A_2 = \begin{vmatrix} \sqrt{2} & 0 & 6/7 \end{vmatrix}.$$

Esempio 3.6

Delle matrici

$$\begin{vmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3\pi^2 \end{vmatrix}, \qquad \begin{vmatrix} 4 & 0 & \pi \\ 0 & 6\sqrt{2} & 27 \\ 0 & 0 & 0 \end{vmatrix}, \qquad \begin{vmatrix} 0 & 0 & 0 \\ 2/3 & 0 & 0 \\ e+1 & 2/3 & \sqrt{7} \end{vmatrix},$$

la prima è diagonale, la seconda triangolare superiore e la terza triangolare inferiore.

3.2 Sistemi triangolari superiori

In questo paragrafo ci occuperemo di un tipo molto semplice di sistema lineare.

Definizione 3.5

Un sistema lineare *quadrato* (cioè con tante equazioni quante incognite) si dice *triangolare superiore* se la sua matrice dei coefficienti è triangolare superiore.

Negli Esempi 3.2-3.4 abbiamo studiato alcuni sistemi lineari quadrati riducendoli a forma triangolare superiore e poi risolvendoli (quando possibile) a occhio. Vedremo che si può sempre procedere così; per questo dimostriamo un risultato che ci dice quando un sistema triangolare superiore ammette soluzione unica.

Proposizione 3.1

Un sistema lineare triangolare superiore $Ax = b$ di n equazioni in n incognite ammette una e una sola soluzione se e solo se tutti gli elementi della diagonale principale della matrice A dei coefficienti sono diversi da zero.

Dimostrazione

Supponiamo prima di tutto che tutti gli elementi $a_{11}, \dots, a_{nn}$ della diagonale principale di A siano non nulli. In questo caso, l'ultima equazione del sistema è della forma $a_{nn}x_n = b_n$, con $a_{nn} \neq 0$; quindi ammette l'unica soluzione $v_n = b_n/a_{nn}$. Sostituendo v_n nella penultima equazione otteniamo

$$a_{n-1,n-1}x_{n-1} = b_{n-1} - a_{n-1,n}v_n,$$

che di nuovo ammette un'unica soluzione v_{n-1}. Procedendo in questo modo giungiamo fino alla prima equazione, che ora è diventata $a_{11}x_1 = b_1 - a_{12}v_2 - \cdots - a_{1n}v_n$, e di nuovo troviamo un'unica soluzione v_1. In conclusione, la n-upla $(v_1, \dots, v_n)$ così determinata è l'unica soluzione del sistema.

Osservazione 3.3

Questa dimostrazione fornisce un metodo per il calcolo della soluzione, detto *risoluzione all'indietro* di un sistema triangolare superiore.

Osservazione 3.4

Finora, abbiamo proceduto per *dimostrazione diretta*: supposta vera l'ipotesi, tramite ragionamenti ne abbiamo dedotto la tesi. Ci sono anche altri modi per dimostrare affermazioni del tipo "$\mathcal{A}$ implica $\mathcal{B}$". La *dimostrazione per assurdo* procede supponendo vere contemporaneamente l'ipotesi "$\mathcal{A}$" e la negazione "non $\mathcal{B}$" della tesi per poi giungere, tramite ragionamenti, a una contraddizione, cioè a qualcosa di sicuramente falso. Quindi "$\mathcal{A}$" e "non $\mathcal{B}$" non possono essere contemporaneamente veri, per cui se "$\mathcal{A}$" è vera anche "$\mathcal{B}$" dev'esserlo. Invece, la *dimostrazione inversa* suppone vera la negazione "non $\mathcal{B}$" della tesi e, tramite ragionamenti, deduce la negazione "non $\mathcal{A}$" dell'ipotesi. Anche questa è una dimostrazione dell'affermazione "$\mathcal{A}$ implica $\mathcal{B}$": infatti, se "non $\mathcal{B}$" implica "non $\mathcal{A}$", allora se "$\mathcal{A}$" e "non $\mathcal{B}$" fossero contemporaneamente vere lo sarebbero anche "$\mathcal{A}$" e "non $\mathcal{A}$", che è impossibile.

Ritornando alla dimostrazione della Proposizione 3.1, utilizzeremo ora una dimostrazione inversa. Dobbiamo dimostrare che se il nostro sistema $Ax = b$ ammette una e una sola soluzione, allora tutti gli elementi sulla diagonale principale di A sono diversi da zero.

Supponiamo che ci sia almeno un coefficiente nullo sulla diagonale principale (la negazione della tesi); dobbiamo dimostrare che il sistema non ammette un'unica soluzione, cioè o ne ammette più di una o non ne ammette neanche una (la negazione dell'ipotesi). Se il nostro sistema non ammette soluzioni, abbiamo finito. Supponiamo allora che una soluzione $(v_1, \ldots, v_n)$ ci sia; ne vogliamo costruire un'altra, diversa. Sia k il minimo indice per cui $a_{kk} = 0$; in particolare, $a_{11}, \ldots, a_{k-1,k-1} \neq 0$. Sostituiamo $v_{k+1}, \ldots, v_n$ al posto di $x_{k+1}, \ldots, x_n$ nella k-esima equazione. Ricordando che $a_{kk} = 0$, troviamo

$$0x_k = b_k - a_{k,k+1}v_{k+1} - \cdots - a_{kn}v_n. \tag{3.6}$$

Siccome abbiamo supposto che il nostro sistema ha almeno una soluzione, si deve per forza avere (ricorda l'Esempio 3.1) $b_k - a_{k,k+1}v_{k+1} - \cdots - a_{kn}v_n = 0$; quindi ogni x_k è soluzione di (3.6). Prendiamo allora $w \neq v_k$ e sostituiamo w, $v_{k+1}, \ldots, v_n$ al posto di $x_k, x_{k+1}, \ldots, x_n$ nelle prime $k-1$ equazioni. Otteniamo così un sistema triangolare superiore di ordine $k-1$ i cui coefficienti sulla diagonale principale sono tutti non nulli; per quanto già visto, questo sistema ammette un'unica soluzione $(u_1, \ldots, u_{k-1})$. Ma allora le n-uple $(v_1, \ldots, v_{k-1}, v_k, v_{k+1}, \ldots, v_n)$ e $(u_1, \ldots, u_{k-1}, w, v_{k+1}, \ldots, v_n)$ sono due soluzioni distinte del nostro sistema, come volevamo. □

Osservazione 3.5

La prima parte della Proposizione 3.1 può essere dimostrata anche usando una tecnica particolare di dimostrazione diretta, detta *dimostrazione per induzione*. L'idea è che per dimostrare che una qualche affermazione "$\mathcal{A}_n$" dipendente da un numero naturale n è vera per tutti gli $n \in \mathbb{N}$, basta dimostrare che è vera per $n = 1$ (oppure $n = 0$, a seconda dei casi) e poi dimostrare che se è vera per un qualunque $n \in \mathbb{N}$ (questa si chiama *ipotesi induttiva*), allora è vera anche per $n+1$. Infatti, in questo modo, sapendo che è vera per 1 dimostriamo che è vera per 2, e poi per 3, per 4 e così via. Per esempio, la dimostrazione per induzione della prima parte della Proposizione 3.1 funziona come segue. Il numero naturale in questione è l'ordine del sistema. Quando il sistema ha ordine 1, è una sola equazione in un'incognita, con coefficiente non nullo; quindi in questo caso l'Esempio 3.1. ci dice che vi è un'unica soluzione (cioè l'affermazione è vera). Supponiamo allora che l'affermazione sia vera per tutti i sistemi triangolari superiori di ordine n (ipotesi induttiva), e consideriamo un sistema triangolare superiore di ordine $n+1$. L'ultima equazione sarà del tipo $a_{n+1,n+1}x_{n+1} = b_{n+1}$ con $a_{n+1,n+1} \neq 0$; quindi ammette un'unica soluzione $v_{n+1} = b_{n+1}/a_{n+1,n+1}$. Sostituiamo v_{n+1} al posto di x_{n+1} nelle restanti n equazioni; otteniamo un sistema triangolare superiore di ordine n con coefficienti non nulli sulla diagonale principale. Per l'ipotesi induttiva, questo sistema ammette un'unica soluzione $(v_1, \ldots, v_n)$; quindi $(v_1, \ldots, v_n, v_{n+1})$ è l'unica soluzione del nostro sistema. Dunque il *principio d'induzione* (che è il nome del ragionamento che si applica nelle dimostrazioni per induzione) ci ha permesso di rimpiazzare la verifica per sistemi di ordine qualunque con una semplice verifica per sistemi di ordine 1 e un ragionamento che ci fa passare dall'ordine n all'ordine $n+1$.

Il principio d'induzione può essere espresso nel modo seguente.

Principio d'induzione 1

Sia $\mathcal{A}_n$ un'affermazione dipendente da un numero naturale n. Supponiamo che:

a) $\mathcal{A}_0$ sia vera;

b) se $\mathcal{A}_n$ è vera, allora anche $\mathcal{A}_{n+1}$ lo è.

Allora $\mathcal{A}_n$ è vera per ogni numero naturale n.

Un'altra forma di cui non è difficile dimostrare l'equivalenza con la precedente è la seguente.

Principio d'induzione 2

Sia $\mathcal{A}_n$ un'affermazione dipendente da un numero naturale n. Supponiamo che:
a) $\mathcal{A}_0$ sia vera;
b) se $\mathcal{A}_k$ è vera per ogni $0 \leq k \leq n$, allora anche $\mathcal{A}_{n+1}$ è vera.
Allora $\mathcal{A}_n$ è vera per ogni numero naturale n.

Useremo più volte il principio d'induzione, nelle due forme appena indicate oppure in altre varianti equivalenti. Vediamo come esempio un'applicazione classica.

Esempio 3.7

Vogliamo dimostrare che "La somma degli angoli interni di un poligono convesso di $n+3$ lati è $(n+1)\pi$ radianti". Utilizziamo il Principio d'induzione 1. L'affermazione $\mathcal{A}_0$ è un teorema classico della geometria euclidea (Teorema 1.1). Supponiamo $\mathcal{A}_n$ vera e consideriamo un poligono convesso P di $n+4=(n+1)+3$ lati. Collegando due vertici con una diagonale saltando un solo vertice intermedio, dividiamo P in un triangolo (la cui somma degli angoli interni è π) e un poligono convesso P' di $n+3$ lati. Per ipotesi induttiva la somma degli angoli interni di P' è $(n+1)\pi$; quindi la somma degli angoli interni di P è $\pi+(n+1)\pi=(n+2)\pi$. Dunque abbiamo dimostrato $\mathcal{A}_{n+1}$ supponendo vera $\mathcal{A}_n$. Siccome $\mathcal{A}_0$ è vera, per il principio d'induzione abbiamo dimostrato il nostro enunciato. ⎿⏋

Torniamo adesso alla risoluzione all'indietro di sistemi triangolari superiori.

Esempio 3.8

Consideriamo il seguente sistema di tre equazioni in tre incognite

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

La matrice dei coefficienti e il vettore dei termini noti sono

$$A = \begin{vmatrix} 3 & -1 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 4 \end{vmatrix}, \qquad b = \begin{vmatrix} 3 \\ 0 \\ -4 \end{vmatrix}.$$

Dunque il sistema è triangolare superiore con elementi non nulli sulla diagonale principale; per la Proposizione 3.1, il sistema ammette una e una sola soluzione, e possiamo trovarla risolvendo all'indietro. L'ultima equazione dice che $z=-1$; sostituendo nella seconda equazione troviamo $-y=2$, cioè $y=-2$; sostituendo y e z nella prima troviamo $3x=3$, cioè $x=1$. Quindi la soluzione è data da $x=1$, $y=-2$, $z=-1$. ⎿⏋

Esempio 3.9

Consideriamo il seguente sistema di tre equazioni in tre incognite

$$\begin{cases} 2x + 4y - 2z = 8, \\ 3z = 6, \\ 2z = -2. \end{cases}$$

La matrice dei coefficienti e il vettore dei termini noti sono

$$A = \begin{vmatrix} 2 & 4 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{vmatrix}, \qquad b = \begin{vmatrix} 8 \\ 6 \\ -2 \end{vmatrix}.$$

Il sistema è triangolare superiore, ma stavolta $a_{22} = 0$. Risolvendo all'indietro vediamo che l'ultima equazione ci dà $z = -1$, e sostituendo nella seconda otteniamo $-3 = 6$: impossibile. Quindi questo sistema non ha soluzioni. ⊔⊓

Esempio 3.10

Consideriamo il seguente sistema di tre equazioni in tre incognite

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

La matrice dei coefficienti e il vettore dei termini noti sono

$$A = \begin{vmatrix} 2 & 4 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{vmatrix}, \qquad b = \begin{vmatrix} 8 \\ -3 \\ -2 \end{vmatrix}.$$

Il sistema è di nuovo triangolare superiore con $a_{22} = 0$; quindi o non ha soluzioni o ne ha infinite. Risolvendo all'indietro troviamo $z = -1$; sostituendo nella seconda equazione otteniamo $-3 = -3$, che è sempre vera. Sostituendo il valore di z nella prima equazione troviamo $x = 3 - 2y$. Dunque stavolta abbiamo una soluzione per ogni valore di y – ovvero (come nell'Esempio 3.4) una retta di soluzioni, di equazioni parametriche $x = 3 - 2t$, $y = t$, $z = -1$. ⊔⊓

3.3 Il metodo d'eliminazione di Gauss

Vogliamo trovare ora un modo per trasformare un qualunque sistema lineare quadrato in un sistema triangolare superiore che abbia esattamente le stesse soluzioni; vedremo nel Paragrafo 6.2 come comportarci per sistemi non quadrati.

Definizione 3.6

Due sistemi lineari (anche non quadrati) dello stesso ordine si dicono *equivalenti* se hanno esattamente le stesse soluzioni.

Chiaramente, due sistemi ottenuti l'uno dall'altro cambiando l'ordine delle equazioni sono equivalenti. Lo scambiare due equazioni è la prima *operazione elementare* che possiamo effettuare su un sistema senza modificare le soluzioni.

Definizione 3.7

Date due equazioni $a_1x_1 + \cdots + a_nx_n = a$ e $b_1x_1 + \cdots + b_nx_n = b$, si dice *combinazione lineare* delle due equazioni di *coefficienti* $h, k \in \mathbb{R}$ l'equazione

$$h(a_1x_1 + \cdots + a_nx_n) + k(b_1x_1 + \cdots + b_nx_n) = ha + kb.$$

I coefficienti di questa nuova equazione sono $ha_j + kb_j$, per $j = 1, \ldots, n$.

Sostituendo un'equazione in un sistema lineare con una sua combinazione lineare con coefficienti non nulli le soluzioni del sistema non cambiano. Infatti vale il seguente risultato:

Lemma 3.2

Sia $Ax = b$ un sistema lineare contenente le equazioni

$$a_1x_1 + \cdots + a_nx_n = a$$

e

$$b_1x_1 + \cdots + b_nx_n = b. \tag{3.7}$$

Sia $\tilde{A}x = \tilde{b}$ il sistema lineare ottenuto sostituendo in $Ax = b$ l'equazione

$$h(a_1x_1 + \cdots + a_nx_n) + k(b_1x_1 + \cdots + b_nx_n) = ha + kb \tag{3.8}$$

al posto dell'equazione (3.7), dove $h, k \in \mathbb{R}$ sono due numeri reali con $k \neq 0$. Allora i sistemi $Ax = b$ e $\tilde{A}x = \tilde{b}$ sono equivalenti.

Dimostrazione

Sia $(v_1, \ldots, v_n)$ una soluzione di $Ax = b$. Questo vuol dire che i numeri $v_1, \ldots, v_n$ soddisfano tutte le equazioni del sistema. In particolare

$$\begin{cases} a_1v_1 + \cdots + a_nv_n = a, \\ b_1v_1 + \cdots + b_nv_n = b, \end{cases} \tag{3.9}$$

per cui $(v_1, \ldots, v_n)$ è chiaramente anche una soluzione di (3.8) – e quindi anche di $\tilde{A}x = \tilde{b}$, in quanto le altre equazioni non sono cambiate.

Viceversa, sia $(v_1, \ldots, v_n)$ una soluzione di $\tilde{A}x = \tilde{b}$. Allora la prima riga di (3.9) è ancora verificata; sostituendola in (3.8) troviamo $k(b_1v_1 + \cdots + b_nv_n) = kb$, e, essendo $k \neq 0$, vediamo che $v_1, \ldots, v_n$ soddisfano anche (3.7) – e quindi sono una soluzione di $Ax = b$. □

Dunque la sostituzione di un'equazione di un sistema con una sua combinazione lineare con un'altra equazione del sistema non cambia le soluzioni (purché il coefficiente che moltiplica l'equazione sostituita – k nell'enunciato del Lemma 3.2 – sia diverso da zero); è la seconda *operazione elementare* che possiamo effettuare su un sistema lineare per ottenere sistemi equivalenti.

Il *metodo di eliminazione di Gauss* è una procedura che trasforma, tramite operazioni elementari, un sistema lineare quadrato in un sistema triangolare superiore equivalente. Risolvendo quest'ultimo con una risoluzione all'indietro, otteniamo la soluzione del sistema di partenza. Il metodo di eliminazione di Gauss consiste in diversi passi (per l'esattezza uno meno dell'ordine del sistema). Il passo i-esimo annulla gli elementi sotto la diagonale principale della colonna i-esima della matrice dei coefficienti, producendo nel contempo un numero reale p_i, l'i-esimo *pivot* del sistema relativo alla data eliminazione di Gauss – e tutto senza cambiare le soluzioni del sistema.

Ma vediamo i dettagli. Partiamo da un sistema lineare quadrato $Ax = b$ di ordine n, dove

$$A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad \text{e} \quad b = \begin{vmatrix} b_1 \\ \vdots \\ b_n \end{vmatrix}.$$

Consideriamo la prima colonna di A. Se contiene solo zeri, poniamo $p_1 = 0$ e passiamo alla seconda colonna. Se invece contiene qualche elemento non nullo, scambiamo se necessario la prima equazione con una delle successive in modo che a essere diverso da zero sia a_{11}. Poniamo $p_1 = a_{11}$, e sommiamo all'equazione j-esima (per $j = 2, \ldots, n$) la prima equazione moltiplicata per $-a_{j1}/p_1$. Otteniamo così un sistema equivalente (grazie al Lemma 3.2) in cui gli elementi sotto la diagonale principale della prima colonna della matrice dei coefficienti sono tutti nulli.

I passi successivi sono molto simili. Supponiamo di aver già trattato le prime $i-1$ colonne del sistema, e consideriamo la colonna i-esima. Se A^i contiene solo zeri dalla riga i-esima compresa in giù, poniamo $p_i = 0$ e passiamo alla colonna successiva. Altrimenti, a meno di scambiare l'i-esima equazione con una sottostante, possiamo supporre che a_{ii} sia diverso da zero. Poniamo allora $p_i = a_{ii}$, e sommiamo alla j-esima equazione (per $j = i+1, \dots, n$) l'equazione i-esima moltiplicata per $-a_{ji}/p_i$. In questo modo otteniamo un sistema equivalente in cui gli elementi sotto la diagonale principale delle prime i colonne della matrice dei coefficienti sono tutti nulli.

Procedendo così giungiamo all'ultima colonna; ponendo infine $p_n = a_{nn}$ abbiamo ottenuto un sistema triangolare superiore, equivalente al sistema di partenza, con gli n pivot $p_1, \dots, p_n$ sulla diagonale principale.

Esempio 3.11

Consideriamo il sistema quadrato di ordine 4

$$\begin{cases} x + 3y + z - w = 1, \\ 3x + 9y + 4z + w = 1, \\ 2x + y + 5z + 2w = 0, \\ y - z - w = 2. \end{cases} \tag{3.10}$$

La matrice dei coefficienti e il vettore dei termini noti sono dati da

$$A = \begin{vmatrix} 1 & 3 & 1 & -1 \\ 3 & 9 & 4 & 1 \\ 2 & 1 & 5 & 2 \\ 0 & 1 & -1 & -1 \end{vmatrix} \quad \text{e} \quad b = \begin{vmatrix} 1 \\ 1 \\ 0 \\ 2 \end{vmatrix}.$$

Per semplicità di scrittura, eseguiremo le operazioni previste dall'eliminazione di Gauss sulle righe della matrice e del vettore dei termini noti, senza riscrivere ogni volta le equazioni. Siccome il primo elemento in alto a sinistra della matrice A è diverso da zero, troviamo subito $p_1 = 1$ senza bisogno di scambi di righe. Poi dobbiamo sottrarre alla seconda riga 3 volte la prima, alla terza 2 volte la prima e lasciare la quarta invariata. Quindi abbiamo effettuato la trasformazione seguente

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 3 & 9 & 4 & 1 & 1 \\ 2 & 1 & 5 & 2 & 0 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right|$$

dove per semplicità il vettore dei termini noti è scritto subito a destra della matrice dei coefficienti. Ora, nella seconda colonna sotto la seconda riga ci sono dei termini non nulli, ma l'elemento sulla diagonale principale è zero; quindi dobbiamo effettuare uno scambio di righe. Qui abbiamo una *scelta*: possiamo portare al secondo posto la terza riga oppure la quarta. Supponiamo di portarvi la terza; otteniamo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right|,$$

e il secondo pivot è $p_2 = -5$. A questo punto dobbiamo lasciare la terza riga invariata e sommare alla quarta $1/5$ della seconda, in modo da ottenere

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & -2/5 & -1/5 & 8/5 \end{array}\right|.$$

Siamo al passo 3. Non servono altri scambi di righe; quindi il terzo pivot è $p_3 = 1$, e dobbiamo semplicemente sommare $2/5$ della terza riga alla quarta, ottenendo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & -2/5 & -1/5 & 8/5 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 7/5 & 4/5 \end{array}\right|.$$

Il quarto pivot è $p_4 = 7/5$. In conclusione, il nostro sistema è equivalente al sistema triangolare superiore

$$\begin{cases} x + 3y + z - w = 1, \\ -5y + 3z + 4w = -2, \\ z + 4w = -2, \\ \dfrac{7}{5} w = \dfrac{4}{5}. \end{cases}$$

Risolvendo all'indietro otteniamo $x = 11$, $y = -12/7$, $z = -30/7$, $w = 4/7$, che è la soluzione (unica) del sistema (3.10) di partenza.

Esempio 3.12

Vediamo che cosa sarebbe successo se avessimo effettuato al secondo passo l'altro scambio di righe possibile. Scambiando la quarta riga con la seconda otteniamo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \end{array}\right|,$$

e il secondo pivot stavolta è $p_2 = 1$. A questo punto dobbiamo sommare 5 volte la seconda riga alla terza e lasciare la quarta invariata, in modo da ottenere

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & -2 & -1 & 8 \\ 0 & 0 & 1 & 4 & -2 \end{array}\right|.$$

Siamo al passo 3. Non servono altri scambi di righe; quindi il terzo pivot è $p_3 = -2$, e dobbiamo semplicemente sommare $1/2$ della terza riga alla quarta, ottenendo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & -2 & -1 & 8 \\ 0 & 0 & 1 & 4 & -2 \end{array}\right| \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 & 2 \\ 0 & 0 & -2 & -1 & 8 \\ 0 & 0 & 0 & 7/2 & 2 \end{array}\right|.$$

Il quarto pivot è $p_4 = 7/2$. Dunque il nostro sistema è equivalente anche al sistema

$$\begin{cases} x + 3y + z - w = 1, \\ y - z - w = 2, \\ -2z - w = 8, \\ \dfrac{7}{2}w = 2, \end{cases}$$

e risolvendo all'indietro troviamo di nuovo $x = 11, y = -12/7, z = -30/7, w = 4/7$. ⊔⊓

Osservazione 3.6

Nell'eliminazione di Gauss il Lemma 3.2 non viene usato a piena potenza. Infatti, ogni volta che un'equazione viene sostituita da una combinazione lineare il coefficiente dell'equazione sostituita (il k del Lemma 3.2) è sempre uguale a 1. Se il nostro obiettivo è semplicemente risolvere il sistema, non c'è motivo di limitarsi in questo modo: possiamo usare combinazioni lineari con $k \neq 0$ qualunque, e risolvere comunque il sistema (vedi l'Esempio 3.13). In seguito, però, dovremo applicare il metodo di eliminazione di Gauss anche per altri scopi (per esempio per calcolare il determinante di una matrice quadrata; vedi il Capitolo 9), e in quelle situazioni sarà fondamentale usare $k = 1$ e non un $k \neq 0$ qualunque.

Esempio 3.13

Riprendiamo di nuovo il sistema (3.10), e iniziamo l'eliminazione di Gauss come visto nell'Esempio 3.11. Procediamo come prima fino a ricavare il secondo pivot $p_2 = -5$. A questo punto invece di sommare alla quarta riga $1/5$ della seconda, ci discostiamo dall'eliminazione di Gauss usuale sostituendo alla quarta riga una combinazione lineare fatta dalla seconda riga più cinque volte la quarta (quindi $k = 5$ nel Lemma 3.2). Otteniamo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 1 & -1 & -1 & 2 \end{array}\right. \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & -2 & -1 & 8 \end{array}\right. .$$

Per concludere ora basta sommare il doppio della terza riga alla quarta ottenendo

$$\left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & -2 & -1 & 8 \end{array}\right. \longrightarrow \left|\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 1 \\ 0 & -5 & 3 & 4 & -2 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 7 & 4 \end{array}\right. .$$

Dunque il sistema (3.10) è equivalente anche al sistema triangolare superiore

$$\begin{cases} x + 3y + z - w = 1, \\ -5y + 3z + 4w = -2, \\ z + 4w = -2, \\ 7w = 4. \end{cases}$$

Procedendo in questo modo abbiamo semplificato i calcoli e ottenuto comunque un sistema triangolare superiore equivalente; ma, come vedremo nel Capitolo 9, abbiamo modificato alcune caratteristiche importanti della matrice dei coefficienti (per esempio il determinante), con il rischio, per giunta, di sbagliare (vedi l'Esercizio 3.27).

Nota che i pivot e le varie operazioni effettuate dipendono solo dalla matrice dei coefficienti; in particolare, quindi, potremo applicare l'eliminazione di Gauss semplicemente a una matrice quadrata e parlare di *pivot* di una matrice quadrata. Inoltre, il sistema triangolare superiore equivalente che troviamo alla fine dell'eliminazione di Gauss ha sulla diagonale principale esattamente i pivot. Quindi la Proposizione 3.1 ci fornisce un criterio per determinare quando un sistema lineare quadrato ammette un'unica soluzione.

Teorema 3.3

Un sistema lineare quadrato ammette un'unica soluzione se e solo se i pivot della sua matrice dei coefficienti sono tutti non nulli.

Dimostrazione

Infatti il sistema ammette un'unica soluzione se e solo se il sistema triangolare superiore a esso equivalente ottenuto tramite un'eliminazione di Gauss ammette un'unica soluzione, per cui la tesi segue dalla Proposizione 3.1. □

In particolare, quindi, l'unicità della soluzione dipende solo dalla matrice dei coefficienti, come avevamo suggerito nell'Osservazione 3.1.

Osservazione 3.7

Dagli esempi precedenti emerge che i pivot di una matrice quadrata *non sono univocamente determinati*, ma cambiano a seconda di quali scambi di riga vengono effettuati. Comunque, non sono completamente arbitrari; infatti, per il Teorema 3.3, se i pivot ottenuti con un'eliminazione di Gauss sono tutti non nulli, allora anche i pivot ottenuti con un'altra eliminazione di Gauss lo sono (perché?). Nel Capitolo 9 vedremo che il valore assoluto del prodotto di tutti i pivot è indipendente dall'eliminazione di Gauss effettuata (risulterà essere il valore assoluto del determinante della matrice).

Questo ci porta alla seguente definizione.

Definizione 3.8

Una matrice quadrata è *non singolare* se tutti i suoi pivot (rispetto a un'eliminazione di Gauss) sono non nulli; è *singolare* altrimenti.

Quindi il Teorema 3.3 si può anche esprimere dicendo che *un sistema quadrato ammette soluzione unica se e solo se la matrice dei coefficienti è non singolare.*

Osservazione 3.8

Un'altra conseguenza del metodo d'eliminazione di Gauss è che ogni sistema lineare quadrato è equivalente a un sistema triangolare superiore. Siccome sappiamo (Proposizione 3.1 ed Esercizio 3.1) che un sistema triangolare superiore o ha soluzione unica, o non ne ha o ne ha infinite, questo accade per qualunque sistema lineare quadrato.

Esercizi 3

3.1 Dimostra che se un sistema triangolare superiore ammette due soluzioni distinte allora ne ammette infinite.

3.2 Studia (cioè vedi se hanno soluzioni, e in tal caso trovale) i sistemi lineari:

$$\begin{cases} x+y-3z=0, \\ 3y-z=1, \\ 4z=8; \end{cases} \qquad \begin{cases} 2x_1-5x_2+x_3=2, \\ 2x_2-x_3=7, \\ 3x_3=-3. \end{cases}$$

3.3 Studia i sistemi lineari:

$$\begin{cases} x+y+z+w=6, \\ 2y+z-w=6, \\ 3z-2w=-3, \\ 2w=6; \end{cases} \qquad \begin{cases} x_1-x_2-2x_3+2x_4=3, \\ 8x_2+3x_3-11x_4=4, \\ 2x_3-9x_4=1, \\ \pi x_4=\pi. \end{cases}$$

3.4 Studia i sistemi lineari:

$$\begin{cases} x+y=3, \\ 3x-y=1; \end{cases} \qquad \begin{cases} 3x_1-x_2=2, \\ -2x_1+x_2=1; \end{cases} \qquad \begin{cases} \alpha+2\beta=5, \\ 2\alpha+\beta=-2. \end{cases}$$

3.5 Studia i sistemi lineari:

$$\begin{cases} x+y+z=3, \\ \sqrt{3}x+2y-z=0, \\ 3x-5y+3z=1; \end{cases} \qquad \begin{cases} 3x_1+x_2-x_3=2, \\ x_1+x_2=\sqrt{3}, \\ 5x_1-x_2-3x_3=1. \end{cases}$$

3.6 Studia i sistemi lineari:

$$\begin{cases} 3x+y-z=4, \\ x+y=3, \\ 5x-y-3z=0; \end{cases} \qquad \begin{cases} 3x_1-x_2-x_3=1, \\ -2x_1-x_2+x_3=-1, \\ 4x_1+2x_2-x_3=6. \end{cases}$$

3.7 Studia i sistemi lineari:

$$\begin{cases} x_1-2x_2+x_3=1, \\ x_1-x_2+4x_3=7, \\ x_1-2x_2+2x_3=3; \end{cases} \qquad \begin{cases} x+y+z=1, \\ 2x+2y+z=1, \\ 3y+z=1; \end{cases} \qquad \begin{cases} 5x_1+3x_2-2x_3=1, \\ x_2-2x_3=-2, \\ x_2+2x_3=1. \end{cases}$$

3.8 Studia i sistemi lineari:

$$\begin{cases} x+y-z+3w=1, \\ 2x+\sqrt{3}y+w=3, \\ 4x-y-z-w=0, \\ 7x+5y-z-w=3; \end{cases} \qquad \begin{cases} 3x_1+x_2+x_4=3, \\ x_1+2x_2+x_3+2x_4=1, \\ -2x_2+x_3+2x_4=6, \\ 4x_1+x_2+2x_3+5x_4=1. \end{cases}$$

3.9 Risolvi con l'eliminazione di Gauss i seguenti sistemi:

$$\begin{cases} 2x_2-4x_3+x_4=1, \\ x_1-3x_2-x_3+x_4=0, \\ x_1-x_2+4x_3-2x_4=-1, \\ 2x_1-2x_2-x_3+2x_4=0; \end{cases} \qquad \begin{cases} x+y+z+t=1, \\ 8x+4y+2z+t=5, \\ 27x+9y+3z+t=14, \\ 64x+16y+4z+t=30. \end{cases}$$

3.10 Studia i seguenti sistemi lineari:

$$\begin{cases} 19x-y+5z+t=3, \\ 18x+5z+t=1, \\ 6x+9y+t=1, \\ 12x+18y+3t=3; \end{cases} \qquad \begin{cases} 14x+3z+t=3, \\ 13x+y+3z+t=1, \\ 6x+8y+2t=1, \\ 6x+7y+t=1; \end{cases} \qquad \begin{cases} 9x+y+z+t=3, \\ 8x+2y+z+t=1, \\ 6x+7y+3t=1, \\ 6x+5y+t=1. \end{cases}$$

3.11 Studia i seguenti sistemi lineari al variare del parametro $k \in \mathbb{R}$:

$$\begin{cases} x+ky=1, \\ -kx+y=1; \end{cases} \qquad \begin{cases} x_1+kx_2=k, \\ kx_1+x_2=2k+3; \end{cases} \qquad \begin{cases} k\alpha+(k+1)\beta=1, \\ (k-1)\alpha+k\beta=k. \end{cases}$$

3.12 Studia i seguenti sistemi al variare del parametro $k \in \mathbb{R}$:

$$\begin{cases} 2x_1+kx_2=2, \\ kx_1+2x_2=k, \\ kx_2+kx_3=k; \end{cases} \qquad \begin{cases} 3x+2y+kz=11, \\ 2x-6y-3z=0, \\ kx+4y+2x=7; \end{cases} \qquad \begin{cases} y+z=k, \\ 2x+3y+7z=5, \\ x-3y-z=-2. \end{cases}$$

3.13 Studia il sistema lineare al variare del parametro $k \in \mathbb{R}$:

$$\begin{cases} x+ky+z+3w=5, \\ x+(k+2)y+(k+1)z+5w=k+5, \\ 2x+2(k-1)y+2z+7w=12-k, \\ 2x+2z+(2k+5)w=2k+9. \end{cases}$$

3.14 Stabilisci se le seguenti matrici sono singolari o meno:

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 2 & 2 \\ 5 & 1 & -1 \\ 4 & 1 & 2 \end{vmatrix}.$$

3.15 Determina se le seguenti matrici sono singolari o meno:

$$\begin{vmatrix} 1 & 0 & 3 & 7 \\ -1 & 1 & 0 & 2 \\ 2 & 2 & 0 & 2 \\ 3 & 1 & 4 & -1 \end{vmatrix}, \qquad \begin{vmatrix} 1 & -1 & 2 & -2 \\ 2 & -2 & 1 & -1 \\ 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 \end{vmatrix}.$$

3.16 Studia i sistemi seguenti al variare dei parametri a, b, $c \in \mathbb{R}$:

$$\begin{cases} x + 2y + az = 1, \\ 2x + ay + 8z = -1, \\ 4x + 7y + z = b; \end{cases} \qquad \begin{cases} 2x + y = a, \\ 2x + z = b, \\ 4x + y + z = c. \end{cases}$$

3.17 Scrivi un sistema lineare di tre equazioni in tre incognite, dipendente da un parametro k, in modo che se $k \neq 0$ allora il sistema ammette come unica soluzione la terna $(1, 0, 1)$, mentre se $k = 0$ il sistema ammette una retta di soluzioni.

3.18 Sia $Ax = b$ un sistema lineare quadrato con matrice A non singolare. Dimostra che il sistema è equivalente a un sistema lineare quadrato con matrice dei coefficienti diagonale.

3.19 Verifica che puoi trovare i pivot della seguente matrice senza bisogno di effettuare scambi di righe e calcolane il valore:

$$\begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 5 \\ 3 & 5 & 9 \end{vmatrix}.$$

3.20 Verifica che puoi trovare i pivot della seguente matrice senza bisogno di effettuare scambi di righe e calcolane il valore:

$$\begin{vmatrix} 2 & 3 & 1 \\ 1 & 5 & 2 \\ 1 & 12 & -3 \end{vmatrix}.$$

3.21 Calcola i pivot della matrice

$$A = \begin{vmatrix} 1 & 0 & 2 \\ 2 & 3 & -1 \\ 0 & 3 & 4 \end{vmatrix}$$

effettuando tutti gli scambi di riga possibili, e verifica che il valore assoluto del prodotto dei pivot non cambia. Quali combinazioni di scambi di righe danno come prodotto dei pivot l'opposto di quello ottenuto effettuando un'eliminazione di Gauss senza scambi di righe?

3.22 Calcola i pivot della matrice

$$\begin{vmatrix} 1 & 2 & 5 \\ 2 & 2 & 9 \\ 2 & 6 & -3 \end{vmatrix}$$

prima senza effettuare scambi di righe, e poi scambiando la prima e la terza riga. Infine, confronta la somma dei pivot che hai ottenuto nei due casi.

3.23 Determina per quali valori di $k \in \mathbb{R}$ le seguenti matrici sono singolari:

$$\begin{vmatrix} k & k+1 \\ k-1 & k \end{vmatrix}, \quad \begin{vmatrix} 2k+1 & k-10 \\ k+1 & k-7 \end{vmatrix}.$$

3.24 Trova per quali valori del parametro $k \in \mathbb{R}$ è singolare la matrice
$$\begin{vmatrix} 0 & -1 & 0 \\ 1 & 3 & 0 \\ -4 & -4 & k \end{vmatrix}.$$

3.25 Determina per quali valori di $k \in \mathbb{R}$ le seguenti matrici sono singolari:
$$\begin{vmatrix} 1 & -1 & 2 \\ k & 1 & 1 \\ 2k & 2 & 1 \end{vmatrix}, \quad \begin{vmatrix} k & 1 & -1 \\ k+1 & 0 & 1 \\ k & 3 & -1 \end{vmatrix}.$$

3.26 Determina per quali valori di $k \in \mathbb{R}$ le seguenti matrici sono singolari:
$$\begin{vmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & k & 4 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 5 & 3 \end{vmatrix}, \quad \begin{vmatrix} k & 1 & 3 & -2 \\ k & k & 1 & -1 \\ k & 2 & 0 & k+1 \\ k & 1 & 2 & -2 \end{vmatrix}.$$

3.27 Dato $\beta \in \mathbb{R}$, sia $A \in M_{3,3}(\mathbb{R})$ la matrice
$$A = \begin{vmatrix} \beta & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \end{vmatrix}.$$

Dimostra che A è non singolare per ogni β; pertanto, il sistema $Ax = O$ ha sempre come unica soluzione la terna $(0, 0, 0)$. Considera poi la seguente "eliminazione di Gauss generalizzata" (nel senso dell'Osservazione 3.6 e dell'Esempio 3.13).
$$\begin{vmatrix} \beta & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \end{vmatrix} \longrightarrow \begin{vmatrix} \beta & 0 & -1 \\ 0 & \beta & -\beta+1 \\ 0 & \beta & -\beta-1 \end{vmatrix} \longrightarrow \begin{vmatrix} \beta & 0 & -1 \\ 0 & \beta & -\beta+1 \\ 0 & 0 & -2 \end{vmatrix} = B,$$

dove prima abbiamo sostituito alla riga A_2 la combinazione lineare $\beta A_2 - A_1$ e alla riga A_3 la combinazione lineare $\beta A_3 + A_1$, e poi abbiamo sottratto la seconda riga dalla terza. Per il Lemma 3.2, il sistema $Bx = O$ è equivalente al sistema $Ax = O$; quindi ha come unica soluzione la terna $(0, 0, 0)$. Sfortunatamente, per $\beta = 0$ tutte le terne della forma $(x_1, x_2, 0)$ sono soluzione di $Bx = O$. Dov'è l'errore?

3.28 Dimostra per induzione le seguenti uguaglianze:
$$1 + \cdots + n = \frac{n(n+1)}{2}; \qquad 1^2 + \cdots + n^2 = \frac{2n^3 + 3n^2 + n}{6};$$
$$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \cdots + \frac{n}{2^n} = 2 - \frac{n+2}{2^n}.$$

3.29 Dimostra la seguente uguaglianza, dove q è un numero reale diverso da 1:
$$(1+q)(1+q^2)(1+q^4)\cdots(1+q^{2^n}) = \frac{1 - q^{2^{n+1}}}{1-q}.$$

3.30 Dimostra per induzione la seguente uguaglianza:
$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n \cdot (n+1)} = \frac{n}{n+1}.$$

Spazi vettoriali

4

Sommario

È giunto ora il momento di introdurre gli spazi vettoriali, la struttura principe dell'Algebra Lineare, di cui $\mathcal{V}_O^2$, $\mathcal{V}_O^3$ e l'insieme delle soluzioni di un sistema lineare omogeneo (cioè con vettore dei termini noti nullo) sono solo i primi esempi. Parleremo di spazi e sottospazi vettoriali, di dipendenza e indipendenza lineare, di basi e dimensione, di intersezione e somma, senza trascurare alcune conseguenze riguardanti i sistemi lineari. Studieremo poi un esempio particolare di spazio vettoriale reale, quello costituito dai numeri complessi. Infine, i Complementi a questo capitolo contengono una dimostrazione del Teorema fondamentale dell'Algebra.

4.1 Spazi e sottospazi

Abbiamo visto che le soluzioni di un sistema lineare con n incognite sono liste di n numeri reali: quei numeri che, sostituiti nell'ordine alle incognite, soddisfano tutte le equazioni del sistema. L'insieme delle liste di n numeri reali sarà indicato con $\mathbb{R}^n$:

$$\mathbb{R}^n = \left\{ v = \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} \;\middle|\; v_1, \ldots, v_n \in \mathbb{R} \right\}.$$

Ricordando quanto visto nel Paragrafo 2.2, dove a ogni punto del piano veniva associata una coppia di numeri reali e a ogni punto dello spazio una terna di numeri reali, l'insieme $\mathbb{R}^n$ può essere considerato come uno spazio geometrico n-dimensionale(1), in cui ogni punto è rappresentato da una n-upla di numeri reali. Gli elementi (che d'ora in poi chiameremo anche *vettori*) di $\mathbb{R}^n$ verranno scritti quasi sempre come colonne piuttosto che come righe di numeri, per motivi che vedremo nel Capitolo 5 (essenzialmente per poterli moltiplicare a sinistra per una matrice).

Osservazione 4.1

Forse ti starai chiedendo perché mai ci occupiamo di spazi n-dimensionali in un testo di Geometria, visto che lo spazio in cui viviamo è tridimensionale. Il punto è che anche per rappresentare sensatamente il comportamento di un oggetto nello spazio tre numeri non bastano. Per esempio, un punto materiale in movimento è descritto dalla sua posizione e dalla sua velocità, per cui servono un totale di 6 coordinate – e quindi siamo in $\mathbb{R}^6$. Per descrivere una squadra di calcio di punti materiali allora dobbiamo lavorare almeno in $\mathbb{R}^{66}$; ma anche lasciando perdere il calcio, nel corso dei tuoi studi incontrerai dozzine di situazioni che richiederanno l'uso di $\mathbb{R}^n$ con $n > 3$.

Fra le varie operazioni che si possono definire su $\mathbb{R}^n$ due ci interessano in modo particolare: la somma e il prodotto per un numero reale.

Definizione 4.1

La *somma* di due vettori di $\mathbb{R}^n$ è semplicemente la somma componente per componente

$$\forall v = \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} \in \mathbb{R}^n \quad \forall w = \begin{vmatrix} w_1 \\ \vdots \\ w_n \end{vmatrix} \in \mathbb{R}^n \qquad v + w = \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} + \begin{vmatrix} w_1 \\ \vdots \\ w_n \end{vmatrix} = \begin{vmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{vmatrix}.$$

Anche il prodotto per un numero reale (o, come diremo, il *prodotto per uno scalare*) è definito componente per componente

$$\forall \lambda \in \mathbb{R} \quad \forall v = \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} \in \mathbb{R}^n \qquad \lambda v = \lambda \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} = \begin{vmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{vmatrix};$$

non abbiamo fatto altro che applicare a $\mathbb{R}^n$ le definizioni viste per $\mathbb{R}^2$ e $\mathbb{R}^3$.

---

(1) Ma vedi il Paragrafo 4.4 per il significato della parola "dimensione" in questo contesto.

Ispirati dal Paragrafo 1.3, vediamo quali proprietà hanno la somma e il prodotto per scalari. Prima di tutto, $\mathbb{R}^n$ con la somma è un *gruppo commutativo* (come sicuramente avrai cura di verificare), il cui elemento neutro è il *vettore nullo*

$$O = \begin{vmatrix} 0 \\ \vdots \\ 0 \end{vmatrix},$$

e dove l'opposto di un vettore è dato da

$$-\begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} = \begin{vmatrix} -v_1 \\ \vdots \\ -v_n \end{vmatrix}.$$

Il prodotto per scalari invece soddisfa le seguenti proprietà

$$\forall \lambda \in \mathbb{R} \ \forall v, w \in \mathbb{R}^n \quad \lambda(v + w) = \lambda v + \lambda w;$$

$$\forall \lambda, \mu \in \mathbb{R} \ \forall v \in \mathbb{R}^n \quad (\lambda + \mu)v = \lambda v + \mu v;$$

$$\forall \lambda, \mu \in \mathbb{R} \ \forall v \in \mathbb{R}^n \quad (\lambda\mu)v = \lambda(\mu v);$$

$$\forall v \in \mathbb{R}^n \quad 1v = v \quad \text{e} \quad 0v = O$$

(sono tutte ovvie, ma tu dimostrale lo stesso, così da essere certo di averle capite).

L'insieme $\mathbb{R}^n$ non è l'unico insieme che conosciamo su cui sono definite due operazioni che godono delle stesse proprietà: abbiamo già incontrato $\mathcal{V}_O^2$ e $\mathcal{V}_O^3$, ma anche i polinomi in una (o più) variabili con le usuali operazioni di somma e prodotto per un numero reale le soddisfano. E allora tanto vale introdurre la seguente definizione.

Definizione 4.2

Uno *spazio vettoriale* (o *spazio lineare*) su $\mathbb{R}$ è un insieme V su cui sono definite due operazioni: una *somma* (che a due elementi di V associa un elemento di V) e un *prodotto per scalari* (che a un elemento di $\mathbb{R}$ e un elemento di V associa un elemento di V), soddisfacenti le proprietà seguenti:

(1) $\forall u, v, w \in V \quad (u + v) + w = u + (v + w)$ (*associatività della somma*);

(2) $\exists O \in V : \forall v \in V \quad v + O = O + v = v$ (*esiste l'elemento neutro per la somma*);

(3) $\forall v \in V \ \exists -v \in V : \quad v + (-v) = (-v) + v = O$ (*esiste l'opposto per la somma*);

(4) $\forall v, w \in V \quad v + w = w + v$ (*commutatività della somma*);

(5) $\left.\begin{array}{l} \forall \lambda \in \mathbb{R} \ \forall v, w \in V \ \lambda(v + w) = \lambda v + \lambda w; \\ \forall \lambda, \mu \in \mathbb{R} \ \forall v \in V \ (\lambda + \mu)v = \lambda v + \mu v; \end{array}\right\}$ (*distributività del prodotto per scalari*);

(6) $\forall \lambda, \mu \in \mathbb{R} \ \forall v \in V \ (\lambda\mu)v = \lambda(\mu v)$ (*associatività del prodotto per scalari*);

(7) $\forall v \in V \quad 1v = v \quad \text{e} \quad 0v = O$.

In particolare, V con la somma è un gruppo commutativo. Inoltre, se al posto di $\mathbb{R}$ sostituiamo ovunque un altro campo $\mathbb{K}$ abbiamo la definizione di *spazio vettoriale sul campo* $\mathbb{K}$. Gli elementi di uno spazio vettoriale sono detti *vettori*, e l'elemento neutro per la somma O si chiama *vettore nullo*.

Esempio 4.1

Abbiamo visto che $\mathcal{V}_O^2$ (Proposizioni 2.1 e 2.2), $\mathcal{V}_O^3$ (Esercizio 2.3), $\mathcal{V}^2$ (Paragrafo 2C.1) e $\mathbb{R}[t]$ (Esercizio 1.30) sono spazi vettoriali. Anche l'insieme $M_{m,n}(\mathbb{R})$ delle matrici $m \times n$ è uno spazio vettoriale su $\mathbb{R}$, con le operazioni

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{vmatrix} = \begin{vmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{vmatrix},$$

$$\lambda\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{vmatrix} = \begin{vmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{vmatrix}.$$

Se queste definizioni ti fanno sospettare che $M_{m,n}(\mathbb{R})$ e $\mathbb{R}^{mn}$ siano molto simili – semplicemente, uno è scritto in tabelle rettangolari e l'altro in colonna –, hai perfettamente ragione: come vedremo nel Capitolo 7, questi due spazi sono isomorfi, in maniera anche più stringente di quanto non lo fossero $\mathcal{V}_O^2$ e $\mathbb{R}^2$.

Esempio 4.2

Sia A un insieme qualunque, e V l'insieme di tutte le funzioni da A in $\mathbb{R}$. Definiamo una somma e un prodotto per scalari su V in questo modo: se $f, g \in V$ e $\lambda \in \mathbb{R}$, allora $f+g, \lambda f\colon A \to \mathbb{R}$ sono definite da

$$\forall a \in A \quad (f+g)(a) = f(a) + g(a) \quad \text{e} \quad (\lambda f)(a) = \lambda f(a).$$

Si verifica immediatamente (esercizio) che V con queste operazioni è uno spazio vettoriale. Se $A = \mathbb{R}$ possiamo considerare il sottoinsieme di V delle funzioni continue, o quello delle funzioni derivabili: con le stesse operazioni, sono anch'essi esempi di spazi vettoriali.

Negli spazi vettoriali ci sono dei sottoinsiemi piuttosto importanti, che sono l'analogo delle rette (e dei piani) per l'origine in $\mathcal{V}_O^3$.

Definizione 4.3

Un *sottospazio vettoriale* di uno spazio vettoriale V è un sottoinsieme $W \subseteq V$ chiuso rispetto alla somma e al prodotto per scalari, cioè tale che

$$\forall w, w_1, w_2 \in W \ \forall \lambda \in \mathbb{R} \qquad w_1 + w_2 \in W \quad \text{e} \quad \lambda w \in W.$$

Osservazione 4.2

Il vettore nullo O è contenuto in qualunque sottospazio vettoriale W: infatti, se w è un qualunque elemento di W, allora $O = 0w \in W$. In particolare, un sottoinsieme di uno spazio vettoriale che non contiene il vettore nullo non può essere un sottospazio vettoriale. Analogamente, se $v \in W$ allora $-v = (-1)v \in W$, per cui un sottospazio vettoriale contiene anche l'opposto di ogni suo elemento.

Osservazione 4.3

Un sottospazio vettoriale, considerato con la somma e il prodotto per scalari, è uno spazio vettoriale a tutti gli effetti (perché?), contenuto dentro uno spazio vettoriale più grande.

Ovviamente, uno spazio vettoriale V è sempre sottospazio di se stesso; all'estremo opposto, l'insieme $\{O\}$ costituito dal solo vettore nullo è anch'esso sempre un sottospazio. Altri esempi meno banali sono descritti qui di seguito.

Esempio 4.3

Fissiamo un riferimento affine in $\mathcal{A}^2$, in modo da identificarlo con $\mathbb{R}^2$. Sappiamo che $\{O\}$ e $\mathbb{R}^2$ sono sottospazi; vediamo quali sono gli altri. Sia V un sottospazio di $\mathbb{R}^2$; se $V \neq \{O\}$, deve contenere un vettore $v \neq O$. Allora, essendo chiuso rispetto al prodotto per scalari, V deve contenere tutti i multipli di v, cioè tutti i vettori della forma λv con $\lambda \in \mathbb{R}$. Ma ora l'insieme $\mathbb{R}v = \{\lambda v \mid \lambda \in \mathbb{R}\}$ (che geometricamente è una retta per l'origine) è un sottospazio di $\mathbb{R}^2$: infatti $\lambda v + \mu v = (\lambda + \mu)v$ e $\lambda(\mu v) = (\lambda\mu)v$, per cui $\mathbb{R}v$ è chiuso rispetto alla somma e al prodotto per scalari. Quindi le rette per l'origine sono sottospazi di $\mathbb{R}^2$.

Supponiamo ora che V sia ancora più grosso, cioè che oltre alla retta $\mathbb{R}v$ contenga anche un altro vettore $w \notin \mathbb{R}v$. Identificando v e w con vettori applicati nell'origine, vediamo subito che ogni vettore di $\mathbb{R}^2$ si può scrivere come somma di un multiplo di v e di un multiplo di w (per la Proposizione 2.3). Essendo V chiuso rispetto alla somma e al prodotto per scalari, questo vuol dire che $V = \mathbb{R}^2$. Pertanto, i sottospazi di $\mathbb{R}^2$ sono tutti e soli $\{O\}$, le rette per l'origine e $\mathbb{R}^2$ stesso.

Un altro importantissimo esempio di sottospazio vettoriale è fornito dai sistemi lineari.

Definizione 4.4

Un sistema lineare della forma $Ax = O$ (ovvero in cui i termini noti sono tutti nulli) è detto *omogeneo*. Se $Bx = b$ è un sistema lineare qualunque, il sistema $Bx = O$ è il *sistema omogeneo associato* al sistema $Bx = b$.

Chiaramente il vettore nullo O è soluzione di un sistema lineare omogeneo, ma potrebbe non essere l'unico, come capita per esempio nel sistema

$$\begin{cases} 2x_1 + x_2 = 0, \\ -4x_1 - 2x_2 = 0. \end{cases}$$

Proposizione 4.1

Sia $W \subseteq \mathbb{R}^n$ l'insieme delle soluzioni di un sistema lineare omogeneo in n incognite $Ax = O$. Allora W è un sottospazio vettoriale di $\mathbb{R}^n$.

Dimostrazione

Il nostro sistema è della forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ \qquad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0. \end{cases}$$

Siano v e w due soluzioni del sistema; dobbiamo dimostrare che anche $v + w$ e λv lo sono. Ma dire che v e w sono soluzioni del sistema equivale a dire che

$$\begin{cases} a_{11}v_1 + \cdots + a_{1n}v_n = 0, \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n = 0, \end{cases} \quad \text{e} \quad \begin{cases} a_{11}w_1 + \cdots + a_{1n}w_n = 0, \\ \vdots \\ a_{m1}w_1 + \cdots + a_{mn}w_n = 0. \end{cases}$$

Sommando otteniamo

$$\begin{cases} a_{11}(v_1 + w_1) + \cdots + a_{1n}(v_n + w_n) = 0, \\ \vdots \\ a_{m1}(v_1 + w_1) + \cdots + a_{mn}(v_n + w_n) = 0, \end{cases}$$

cioè $v + w$ è soluzione del sistema. Analogamente, moltiplicando per λ otteniamo

$$\begin{cases} a_{11}(\lambda v_1) + \cdots + a_{1n}(\lambda v_n) = 0, \\ \vdots \\ a_{m1}(\lambda v_1) + \cdots + a_{mn}(\lambda v_n) = 0, \end{cases}$$

cioè λv è soluzione del sistema. □

Osservazione 4.4

L'insieme delle soluzioni di un sistema lineare *non* omogeneo *non* è un sottospazio vettoriale, in quanto non contiene il vettore nullo.

4.2 Combinazioni lineari

Introduciamo ora un concetto fondamentale.

Definizione 4.5

Sia V uno spazio vettoriale; la *combinazione lineare* di k vettori $v_1, \ldots, v_k \in V$ con *coefficienti* (o *pesi*) $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ è il vettore

$$\alpha_1 v_1 + \cdots + \alpha_k v_k \in V.$$

Lo *span* dei (o *sottospazio generato* dai) vettori $v_1, \ldots, v_k$ è l'insieme di tutte le possibili combinazioni lineari di $v_1, \ldots, v_k$. In simboli

$$\operatorname{Span}(v_1, \ldots, v_k) = \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_1, \ldots, \alpha_k \in \mathbb{R}\}.$$

Per giustificarne il nome dobbiamo provare che lo span è davvero un sottospazio.

Proposizione 4.2

Sia V uno spazio vettoriale, e $v_1, \ldots, v_k$ vettori di V. Allora $\operatorname{Span}(v_1, \ldots, v_k)$ *è un sottospazio di V.*

Dimostrazione

Per dimostrare che è chiuso rispetto alla somma dobbiamo far vedere che la somma di due combinazioni lineari di $v_1, \ldots, v_k$ è ancora una combinazione lineare di $v_1, \ldots, v_k$. Infatti se $\alpha_1 v_1 + \cdots + \alpha_k v_k$ e $\beta_1 v_1 + \cdots + \beta_k v_k$ sono due combinazioni lineari di $v_1, \ldots, v_k$, anche

$$(\alpha_1 v_1 + \cdots + \alpha_k v_k) + (\beta_1 v_1 + \cdots + \beta_k v_k) = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_k + \beta_k)v_k$$

è una combinazione lineare di $v_1, \ldots, v_k$. Analogamente,

$$\lambda(\alpha_1 v_1 + \cdots + \alpha_k v_k) = (\lambda\alpha_1)v_1 + \cdots + (\lambda\alpha_k)v_k$$

è una combinazione lineare di $v_1, \ldots, v_k$ per ogni $\lambda \in \mathbb{R}$, e quindi $\mathrm{Span}(v_1, \ldots, v_k)$ è un sottospazio. □

Esempio 4.4

Se $v \in V$ è un vettore non nullo, allora

$$\mathrm{Span}(v) = \mathbb{R}v = \{\lambda v \mid \lambda \in \mathbb{R}\}$$

è la retta composta da tutti i multipli di v.

Esempio 4.5

Prendiamo $V = \mathbb{R}^3$ e

$$v_1 = \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix} \quad \text{e} \quad v_2 = \begin{vmatrix} -1 \\ -1 \\ 0 \end{vmatrix}.$$

Allora

$$\mathrm{Span}(v_1, v_2) = \left\{ \alpha_1 \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix} + \alpha_2 \begin{vmatrix} -1 \\ -1 \\ 0 \end{vmatrix} \;\middle|\; \alpha_1, \alpha_2 \in \mathbb{R} \right\} = \left\{ \begin{vmatrix} 3\alpha_1 - \alpha_2 \\ \alpha_1 - \alpha_2 \\ 0 \end{vmatrix} \;\middle|\; \alpha_1, \alpha_2 \in \mathbb{R} \right\}.$$

Geometricamente, $\mathrm{Span}(v_1, v_2)$ è il piano $z = 0$, cioè l'insieme di tutti i vettori della forma $\begin{vmatrix} b_1 \\ b_2 \\ 0 \end{vmatrix}$. Infatti è facile verificare che per ogni $b_1, b_2 \in \mathbb{R}$ esistono $\alpha_1, \alpha_2 \in \mathbb{R}$ tali che

$$\begin{cases} 3\alpha_1 - \alpha_2 = b_1, \\ \alpha_1 - \alpha_2 = b_2, \end{cases} \tag{4.1}$$

per cui

$$\begin{vmatrix} b_1 \\ b_2 \\ 0 \end{vmatrix} = \alpha_1 v_1 + \alpha_2 v_2 \in \mathrm{Span}(v_1, v_2);$$

per l'esattezza, risolvendo il sistema si ottiene $\alpha_1 = (b_1 - b_2)/2$ e $\alpha_2 = (b_1 - 3b_2)/2$.

Esempio 4.6

Scegliamo di nuovo $V = \mathbb{R}^3$, ma stavolta consideriamo i vettori

$$v_1 = \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix}, \quad v_2 = \begin{vmatrix} -1 \\ -1 \\ 0 \end{vmatrix} \quad \text{e} \quad v_3 = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}.$$

Di primo acchito troviamo

$$\mathrm{Span}(v_1, v_2, v_3) = \left\{ \begin{vmatrix} 3\alpha_1 - \alpha_2 + \alpha_3 \\ \alpha_1 - \alpha_2 \\ 0 \end{vmatrix} \;\middle|\; \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} \right\}.$$

Ma ora, ponendo $\beta_1 = \alpha_1 + \alpha_3/2$ e $\beta_2 = \alpha_2 + \alpha_3/2$. otteniamo

$$\begin{vmatrix} 3\alpha_1 - \alpha_2 + \alpha_3 \\ \alpha_1 - \alpha_2 \\ 0 \end{vmatrix} = \beta_1 \begin{vmatrix} 3 \\ 1 \\ 0 \end{vmatrix} + \beta_2 \begin{vmatrix} -1 \\ -1 \\ 0 \end{vmatrix},$$

per cui $\text{Span}(v_1, v_2, v_3) = \text{Span}(v_1, v_2)$.

Ricordiamo che un sistema lineare $Ax = b$ è *compatibile* se ammette soluzioni. Con la terminologia introdotta possiamo dare un criterio (non molto operativo ma che sarà utile in seguito) per decidere quando un sistema è compatibile.

Proposizione 4.3

Sia $Ax = b$ un sistema lineare in m equazioni ed n incognite, e siano $A^1, \dots, A^n \in \mathbb{R}^m$ le colonne della matrice dei coefficienti. Allora il sistema è compatibile se e solo se $b \in \text{Span}(A^1, \dots, A^n)$.

Dimostrazione

Infatti il sistema si può scrivere come

$$x_1 A^1 + \cdots + x_n A^n = b \tag{4.2}$$

(convincitene prima di proseguire; è importante). Quindi il sistema è compatibile se e solo se b è combinazione lineare delle colonne della matrice dei coefficienti, cioè se e solo se $b \in \text{Span}(A^1, \dots, A^n)$. □

4.3 Indipendenza lineare e basi

Nel Paragrafo 2.2 avevamo parlato di basi di $\mathcal{V}_O^2$ e $\mathcal{V}_O^3$; in questo paragrafo vedremo come recuperare questo concetto fondamentale nel contesto degli spazi vettoriali.

Definizione 4.6

Un insieme di vettori $v_1, \dots, v_k \in V$ di uno spazio vettoriale V si dice *linearmente dipendente* se esistono $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ *non tutti nulli* tali che

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = O. \tag{4.3}$$

Viceversa, $v_1, \dots, v_k$ si dicono *linearmente indipendenti* se non sono linearmente dipendenti, ovvero se $\alpha_1 v_1 + \cdots + \alpha_k v_k = O$ implica $\alpha_1 = \cdots = \alpha_k = 0$. Un'eguaglianza della forma (4.3) con gli α_j non tutti nulli sarà chiamata *relazione di dipendenza lineare* fra $v_1, \dots, v_k$.

Esempio 4.7

Se fra i vettori $v_1, \dots, v_k \in V$ vi è il vettore nullo, allora $v_1, \dots, v_k$ sono linearmente dipendenti. Infatti se $v_i = O$ allora

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_k = O$$

è una relazione di dipendenza lineare fra $v_1, \dots, v_k$.

Esempio 4.8

Se fra i vettori $v_1, \ldots, v_k \in V$ ve ne sono due uguali, allora $v_1, \ldots, v_k$ sono linearmente dipendenti. Infatti se si ha $v_i = v_j$ con $i < j$ allora

$$0v_1 + \cdots + 0v_{i-1} + v_i + 0v_{i+1} + \cdots + 0v_{j-1} + (-1)v_j + 0v_{j+1} + \cdots + 0v_k = O$$

è una relazione di dipendenza lineare fra $v_1, \ldots, v_k$.

Esempio 4.9

I vettori v_1 e $v_2 \in \mathbb{R}^5$ dell'Esempio 4.5 sono linearmente indipendenti. Infatti, la relazione $\alpha_1 v_1 + \alpha_2 v_2 = O$ vale se e solo se

$$\begin{cases} 3\alpha_1 - \alpha_2 = 0, \\ \alpha_1 - \alpha_2 = 0, \end{cases}$$

e questo sistema ha come unica soluzione $\alpha_1 = \alpha_2 = 0$.

Esempio 4.10

I vettori v_1, v_2, $v_3 \in \mathbb{R}^3$ dell'Esempio 4.6 sono invece linearmente dipendenti; infatti, $v_1 + v_2 - 2v_3 = O$.

Per vedere se i vettori $v_1, \ldots, v_k \in \mathbb{R}^n$ sono linearmente indipendenti si forma la matrice $A = |v_1 \ \cdots \ v_k| \in M_{n,k}(\mathbb{R})$ e si studia il sistema $Ax = O$: i vettori sono dipendenti se e solo se il sistema ha soluzioni diverse da $x = O$.

Proposizione 4.4

Il sistema $Ax = O$ in n incognite ha soluzione unica (quella banale $x = O$) se e solo se le colonne $A^1, \ldots, A^n$ sono linearmente indipendenti.

Dimostrazione

Per (4.2), una soluzione è data da $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che

$$\alpha_1 A^1 + \cdots + \alpha_n A^n = O.$$

Quindi esiste una soluzione diversa da $x = O$ se e solo se le colonne di A sono linearmente dipendenti. □

Osservazione 4.5

Dire che i vettori $v_1, \ldots, v_k \in V$ sono linearmente dipendenti equivale a dire che è possibile ricavare uno di loro come combinazione lineare degli altri. Infatti, se abbiamo $\alpha_1 v_1 + \cdots + \alpha_k v_k = O$ con, per esempio, $\alpha_1 \neq 0$, allora

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \cdots - \frac{\alpha_k}{\alpha_1} v_k.$$

In particolare, $\mathrm{Span}(v_1, \ldots, v_k) = \mathrm{Span}(v_2, \ldots, v_k)$ (perché? Ricorda l'Esempio 4.6), e quindi v_1 non serve per descrivere il sottospazio generato da $v_1, \ldots, v_k$; gli altri sono sufficienti.

Questo ci suggerisce la nozione di base di uno spazio vettoriale.

Definizione 4.7

Sia V uno spazio vettoriale. Un insieme $\mathcal{B} = \{v_1, \ldots, v_n\}$ di vettori di V è una *base* di V se:

a) $V = \mathrm{Span}(v_1, \ldots, v_n)$, cioè $v_1, \ldots, v_n$ sono un *sistema di generatori* di V;

b) $v_1, \ldots, v_n$ sono linearmente indipendenti.

Esempio 4.11

I vettori $e_1, \ldots, e_n \in \mathbb{R}^n$, dove

$$e_1 = \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}, \; e_2 = \begin{vmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{vmatrix}, \; \ldots, \; e_n = \begin{vmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{vmatrix},$$

sono una base di $\mathbb{R}^n$, la *base canonica*.

Esempio 4.12

Sia

$$V = \left\{ \begin{vmatrix} b_1 \\ b_2 \\ 0 \end{vmatrix} \;\middle|\; b_1, b_2 \in \mathbb{R} \right\}$$

il piano $z = 0$ in $\mathbb{R}^3$. Allora i vettori $\{v_1, v_2\}$ dell'Esempio 4.5 sono una base di V; infatti, abbiamo visto nell'Esempio 4.5 che sono un sistema di generatori di V, e nell'Esempio 4.9 che sono linearmente indipendenti. Non sono l'unica base di V; anche $\{e_1, e_2\} \subset \mathbb{R}^3$ lo è, e non è difficile trovarne infinite altre.

Esempio 4.13

Consideriamo lo spazio vettoriale (Esercizio 4.5) $V = \mathbb{R}_3[t]$ dei polinomi in una variabile a coefficienti reali di grado minore o uguale a 3. Allora

$$p_0 = 1, \, p_1 = t, \, p_2 = t^2, \, p_3 = t^3$$

formano una base di V. Infatti sono un sistema di generatori: un generico polinomio $p = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$ si scrive come combinazione lineare di $p_0, \ldots, p_3$ nel modo più naturale, cioè $p = \alpha_0 p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$. Sono anche linearmente indipendenti: infatti un polinomio è identicamente nullo se e solo se tutti i suoi coefficienti sono zero. Lo stesso ragionamento ci dice che

$$p_0 = 1, \, p_1 = t, \, \ldots, \, p_n = t^n$$

è una base di $\mathbb{R}_n[t]$, lo spazio dei polinomi in una variabile a coefficienti reali di grado minore o uguale a n.

Esempio 4.14

Se i vettori $v_1, \ldots, v_k$ di uno spazio vettoriale V sono linearmente indipendenti, allora sono una base del sottospazio $\mathrm{Span}(v_1, \ldots, v_k) \subseteq V$.

L'importanza delle basi consiste nel fatto che se $\{v_1, \ldots, v_n\}$ è una base di V allora ogni elemento di V può essere scritto *in un solo modo* come combinazione lineare dei vettori $v_1, \ldots, v_n$.

Proposizione 4.5

Siano $v_1, \ldots, v_k$ vettori linearmente indipendenti di uno spazio vettoriale V. Supponiamo che $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{R}$ siano tali che

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \beta_1 v_1 + \cdots + \beta_k v_k. \tag{4.4}$$

Allora $\alpha_j = \beta_j$ per $j = 1, \ldots, k$.

Dimostrazione

L'uguaglianza (4.4) implica che $(\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_k - \beta_k)v_k = 0$; la tesi segue dall'indipendenza lineare dei v_j. □

Sia $\{v_1, \ldots, v_n\}$ una base di uno spazio vettoriale V, e prendiamo $v \in V$. Siccome la base è un sistema di generatori di V, esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$; siccome i v_j sono linearmente indipendenti, per la Proposizione 4.5 gli $\alpha_1, \ldots, \alpha_n$ sono univocamente determinati.

Definizione 4.8

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di uno spazio vettoriale V, e $v \in V$. Gli n numeri reali $(\alpha_1, \ldots, \alpha_n)$ tali che $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ sono le *coordinate* di v *rispetto alla base* $\mathcal{B}$.

Pertanto, *uno spazio vettoriale* (per esempio, l'insieme delle soluzioni di un sistema lineare omogeneo) *è completamente determinato una volta che se ne conosca una base:* infatti ogni suo elemento si ottiene in uno e un sol modo come combinazione lineare degli elementi della base, con le coordinate come coefficienti.

Esempio 4.15

Se $x \in \mathbb{R}^n$ possiamo scrivere

$$x = \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix} = x_1 e_1 + \cdots + x_n e_n,$$

per cui $(x_1, \ldots, x_n)$ sono le coordinate di x rispetto alla base canonica.

Esempio 4.16

Fissiamo un riferimento affine $RA(O, \vec{\imath}, \vec{\jmath})$ di $\mathcal{V}^2_O$; allora $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}\}$ è una base di $\mathcal{V}^2_O$. Infatti la Proposizione 2.3 ci dice che $\mathcal{B}$ è un sistema di generatori, e l'indipendenza lineare (Esercizio 4.18) segue dal fatto che $\vec{\imath}$ e $\vec{\jmath}$ non sono proporzionali. Dunque le coordinate che avevamo introdotto nel Paragrafo 2.2 non sono altro che un caso particolare delle coordinate appena definite.

Esempio 4.17

Sia V lo spazio vettoriale dell'Esempio 4.12, e prendiamo

$$w = \begin{vmatrix} 2 \\ 0 \\ 0 \end{vmatrix} \in V.$$

Per trovare le coordinate di w rispetto alla base $\mathcal{B}_1 = \{v_1, v_2\}$ dobbiamo risolvere il sistema $w = \alpha_1 v_1 + \alpha_2 v_2$, cioè

$$\begin{cases} 3\alpha_1 - \alpha_2 = 2, \\ \alpha_1 - \alpha_2 = 0. \end{cases}$$

La soluzione è $\alpha_1 = \alpha_2 = 1$, per cui le coordinate di w rispetto a $\mathcal{B}_1$ sono $(1, 1)$. Invece, nota che w ha (perché?) coordinate $(2, 0)$ rispetto alla base $\mathcal{B}_2 = \{e_1, e_2\}$.

Osservazione 4.6

Per stabilire se $v_1, \ldots, v_k \in \mathbb{R}^n$ sono un sistema di generatori di $\mathbb{R}^n$, si considera di nuovo la matrice $A = \begin{vmatrix} v_1 & \cdots & v_k \end{vmatrix} \in M_{n,k}(\mathbb{R})$; i vettori dati sono un sistema di generatori se e solo se il sistema $Ax = b$ ha soluzione per ogni $b \in \mathbb{R}^n$. Questa è una tecnica un po' scomoda; un approccio migliore è descritto nell'Esercizio 4.23, e vedremo un sistema davvero efficace nel Paragrafo 6.3.

4.4 Esistenza delle basi

A questo punto una domanda sorge spontanea: ogni spazio vettoriale ha una base? Risposta: no[(2)], non è neanche detto che abbia un sistema finito (cioè composto da un numero finito di vettori) di generatori; per esempio, $\mathbb{R}[t]$ non ce l'ha (Esercizio 4.25). Come vedremo, questa è l'unica difficoltà: ogni qual volta esiste un sistema finito di generatori, esiste anche una base finita.

Definizione 4.9

Sia $\mathcal{A} \subseteq V$ un sottoinsieme di uno spazio vettoriale V. Diremo che $\mathcal{B} \subseteq \mathcal{A}$ è un sottoinsieme *massimale in $\mathcal{A}$ di vettori linearmente indipendenti* se gli elementi di $\mathcal{B}$ sono linearmente indipendenti e aggiungendo a $\mathcal{B}$ un qualunque altro elemento di $\mathcal{A}$ si ottiene un insieme di vettori linearmente dipendenti.

Esempio 4.18

Sia V il solito spazio vettoriale dell'Esempio 4.12, e

$$\mathcal{A} = \left\{ v_1, v_2, \begin{vmatrix} 2 \\ 0 \\ 0 \end{vmatrix}, e_1 \right\} \subset V.$$

Allora $\mathcal{B}_1 = \{v_1, v_2\}$ o $\mathcal{B}_2 = \{e_1, v_1\}$ sono esempi di sottoinsiemi massimali in $\mathcal{A}$ di vettori linearmente indipendenti.

(2) Solo perché abbiamo richiesto che una base contenga un numero finito di vettori. Definendo opportunamente quando un insieme infinito di vettori è una base allora si può dimostrare che ogni spazio vettoriale ha una base.

Una base (se esiste) è un tipico caso di insieme massimale di vettori linearmente indipendenti (vedi anche l'Esercizio 4.19).

Proposizione 4.6

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di uno spazio vettoriale V. Allora $\mathcal{B}$ è un insieme massimale in V di vettori linearmente indipendenti.

Dimostrazione

Per definizione i vettori $v_1, \ldots, v_n$ sono linearmente indipendenti; dobbiamo dimostrare che, per ogni $v \in V$, i vettori $v, v_1, \ldots, v_n$ sono dipendenti. Ma $\mathcal{B}$ è un sistema di generatori; quindi esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Quindi si ha $v - \alpha_1 v_1 - \cdots - \alpha_n v_n = O$, cioè una relazione di dipendenza lineare fra $v, v_1, \ldots, v_n$ (il coefficiente di v è 1, che è non nullo). □

Viceversa, un sottoinsieme massimale di vettori linearmente indipendenti in un sistema di generatori è una base. Per dimostrarlo, ci serve un piccolo lemma.

Lemma 4.7

Sia $\mathcal{B}$ un sottoinsieme finito di uno spazio vettoriale V. Se $\mathrm{Span}(\mathcal{B})$ *contiene un sistema di generatori di V allora $V =$* $\mathrm{Span}(\mathcal{B})$*, cioè anche $\mathcal{B}$ è un sistema di generatori di V.*

Dimostrazione

Sia $\mathcal{A}$ un sistema di generatori di V contenuto in $\mathrm{Span}(\mathcal{B})$, e sia v un qualunque elemento di V. Essendo $\mathcal{A}$ un sistema di generatori di V, possiamo scrivere v come combinazione lineare di elementi di $\mathcal{A}$. Siccome $\mathcal{A} \subset \mathrm{Span}(\mathcal{B})$, possiamo scrivere gli elementi di $\mathcal{A}$ come combinazione lineare di elementi di $\mathcal{B}$; mettendo il tutto insieme, possiamo scrivere v come combinazione lineare di elementi di $\mathcal{B}$, per cui $v \in \mathrm{Span}(\mathcal{B})$. Essendo v un vettore qualsiasi di V, abbiamo fatto vedere che $\mathcal{B}$ è un sistema di generatori di V. □

Teorema 4.8

Sia $\mathcal{A} = \{v_1, \ldots, v_k\}$ un sistema di generatori di uno spazio vettoriale V. Sia $\mathcal{B} \subseteq \mathcal{A}$ un sottoinsieme massimale in $\mathcal{A}$ di vettori linearmente indipendenti. Allora $\mathcal{B}$ è una base di V.

Dimostrazione

Scriviamo $\mathcal{B} = \{v_1, \ldots, v_r\}$ con $r \leq k$ (cosa sempre possibile, a meno di cambiare l'ordine in cui abbiamo elencato gli elementi di $\mathcal{A}$). Essendo gli elementi di $\mathcal{B}$ linearmente indipendenti per ipotesi, dobbiamo solo dimostrare che $V = \mathrm{Span}(\mathcal{B})$, sapendo che $\mathcal{A}$ è un sistema di generatori.

Grazie al Lemma 4.7, è sufficiente far vedere che $\mathcal{A} \subseteq \mathrm{Span}(\mathcal{B})$. Sia w un qualunque elemento di $\mathcal{A}$ che non stia già in $\mathcal{B}$. Per l'ipotesi di massimalità, $v_1, \ldots, v_r, w$ sono linearmente dipendenti; quindi esistono $\alpha_1, \ldots, \alpha_r, \beta \in \mathbb{R}$ non tutti nulli tali che

$$\alpha_1 v_1 + \cdots + \alpha_r v_r + \beta w = O.$$

Ora, se fosse $\beta = 0$ avremmo scritto una relazione di dipendenza lineare fra i vettori $v_1, \ldots, v_r$ che sappiamo essere linearmente indipendenti; quindi dev'essere per forza $\beta \neq 0$. Ma allora

$$w = -\frac{\alpha_1}{\beta} v_1 - \cdots - \frac{\alpha_r}{\beta} v_r \in \mathrm{Span}(\mathcal{B}),$$

ed essendo w generico, otteniamo $\mathcal{A} \subseteq \mathrm{Span}(\mathcal{B})$. □

Siccome ogni insieme finito di generatori contiene (perché?) un insieme massimale di vettori linearmente indipendenti, abbiamo dimostrato il seguente corollario.

Corollario 4.9

Ogni spazio vettoriale contenente un sistema finito di generatori ammette una base.

Rimane il problema di come trovare una base di un dato spazio vettoriale. Il Teorema 4.8 suggerisce una prima procedura. Sia $\mathcal{A} = \{v_1, \ldots, v_k\}$ un sistema di generatori di V. Se $V = \{O\}$ abbiamo finito; altrimenti, $\mathcal{A}$ contiene un elemento non nullo, diciamo v_1. Se v_1 e v_j sono linearmente dipendenti qualunque sia $v_j \in \mathcal{A}$, allora $\{v_1\}$ è un sottoinsieme massimale in $\mathcal{A}$ di vettori linearmente indipendenti, per cui è la base cercata. Altrimenti, ci sarà un altro vettore di $\mathcal{A}$, diciamo v_2, tale che v_1 e v_2 siano linearmente indipendenti. Se $\{v_1, v_2\}$ è un sottoinsieme massimale in $\mathcal{A}$ di vettori linearmente indipendenti, abbiamo trovato la nostra base; altrimenti ci sarà un terzo vettore di $\mathcal{A}$, diciamo v_3, tale che v_1, v_2 e v_3 siano linearmente indipendenti. Si procede così fino a trovare la base voluta. Chiaramente, non è un metodo molto efficace; tecniche migliori le studieremo nel Paragrafo 6.3.

Esempio 4.19

Consideriamo lo spazio vettoriale

$$V = \operatorname{Span}\left(v_1 = \begin{vmatrix} 1 \\ 2 \\ 0 \\ 2 \end{vmatrix}, v_2 = \begin{vmatrix} 2 \\ 1 \\ 1 \\ 0 \end{vmatrix}, v_3 = \begin{vmatrix} 5 \\ 4 \\ 2 \\ 2 \end{vmatrix} \right) \subset \mathbb{R}^4;$$

vogliamo una base di V. Prima di tutto, $V \neq \{O\}$; quindi possiamo cominciare a costruire la nostra base prendendo v_1, che è un vettore non nullo. Verifichiamo se v_1 e v_2 sono linearmente indipendenti; il sistema $\alpha_1 v_1 + \alpha_2 v_2 = O$ ammette solo la soluzione $\alpha_1 = \alpha_2 = 0$ (come si vede facilmente considerando per esempio le prime due equazioni), per cui v_1 e v_2 sono linearmente indipendenti. Rimane da vedere cosa succede aggiungendo v_3. Il sistema $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = O$ ammette la soluzione $\alpha_1 = 1$, $\alpha_2 = 2$ e $\alpha_3 = -1$ (come si vede risolvendo il sistema composto dalle prime tre equazioni e poi sostituendo il risultato nella quarta), per cui v_1, v_2 e v_3 sono linearmente dipendenti. Questo vuol dire che $\{v_1, v_2\}$ è un sottoinsieme massimale di vettori linearmente indipendenti in un sistema di generatori, e quindi una base di V. ⊓⊔

Ora, è facile vedere che ogni spazio vettoriale ha un'infinità di basi distinte (esercizio: se $\{v_1, \ldots, v_n\}$ è una base di V, anche $\{2v_1, v_2, \ldots, v_n\}$ lo è). Ma due diverse basi di uno stesso spazio vettoriale hanno una caratteristica fondamentale in comune: hanno sempre lo stesso numero di elementi. Questa è una conseguenza del seguente (e importante) *Teorema del completamento.*

Teorema 4.10 (del completamento)

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di uno spazio vettoriale V, e siano $w_1, \ldots, w_p \in V$ (con $p \leq n$) vettori linearmente indipendenti. Allora esistono $n - p$ vettori di $\mathcal{B}$ che insieme a $w_1, \ldots, w_p$ formano una base di V.

Dimostrazione

Procediamo per induzione su p. Sia $p = 1$; siccome $\mathcal{B}$ è una base di V, esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che

$$w_1 = \alpha_1 v_1 + \ldots + \alpha_n v_n. \tag{4.5}$$

Nel caso $p = 1$, l'ipotesi d'indipendenza lineare si riduce a $w_1 \neq O$; quindi almeno uno degli α_j è non nullo. A meno dell'ordine, possiamo supporre $\alpha_1 \neq 0$ ricavando

$$v_1 = \frac{1}{\alpha_1} w_1 - \frac{\alpha_2}{\alpha_1} v_2 - \cdots - \frac{\alpha_n}{\alpha_1} v_n \in \mathrm{Span}(w_1, v_2, \ldots, v_n).$$

Dunque $\mathcal{B} \subset \mathrm{Span}(w_1, v_2, \ldots, v_n)$ per cui, grazie al Lemma 4.7, $\{w_1, v_2, \ldots, v_n\}$ è un sistema di generatori di V; dimostriamo ora che è composto da vettori linearmente indipendenti.

Supponiamo che si abbia $\beta_1 w_1 + \beta_2 v_2 + \cdots + \beta_n v_n = O$ per certi $\beta_1, \ldots, \beta_n \in \mathbb{R}$; dobbiamo dimostrare che $\beta_1 = \cdots = \beta_n = 0$. Usando la (4.5) otteniamo

$$\begin{aligned} O &= \beta_1(\alpha_1 v_1 + \cdots + \alpha_n v_n) + \beta_2 v_2 + \cdots + \beta_n v_n \\ &= (\beta_1 \alpha_1) v_1 + (\beta_1 \alpha_2 + \beta_2) v_2 + \cdots + (\beta_1 \alpha_n + \beta_n) v_n. \end{aligned}$$

Siccome $v_1, \ldots, v_n$ sono linearmente indipendenti per ipotesi, i coefficienti di questa combinazione lineare devono essere tutti nulli; essendo $\alpha_1 \neq 0$ ne segue che $\beta_1 = 0$ e quindi $\beta_2 = \cdots = \beta_n = 0$. In conclusione, $w_1, v_2, \ldots, v_n$ sono linearmente indipendenti, e il caso $p = 1$ è dimostrato.

Supponiamo ora la tesi vera per $p - 1$ vettori, e dimostriamola per p vettori. Per ipotesi, possiamo trovare $n - (p - 1)$ elementi di $\mathcal{B}$ (che, a meno dell'ordine, possiamo supporre essere $v_p, \ldots, v_n$) tali che $\{w_1, \ldots, w_{p-1}, v_p, \ldots, v_n\}$ sia una base di V. Procediamo come prima. Devono esistere $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che

$$w_p = \alpha_1 w_1 + \cdots + \alpha_{p-1} w_{p-1} + \alpha_p v_p + \cdots + \alpha_n v_n.$$

Non tutti gli α_j con $j = p, \ldots, n$ possono essere nulli, perché altrimenti avremmo scritto una relazione di dipendenza lineare fra $w_1, \ldots, w_p$; a meno dell'ordine possiamo supporre che $\alpha_p \neq 0$. Dunque

$$v_p = \frac{1}{\alpha_p} w_p - \frac{\alpha_1}{\alpha_p} w_1 - \cdots - \frac{\alpha_{p-1}}{\alpha_p} w_{p-1} - \frac{\alpha_{p+1}}{\alpha_p} v_{p+1} - \cdots - \frac{\alpha_n}{\alpha_p} v_n$$

appartiene a $\mathrm{Span}(w_1, \ldots, w_p, v_{p+1}, \ldots, v_n)$. Di nuovo, questo implica che i vettori $w_1, \ldots, w_p, v_{p+1}, \ldots, v_n$ formano un sistema di generatori, per il Lemma 4.7; rimane da dimostrare che sono anche linearmente indipendenti.

Siano $\beta_1, \ldots, \beta_n \in \mathbb{R}$ tali che $\beta_1 w_1 + \cdots + \beta_p w_p + \beta_{p+1} v_{p+1} + \cdots + \beta_n v_n = O$. Procedendo come nel caso $p = 1$, troviamo

$$\begin{aligned} O &= \beta_1 w_1 + \cdots + \beta_{p-1} w_{p-1} + \beta_p(\alpha_1 w_1 + \cdots + \alpha_{p-1} w_{p-1} + \alpha_p v_p + \cdots + \alpha_n v_n) \\ &\quad + \beta_{p+1} v_{p+1} + \cdots + \beta_n v_n \\ &= \gamma_1 w_1 + \cdots + \gamma_{p-1} w_{p-1} + (\beta_p \alpha_p) v_p + \gamma_{p+1} v_{p+1} + \cdots + \gamma_n v_n, \end{aligned}$$

dove abbiamo posto $\gamma_j = \beta_j + \beta_p \alpha_j$. Per ipotesi $w_1, \ldots, w_{p-1}, v_p, \ldots, v_n$ sono linearmente indipendenti; quindi (essendo $\alpha_p \neq 0$) deduciamo $\beta_p = 0$, e dunque tutti i β_j sono zero. Pertanto $w_1, \ldots, w_p, v_{p+1}, \ldots, v_n$ sono linearmente indipendenti, come volevamo. $\square$

Quando applichiamo questo teorema diremo che *completiamo* $\{w_1, \ldots, w_p\}$ *a una base di* V.

Osservazione 4.7

Se $w_1, \ldots, w_p$ sono vettori linearmente indipendenti di $\mathbb{R}^n$, per completarli a una base basta aggiungervi $n-p$ vettori della base canonica. Infatti, basta estrarre una base dall'insieme $\{w_1, \ldots, w_p, e_1, \ldots, e_n\}$ come indicato nel prossimo esempio, o, meglio, usare le tecniche che vedremo nel Paragrafo 6.3.

Esempio 4.20

Prendiamo i vettori

$$w_1 = \begin{vmatrix} 1 \\ 0 \\ 1 \\ 0 \end{vmatrix}, \quad w_2 = \begin{vmatrix} 1 \\ 2 \\ 3 \\ 4 \end{vmatrix} \in \mathbb{R}^4;$$

vogliamo dimostrare che sono linearmente indipendenti e poi completare $\{w_1, w_2\}$ a una base di $\mathbb{R}^4$. Per vedere che sono linearmente indipendenti dobbiamo risolvere il sistema $\alpha_1 w_1 + \alpha_2 w_2 = O$, cioè

$$\begin{cases} \alpha_1 + \alpha_2 = 0, \\ 2\alpha_2 = 0, \\ \alpha_1 + 3\alpha_2 = 0, \\ 4\alpha_2 = 0, \end{cases}$$

che chiaramente ammette come unica soluzione $\alpha_1 = \alpha_2 = 0$, per cui i nostri due vettori sono linearmente indipendenti. Per completare $\{w_1, w_2\}$ a una base di $\mathbb{R}^4$ procediamo come nella dimostrazione del Teorema 4.10. Cominciamo con lo scrivere w_1 come combinazione lineare dei vettori della base canonica di $\mathbb{R}^4$:

$$w_1 = 1e_1 + 0e_2 + 1e_3 + 0e_4. \tag{4.6}$$

La dimostrazione del Teorema 4.10 ci dice che possiamo mettere w_1 al posto di uno qualunque dei vettori della base canonica che compaiono con coefficiente non nullo in (4.6), ottenendo ancora una base di $\mathbb{R}^4$. Per esempio, mettendo w_1 al posto di e_1 otteniamo la base $\mathcal{B}_1 = \{w_1, e_2, e_3, e_4\}$. Adesso scriviamo w_2 come combinazione lineare dei vettori di $\mathcal{B}_1$

$$w_2 = 1w_1 + 2e_2 + 2e_3 + 4e_4. \tag{4.7}$$

La dimostrazione del Teorema 4.10 ci dice che adesso possiamo mettere w_2 al posto di uno qualunque dei vettori di $\mathcal{B}_1$ (w_1 escluso, ovviamente) che compaiono con un coefficiente non nullo in (4.7). Per esempio, sostituendo w_2 al vettore e_3 otteniamo la base $\{w_1, w_2, e_2, e_4\}$, che completa $\{w_1, w_2\}$ come volevamo.

E ora una conseguenza fondamentale del precedente teorema.

Corollario 4.11

Due basi di uno spazio vettoriale V hanno sempre lo stesso numero di elementi.

Dimostrazione

Siano $\mathcal{B} = \{v_1, \ldots, v_k\}$ e $\mathcal{C} = \{w_1, \ldots, w_h\}$ due basi di V e supponiamo per assurdo che $k > h$. Allora il Teorema 4.10 applicato con $n = k$ e $p = h$ ci dice che esistono $k - h$ elementi di $\mathcal{B}$ che aggiunti a $\mathcal{C}$ formano una base. Ma $\mathcal{C}$ era già una base, e quindi un insieme *massimale* di vettori linearmente indipendenti (Proposizione 4.6), il che è una contraddizione. Un analogo ragionamento esclude la possibilità che h sia maggiore di k; quindi $h = k$. □

Dunque il numero di elementi di una base è una caratteristica ben definita dello spazio vettoriale: la sua dimensione.

Definizione 4.10

Se $\{v_1, \ldots, v_n\}$ è una base di uno spazio vettoriale V, il numero n (che per il Corollario 4.11 non dipende dalla base scelta), si chiama *dimensione* di V; scriveremo $n = \dim V$ (oppure $n = \dim_{\mathbb{K}} V$, se è necessario specificare che V è uno spazio vettoriale sul campo $\mathbb{K}$). Lo spazio vettoriale $\{O\}$ composto dal solo vettore nullo ha dimensione zero. Infine, diremo che uno spazio vettoriale privo di sistemi di generatori finiti ha *dimensione infinita;* ma in questo libro ci occuperemo esclusivamente di spazi vettoriali di dimensione finita.

Per calcolare la dimensione di uno spazio vettoriale basta trovarne una base. Per esempio, grazie alla base canonica (Esempio 4.11) vediamo subito che $\mathbb{R}^n$ ha dimensione n; in particolare, una retta ha dimensione 1, un piano ha dimensione 2 e così via, coerentemente con la nozione intuitiva di dimensione.

Osservazione 4.8

In un certo senso, la dimensione di uno spazio vettoriale V è il numero dei parametri necessari per descriverlo. Infatti, una volta scelta una base, ogni elemento di V è combinazione lineare degli $n = \dim V$ *fissati* elementi della base con coefficienti le n *generiche* coordinate. Per chiarire il concetto ecco un esempio.

Esempio 4.21

Consideriamo l'insieme $V = \{A \in M_{2,2}(\mathbb{R}) \mid a_{11} + a_{22} = 0\}$; è facile verificare (esercizio) che V è un sottospazio di $M_{2,2}(\mathbb{R})$. Il generico elemento di V è della forma

$$A = \begin{vmatrix} a & b \\ c & -a \end{vmatrix}, \tag{4.8}$$

per cui dipende da tre parametri. Questo suggerisce l'idea che V abbia dimensione 3; per dimostrarlo, cerchiamone una base. Se vogliamo che i parametri a, b, c siano le coordinate rispetto a una base, la matrice generica $A \in V$ si deve poter scrivere come $A = aM_1 + bM_2 + cM_3$, dove $\{M_1, M_2, M_3\}$ è la base che stiamo cercando. Guardando (4.8) è chiaro che questo è possibile solo prendendo

$$M_1 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \qquad M_2 = \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}, \qquad M_3 = \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix}.$$

Per costruzione, $\{M_1, M_2, M_3\}$ è un sistema di generatori di V; rimane da vedere che sono anche linearmente indipendenti. Ma la relazione di dipendenza lineare data da $\alpha_1 M_1 + \alpha_2 M_2 + \alpha_3 M_3 = O$ implica subito (perché?) $\alpha_1 = \alpha_2 = \alpha_3 = 0$, e ci siamo. ⊓⊔

Ecco qualche altro corollario del Teorema 4.10.

Corollario 4.12

Sia V uno spazio vettoriale di dimensione n. Allora ogni n-upla di vettori linearmente indipendenti di V è una base.

Dimostrazione

Basta applicare il Teorema 4.10 con $p = n$. □

Corollario 4.13

Sia V uno spazio vettoriale di dimensione n, e $w_1, \ldots, w_p \in V$. Se $p > n$, allora $w_1, \ldots, w_p$ sono linearmente dipendenti.

Dimostrazione

Se $w_1, \ldots, w_n$ sono linearmente dipendenti, abbiamo finito. Altrimenti, per il Corollario 4.12 sono una base di V e quindi un insieme massimale in V di vettori linearmente indipendenti (Proposizione 4.6); ma in tal caso i vettori $w_1, \ldots, w_n, w_{n+1}, \ldots, w_p$ devono necessariamente essere linearmente dipendenti. □

Osservazione 4.9

Dunque la dimensione di uno spazio vettoriale è, in un certo senso, il massimo numero di vettori linearmente indipendenti; analogamente si può vedere (Esercizio 4.21) che è il minimo numero di vettori in un sistema di generatori.

Concludiamo questo paragrafo confrontando la dimensione di un sottospazio con quella dello spazio che lo contiene. Il risultato è prevedibile.

Proposizione 4.14

Sia V uno spazio vettoriale di dimensione n, e W un suo sottospazio. Allora:
(1) W ha dimensione finita minore o uguale a n;
(2) $\dim W = \dim V$ se e soltanto se $W = V$.

Dimostrazione

(1) Se $W = \{O\}$ abbiamo finito. Altrimenti, sia w_1 un elemento non nullo di W. Se $\{w_1\}$ non è un insieme massimale in W di vettori linearmente indipendenti, troviamo un $w_2 \in W$ tale che w_1 e w_2 sono linearmente indipendenti. Se $\{w_1, w_2\}$ non è un insieme massimale in W di vettori linearmente indipendenti, troviamo un $w_3 \in W$ tale che w_1, w_2 e w_3 sono linearmente indipendenti. Continuando così, prima o poi dovremo fermarci, perché in V (che contiene W) non possono esistere più di n elementi linearmente indipendenti fra loro (Corollario 4.13). Quindi troveremo un insieme $\{w_1, \ldots, w_m\}$ (con $m \leq n$) massimale in W di vettori linearmente indipendenti, ovvero una base di W (Esercizio 4.19).
(2) Se $\dim W = n$, esiste una base di W composta da n elementi. Per il Corollario 4.12 questa è anche una base di V e quindi $W = V$. Il viceversa è ovvio. □

Osservazione 4.10

Se avessimo saputo a priori che il sottospazio W aveva un sistema finito di generatori – e quindi una base finita – il Corollario 4.13 ci avrebbe permesso subito di concludere che $\dim W \leq \dim V$. Tutto il punto della dimostrazione della Proposizione 4.14 è esattamente far vedere che ogni sottospazio di uno spazio vettoriale di dimensione finita ammette un sistema finito di generatori.

4.5 Somma e intersezione di sottospazi

Sia V uno spazio vettoriale, e U e W suoi sottospazi. Possiamo allora costruire:

a) il sottospazio *intersezione* $U \cap W$;

b) il sottospazio *somma* $U + W = \{u + w \mid u \in U, w \in W\}$.

Ovviamente in entrambi i casi bisogna controllare che si tratti effettivamente di sottospazi. La verifica per l'intersezione è un esercizio per te; vediamo invece il caso della somma. Siano $u_1 + w_1$, $u_2 + w_2 \in U + W$; allora

$$(u_1 + w_1) + (u_2 + w_2) = (u_1 + u_2) + (w_1 + w_2) \in U + W$$

e

$$\lambda(u_1 + w_1) = (\lambda u_1) + (\lambda w_1) \in U + W$$

qualunque sia $\lambda \in \mathbb{R}$, per cui $U + W$ è un sottospazio.

Osservazione 4.11

Vale la pena notare che, in generale, l'unione insiemistica $U \cup W$ *non* è un sottospazio. Per esempio, se $V = \mathbb{R}^2$ e U e W sono due distinte rette per l'origine, $U \cup W$ non è chiuso rispetto alla somma, per cui non è un sottospazio di $\mathbb{R}^2$.

Un problema naturale è se sia possibile calcolare la dimensione di $U + W$ e $U \cap W$ a partire dalle dimensioni di U e W. Per rispondere ci serve il seguente lemma.

Lemma 4.15

Siano U, W sottospazi dello spazio vettoriale V. Se $\mathcal{B}$ è un sistema di generatori di U e $\mathcal{C}$ un sistema di generatori di W, allora $\mathcal{B} \cup \mathcal{C}$ è un sistema di generatori di $U + W$.

Dimostrazione

Ogni elemento di U (rispettivamente, W) è combinazione lineare di vettori di $\mathcal{B}$ (rispettivamente $\mathcal{C}$). Quindi il generico elemento di $U + W$ si scrive come somma di una combinazione lineare di vettori di $\mathcal{B}$ e di una combinazione lineare di vettori di $\mathcal{C}$ – ovvero come combinazione lineare di vettori di $\mathcal{B} \cup \mathcal{C}$. □

Osservazione 4.12

Se $\mathcal{B}$ è una base di U e $\mathcal{C}$ è una base di W, non è detto che $\mathcal{B} \cup \mathcal{C}$ sia una base di $U + W$, in quanto i vettori di $\mathcal{B} \cup \mathcal{C}$ potrebbero essere linearmente dipendenti (Esercizio 4.30).

La differenza fra $\dim U + \dim W$ e $\dim(U + W)$ è data dalla dimensione dell'intersezione. Questo è il contenuto dell'importante *Teorema di Grassmann*.

Teorema 4.16 (Grassmann)

Siano U e W sottospazi di uno spazio vettoriale di dimensione finita V. Allora

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Dimostrazione

Sia $\{v_1, \ldots, v_p\}$ una base di $U \cap W$, dove $p = \dim(U \cap W)$. Usando il Teorema 4.10, possiamo completarla in U a una base $\{v_1, \ldots, v_p, u_{p+1}, \ldots, u_r\}$ di U, con $r = \dim U$, e in W a una base $\{v_1, \ldots, v_p, w_{p+1}, \ldots, w_s\}$ di W, dove $s = \dim W$. Allora ci basta provare che $\mathcal{B} = \{v_1, \ldots, v_p, u_{p+1}, \ldots, u_r, w_{p+1}, \ldots, w_s\}$ è una base di $U + W$; infatti, in tal caso avremo

$$\dim(U + W) = p + (r - p) + (s - p) = r + s - p = \dim U + \dim W - \dim(U \cap W),$$

come voluto. Il Lemma 4.15 ci assicura che $\mathcal{B}$ è un sistema di generatori di $U + W$; dimostriamo ora che sono linearmente indipendenti. Supponiamo allora che

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_{p+1} u_{p+1} + \cdots + \beta_r u_r + \gamma_{p+1} w_{p+1} + \cdots + \gamma_s w_s = O; \tag{4.9}$$

dobbiamo provare che tutti gli α_i, i β_j e i γ_k sono zero. Consideriamo i tre vettori

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_p v_p \in U \cap W, \\ u &= \beta_{p+1} u_{p+1} + \cdots + \beta_r u_r \in U, \\ w &= \gamma_{p+1} w_{p+1} + \cdots + \gamma_s w_s \in W. \end{aligned}$$

La nostra ipotesi è $v + u + w = O$; questo implica che $w = -u - v \in U$, in quanto $u \in U$ e $v \in U \cap W \subseteq U$. Ne segue che $w \in U \cap W$, per cui si deve poter scrivere come combinazione lineare dei v_j, diciamo $w = \delta_1 v_1 + \cdots + \delta_p v_p$. Grazie alla definizione di w si ha $\gamma_{p+1} w_{p+1} + \cdots + \gamma_s w_s = \delta_1 v_1 + \cdots + \delta_p v_p$; essendo i vettori $v_1, \ldots, v_p, w_{p+1}, \ldots, w_s$ linearmente indipendenti, ciò implica $\gamma_{p+1} = \ldots = \gamma_s = \delta_1 = \ldots = \delta_p = 0$. Ci siamo: infatti ora (4.9) ci dice che $\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_{p+1} u_{p+1} + \cdots + \beta_r u_r = O$, e l'indipendenza lineare di $v_1, \ldots, v_p, u_{p+1}, \ldots, u_r$ implica che $\alpha_1 = \ldots = \alpha_p = \beta_{p+1} = \ldots = \beta_r = 0$. □

Osservazione 4.13

Per trovare una base di $U + W$, basta unire una base di U a una di W, e poi estrarre un insieme massimale di vettori linearmente indipendenti. Vedremo nel Paragrafo 6.3 un metodo generale per farlo – assieme anche a un modo per trovare una base di $U \cap W$.

Se U e W sono sottospazi di uno spazio vettoriale V tali che $U \cap W = \{O\}$, il Teorema di Grassmann ci dice che $\dim(U + W) = \dim U + \dim W$. In questo caso scriveremo $U \oplus W$ e diremo che $U \oplus W$ è la *somma diretta* di U e W.

Definizione 4.11

Due sottospazi U e W di uno spazio vettoriale V si dicono *supplementari* (e l'uno un *supplementare* dell'altro) se $V = U \oplus W$.

Questa è una situazione importante, perché in questo caso ogni vettore di V si scrive in modo unico come somma di un vettore di U e di uno di W.

Proposizione 4.17

Siano U e W due sottospazi di uno spazio vettoriale V tali che $U \cap W = \{O\}$, e prendiamo u_1, $u_2 \in U$ e w_1, $w_2 \in W$. Allora $u_1 + w_1 = u_2 + w_2$ se e solo se $u_1 = u_2$ e $w_1 = w_2$.

Dimostrazione

Infatti $u_1 + w_1 = u_2 + w_2$ implica $u_1 - u_2 = w_2 - w_1 \in U \cap W$; la tesi segue allora da $U \cap W = \{O\}$. □

I supplementari esistono sempre.

Proposizione 4.18

Ogni sottospazio U di uno spazio vettoriale V di dimensione finita ha un supplementare.

Dimostrazione

Sia $\{u_1, \dots, u_p\}$ una base di U; completiamola a una base $\{u_1, \dots, u_p, w_{p+1}, \dots, w_n\}$ di V e poniamo $W = \text{Span}(w_{p+1}, \dots, w_n)$. Chiaramente si ha $V = U + W$; quindi ci rimane solo da dimostrare che $U \cap W = \{O\}$. Ma $\dim U = p$ e $\dim W = n - p$ per costruzione; quindi il Teorema 4.16 ci dice che

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) = p + (n - p) - n = 0,$$

da cui $U \cap W = \{O\}$. □

Osservazione 4.14

Questa dimostrazione fornisce anche un modo per trovare un supplementare di un sottospazio dato U. Prendi una base $\{u_1, \dots, u_p\}$ di U e completala a una base $\{u_1, \dots, u_p, w_{p+1}, \dots, w_n\}$ di V; allora $W = \text{Span}(w_{p+1}, \dots, w_n)$ è un supplementare. In particolare, il supplementare di U non è certo unico (perché?).

4.6 Numeri complessi

Facciamo ora una breve digressione per affrontare un problema che forse ti sarai posto: perché non abbiamo tentato di definire su $\mathbb{R}^n$ un prodotto vero e proprio, invece di limitarci al prodotto per scalari? Il motivo è che per $n \geq 3$ *non* esiste alcun prodotto che, assieme alla somma usuale, renda $\mathbb{R}^n$ un campo(3). In $\mathbb{R}^2$ invece un tale prodotto esiste, e si ottiene il campo $\mathbb{C}$ dei numeri complessi, di cui studieremo qui le prime proprietà. Per definirlo, esaminiamo più in dettaglio la struttura dei vettori del piano.

Definizione 4.12

Sia $v = (x, y) \in \mathbb{R}^2$ un vettore non nullo (visto come vettore applicato nell'origine). Il *modulo* (o *lunghezza*) di v è il numero positivo

$$\rho = |v| = \sqrt{x^2 + y^2}.$$

L'*argomento* di v è invece quell'angolo $\theta \in [0, 2\pi)$, misurato in radianti, tale che si abbia $x = \rho \cos\theta$ e $y = \rho \sin\theta$, cioè (vedi la Figura 4.1)

$$\cos\theta = \frac{x}{\sqrt{x^2 + y^2}} \qquad \text{e} \qquad \sin\theta = \frac{y}{\sqrt{x^2 + y^2}};$$

in particolare, se $x \neq 0$ si ha $\tan\theta = y/x$. La coppia (ρ, θ) ci dà le *coordinate polari* del vettore (o punto del piano) $v \in \mathbb{R}^2$.

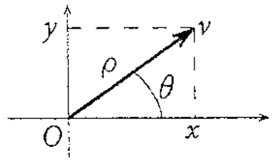

Figura 4.1 - Le coordinate polari.

(3) La dimostrazione di questo risultato, come puoi immaginare, è tutt'altro che semplice.

Osservazione 4.15

Il modulo dell'origine è 0, ma il suo argomento non è univocamente determinato; di conseguenza, le coordinate polari non sono definite nell'origine. Inoltre, se all'argomento di $z = (x, y) \in \mathbb{R}^2$ sommiamo un multiplo intero di 2π, otteniamo un numero θ' che soddisfa ancora le condizioni $x = \rho\cos\theta'$ e $y = \rho\sin\theta'$; con un lieve abuso di linguaggio, anche un tale θ' sarà detto "argomento di z".

Esempio 4.22

Se $v = (2, -2)$, allora v ha modulo $\rho = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}$ e argomento $\theta = 7\pi/4$ (in quanto si deve avere $\tan\theta = -1$ e $\cos\theta > 0$). ⊔⊓

Un vettore v_1 di coordinate polari $(\rho_1, \theta_1) \in \mathbb{R}^+ \times [0, 2\pi)$ ha coordinate cartesiane

$$(x_1, y_1) = (\rho_1\cos\theta_1, \rho_1\sin\theta_1) = \rho_1(\cos\theta_1, \sin\theta_1).$$

In un certo senso, le coordinate polari suddividono il vettore in una componente scalare, il modulo, e in una componente di rotazione, l'argomento. Questo suggerisce un'idea. Possiamo fare agire un vettore v_1 su un altro vettore v_2 ruotando v_2 di un angolo θ_1 e moltiplicandone la lunghezza per ρ_1, in modo da tenere conto di entrambe le componenti di v_1 (vedi la Figura 4.2). Se v_2 ha coordinate polari (ρ_2, θ_2) possiamo quindi definire un possibile prodotto fra v_1 e v_2 ponendo

$$\begin{aligned} v_1 \cdot v_2 &= \rho_1\rho_2\big(\cos(\theta_1 + \theta_2), \sin(\theta_1 + \theta_2)\big) \\ &= \rho_1\rho_2\big(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2, \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2\big). \end{aligned} \tag{4.10}$$

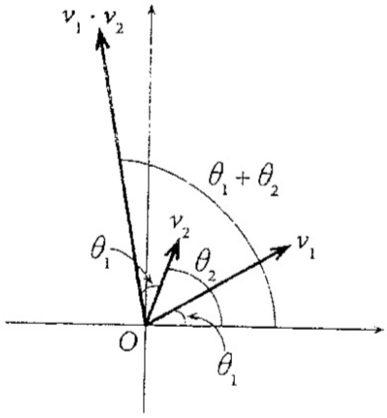

Figura 4.2 - Il prodotto in coordinate polari.

In coordinate cartesiane, (4.10) si scrive

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1). \tag{4.11}$$

Questo è esattamente il prodotto che cercavamo.

Proposizione 4.19

$(\mathbb{R}^2, +, \cdot)$ *è un campo.*

Dimostrazione

Che $\mathbb{R}^2$ con la somma usuale sia un gruppo commutativo è ormai assodato. Le proprietà distributive si verificano (esercizio) immediatamente guardando la (4.11), che mostra anche che il vettore $(1,0)$ è l'elemento neutro per il prodotto. L'associatività del prodotto segue dall'associatività degli ordinari somma e prodotto, grazie alla (4.10), o la possiamo anche verificare direttamente usando la (4.11) come segue

$$\begin{aligned}
[(x_1,y_1)\cdot(x_2,y_2)]\cdot(x_3,y_3) &= (x_1x_2-y_1y_2, x_1y_2+x_2y_1)\cdot(x_3,y_3)\\
&= \big((x_1x_2-y_1y_2)x_3-(x_1y_2+x_2y_1)y_3, (x_1x_2-y_1y_2)y_3+x_3(x_1y_2+x_2y_1)\big)\\
&= \big(x_1(x_2x_3-y_2y_3)-y_1(x_2y_3+x_3y_2), x_1(x_2y_3+x_3y_2)+y_1(x_2x_3-y_2y_3)\big)\\
&= (x_1,y_1)\cdot(x_2x_3-y_2y_3, x_2y_3+x_3y_2) = (x_1,y_1)\cdot[(x_2,y_2)\cdot(x_3,y_3)].
\end{aligned}$$

L'inverso di (x,y) deve avere modulo inverso e argomento opposto, come si vede da (4.10); quindi, essendo $\cos(-\theta)=\cos\theta$ e $\sin(-\theta)=-\sin\theta$, l'inverso è

$$(x,y)^{-1}=\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right),$$

come si verifica anche con (4.11). Infine, la commutatività del prodotto è ovvia. □

Definizione 4.13

Il campo $\mathbb{C}$ dei *numeri complessi* è $\mathbb{R}^2$ con la somma usuale e il prodotto appena definito.

Geometricamente è naturale identificare $\mathbb{R}$ con l'asse delle ascisse, cioè considerare l'applicazione $T\colon\mathbb{R}\to\mathbb{C}$ data da $T(x)=(x,0)$. Guardando (4.11) risulta evidente che T è un *isomorfismo* fra $\mathbb{R}$ e l'asse delle ascisse, nel senso che conserva tutte le operazioni: si ha

$$T(x+y)=T(x)+T(y) \quad \text{e} \quad T(xy)=T(x)\cdot T(y)$$

per ogni x, $y\in\mathbb{R}$. Pertanto d'ora in poi considereremo $\mathbb{R}$ come un sottoinsieme di $\mathbb{C}$, identificandolo con l'asse delle ascisse tramite T, e scriveremo spesso x al posto di $(x,0)$.

Il campo $\mathbb{C}$, essendo $\mathbb{R}^2$ sotto mentite spoglie, è uno spazio vettoriale di dimensione 2 su $\mathbb{R}$. La base canonica di $\mathbb{C}$ su $\mathbb{R}$ è data da $\{(1,0),(0,1)\}$. Il vettore $(1,0)$ è l'unità moltiplicativa di $\mathbb{C}$, e come appena detto verrà indicato semplicemente con 1. Il vettore $(0,1)$, detto *unità immaginaria*, verrà invece indicato sempre con i. Nota che

$$i^2=i\cdot i=(0,1)\cdot(0,1)=(-1,0)=-1$$

(infatti se ruotiamo i di $\pi/2$ radianti arriviamo su -1), cioè i è una soluzione dell'equazione $x^2+1=0$, che su $\mathbb{R}$ non ne aveva nessuna! Si verifica facilmente che un'altra soluzione è $-i$.

Utilizzando le due unità 1 e i possiamo scrivere i numeri complessi in una forma più comoda per i calcoli. Infatti un generico numero complesso (a,b) può venire ora espresso come

$$(a,b)=(a,0)+(0,b)=a(1,0)+b(0,1)=a\cdot 1+b\cdot i=a+ib,$$

che è il modo con cui spesso vengono presentati i numeri complessi. Questa notazione è comoda perché per fare calcoli con essa basta utilizzare le solite proprietà (associativa, distributiva, eccetera) e ricordarsi che $i^2=-1$. Per esempio, ritroviamo le formule per il prodotto e l'inverso come segue:

$$(a+ib)(c+id)=ac+iad+ibc+i^2bd=(ac-bd)+i(ad+bc); \tag{4.12}$$

$$\frac{1}{a+ib}=\frac{a-ib}{(a+ib)(a-ib)}=\frac{a-ib}{a^2-iab+iab-i^2b^2}=\frac{a-ib}{a^2+b^2}. \tag{4.13}$$

Definizione 4.14

Se $z = a + ib \in \mathbb{C}$ è un numero complesso, la *parte reale* $\operatorname{Re} z$ di z è il numero reale a, mentre la *parte immaginaria* $\operatorname{Im} z$ di z è il numero reale b. Se $\operatorname{Re} z = 0$ (cioè se $z = ib$), il numero complesso z si dice *immaginario puro.* Infine, il *modulo* di z è il numero reale $|z| = \sqrt{a^2 + b^2}$.

La formula dell'inverso invita a considerare anche un'altra definizione.

Definizione 4.15

Il *complesso coniugato* di un numero complesso $z = a + ib \in \mathbb{C}$ è il numero complesso $\overline{z} = a - ib$. L'applicazione $\overline{} : \mathbb{C} \to \mathbb{C}$ che associa a un numero complesso il suo coniugato si chiama *coniugio.*

Nelle prossime due proposizioni raccogliamo le proprietà di coniugio e modulo.

Proposizione 4.20

Siano $z_0, z_1, z_2 \in \mathbb{C}$ numeri complessi. Allora:

(1) $\overline{\overline{z_0}} = z_0$;

(2) $z_0 + \overline{z_0} = 2 \operatorname{Re} z_0$;

(3) $z_0 - \overline{z_0} = 2i \operatorname{Im} z_0$;

(4) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$;

(5) $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$;

(6) $\overline{(1/z_0)} = 1/\overline{z_0}$ se $z_0 \neq 0$;

(7) $\overline{z_0} = z_0$ se e solo se $z_0 \in \mathbb{R}$;

(8) $\overline{z_0} = -z_0$ *se e solo se z_0 è immaginario puro;*

(9) $z_0\overline{z_0} = |z_0|^2 \geq 0$, e $z_0\overline{z_0} = 0$ se e solo se $z_0 = 0$;

(10) $z_0^{-1} = \overline{z_0}/|z_0|^2$.

Dimostrazione

(1), (2), (3) e (4) seguono immediatamente dalla definizione. La (5) si ottiene notando che il coniugio consiste nel cambiare segno alla parte immaginaria (cioè alla y), e guardando quindi la (4.11). La (6) è conseguenza della (5), mentre (7) e (8) sono di nuovo ovvie. Per la (9), se $z_0 = a + ib$ si ha

$$z_0\overline{z_0} = (a + ib)(a - ib) = a^2 + b^2 + i(-ab + ba) = a^2 + b^2,$$

e ci siamo. Infine la (10) segue dalla (4.13). □

Proposizione 4.21

Siano $z_0, z_1, z_2 \in \mathbb{C}$ numeri complessi. Allora:

(1) $|z_0| \geq 0$, *e* $|z_0| = 0$ *se e solo se* $z_0 = 0$;

(2) $|\overline{z_0}| = |z_0|$;

(3) $|\operatorname{Re} z_0| \leq |z_0|$, $|\operatorname{Im} z_0| \leq |z_0|$ e $|z_0| \leq |\operatorname{Re} z_0| + |\operatorname{Im} z_0|$;

(4) $|z_1 z_2| = |z_1| \cdot |z_2|$;

(5) $|1/z_0| = 1/|z_0|$ se $z_0 \neq 0$;

(6) $|z_1 + z_2| \leq |z_1| + |z_2|$ (*disuguaglianza triangolare*);
(7) $|z_1 + z_2| \geq \big||z_1| - |z_2|\big|$.

Dimostrazione

La (1) segue dalla Proposizione 4.20.(9), mentre la (2) è ovvia. Se $z_0 = a + ib$, la (3) dice semplicemente che

$$|a| \leq \sqrt{a^2+b^2}, \qquad |b| \leq \sqrt{a^2+b^2}, \qquad \sqrt{a^2+b^2} \leq |a| + |b|,$$

formule che si verificano subito elevando al quadrato.

Per la (4) basta notare che $|z_1z_2|^2 = (z_1z_2)(\overline{z_1z_2}) = (z_1\overline{z_1})(z_2\overline{z_2}) = |z_1|^2|z_2|^2$. La (5) segue subito dalla (4); rimangono la (6) e la (7).

Poniamo $z_j = x_j + iy_j$ per $j = 1, 2$. Elevando al quadrato, vediamo che (6) e (7) valgono se e solo se

$$\big||z_1| - |z_2|\big|^2 \leq |z_1 + z_2|^2 \leq \big(|z_1| + |z_2|\big)^2,$$

cioè se e solo se

$$\begin{aligned}(x_1^2 + y_1^2) + (x_2^2 + y_2^2) - 2\sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} &\leq (x_1 + x_2)^2 + (y_1 + y_2)^2 \\ &\leq (x_1^2 + y_1^2) + (x_2^2 + y_2^2) + 2\sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)}.\end{aligned}$$

Sviluppando i conti vediamo che questo accade se e solo se

$$-\sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} \leq x_1x_2 + y_1y_2 \leq \sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)},$$

che equivale a $|x_1x_2 + y_1y_2| \leq \sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)}$. A sua volta, elevando al quadrato vediamo che questo succede se e solo se

$$x_1^2x_2^2 + y_1^2y_2^2 + 2x_1x_2y_1y_2 \leq x_1^2x_2^2 + x_1^2y_2^2 + y_1^2x_2^2 + y_1^2y_2^2,$$

cioè se e solo se $0 \leq (x_1y_2 - x_2y_1)^2$, che è vero, ed è fatta. □

Osservazione 4.16

Le definizioni sono state scelte in modo che $|w_2 - w_1|$ sia la distanza euclidea fra i punti w_1 e w_2 del piano (come si vede facilmente). Quindi se nella Proposizione 4.21.(6) prendiamo $z_1 = w_3 - w_2$ e $z_2 = w_2 - w_1$ otteniamo

$$|w_3 - w_1| \leq |w_3 - w_2| + |w_2 - w_1|,$$

per cui nel triangolo di vertici w_1, w_2 e w_3 la lunghezza di un lato è minore o uguale alla somma delle lunghezze degli altri due. Per questo motivo la Proposizione 4.21.(6) è chiamata *disuguaglianza triangolare*.

4.7 Potenze e radici

Abbiamo visto che la relazione fra coordinate polari e coordinate cartesiane è espressa dalla formula

$$(x, y) = \rho(\cos\theta, \sin\theta);$$

allora un numero complesso $z = x + iy$ si può scrivere in *forma trigonometrica* come

$$z = |z|(\cos\theta + i\sin\theta). \tag{4.14}$$

Il prodotto e il quoziente si scrivono bene in forma trigonometrica: se $z_j \in \mathbb{C}$ ha coordinate polari (ρ_j, θ_j) per $j = 1, 2$, allora

$$\begin{aligned} z_1 z_2 &= \rho_1\rho_2\big(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\big), \\ \frac{z_1}{z_2} &= \frac{\rho_1}{\rho_2}\big(\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)\big). \end{aligned} \tag{4.15}$$

Una conseguenza importante della (4.15) è la *formula di De Moivre.*

Lemma 4.22 (De Moivre)

Sia $z = \rho(\cos\theta + i\sin\theta) \in \mathbb{C}$. Allora

$$\forall n \in \mathbb{N} \qquad z^n = \rho^n[\cos(n\theta) + i\sin(n\theta)].$$

Dimostrazione

Usiamo l'induzione su n. Se $n = 1$ la formula è ovvia. Sia vera per $n-1$; allora z^{n-1} ha modulo ρ^{n-1} e argomento $(n-1)\theta$, e la (4.15) dà

$$\begin{aligned} z^n = z\cdot z^{n-1} &= \rho\rho^{n-1}\big[\cos\big(\theta+(n-1)\theta\big) + i\sin\big(\theta+(n-1)\theta\big)\big] \\ &= \rho^n[\cos(n\theta) + i\sin(n\theta)]. \end{aligned}$$

□

In particolare, quindi

$$\forall n \in \mathbb{N} \qquad (\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta). \tag{4.16}$$

Definizione 4.16

Sia $w \in \mathbb{C}$ e $n \geq 1$. Una *radice n-esima* di w è un numero complesso $z \in \mathbb{C}$ tale che $z^n = w$.

Noi sappiamo che ogni numero reale x ha due radici n-esime reali se n è pari e x è positivo, una sola se n è dispari (e x è qualunque), e nessuna se n è pari e x è negativo. Sui complessi, la situazione cambia drasticamente. Abbiamo già trovato due radici quadrate di -1, cioè i e $-i$; quindi (perché?) sui complessi ogni numero reale diverso da zero ha esattamente due radici quadrate distinte. Ma vale un risultato molto più forte, conseguenza della formula di De Moivre.

Proposizione 4.23

Sia $w \in \mathbb{C}^*$ un numero complesso diverso da zero, e $n \geq 1$. Allora esistono esattamente n radici distinte n-esime $z_0, \ldots, z_{n-1}$ di w. Per la precisione, se $w = \rho(\cos\theta + i\sin\theta)$ allora

$$z_k = \rho^{1/n}\left[\cos\left(\frac{\theta+2k\pi}{n}\right) + i\sin\left(\frac{\theta+2k\pi}{n}\right)\right] \tag{4.17}$$

per $k = 0, \ldots, n-1$, dove $\rho^{1/n}$ indica l'unica radice n-esima reale positiva di ρ.

Dimostrazione

Prima di tutto, grazie alla formula di De Moivre abbiamo

$$(z_k)^n = \rho\big[\cos(\theta+2k\pi) + i\sin(\theta+2k\pi)\big] = \rho(\cos\theta + i\sin\theta) = w$$

per $k = 0, \ldots, n-1$, per cui $z_0, \ldots, z_{n-1}$ sono effettivamente radici n-esime di w.

Viceversa, se $z = r(\cos\psi + i\sin\psi)$ è tale che $z^n = w$, allora si deve avere $r^n = \rho$, $\cos(n\psi) = \cos\theta$ e $\sin(n\psi) = \sin\theta$, cioè $r = \rho^{1/n}$ e $n\psi = \theta + 2k\pi$ per qualche $k \in \mathbb{Z}$, per cui z è della forma (4.17). Siccome (esercizio) $z_{k+\alpha n} = z_k$ per ogni k, $\alpha \in \mathbb{Z}$, rimane soltanto da dimostrare che $z_0, \ldots, z_{n-1}$ sono tutti diversi. Supponiamo che esistano $0 \le h, k \le n-1$ tali che $z_h = z_k$. Allora

$$\cos\left(\frac{\theta + 2h\pi}{n}\right) = \cos\left(\frac{\theta + 2k\pi}{n}\right), \qquad \sin\left(\frac{\theta + 2h\pi}{n}\right) = \sin\left(\frac{\theta + 2k\pi}{n}\right),$$

e quindi deve esistere $l \in \mathbb{Z}$ tale che $\dfrac{\theta + 2h\pi}{n} = \dfrac{\theta + 2k\pi}{n} + 2l\pi$, cioè $h - k$ dev'essere un multiplo intero di n. Ma h e k sono entrambi compresi fra 0 e $n-1$; quindi necessariamente $h = k$. □

Osservazione 4.17

Tutte le radici n-esime di w hanno lo stesso modulo, $|w|^{1/n}$; quindi stanno tutte sulla circonferenza di centro l'origine e raggio $|w|^{1/n}$.

Esempio 4.23

Calcoliamo le radici terze di 1. La formula (4.17) ci dà

$$\begin{aligned} z_0 &= \cos 0 + i\sin 0 = 1, \\ z_1 &= \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = \frac{1}{2}(-1 + i\sqrt{3}), \\ z_2 &= \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2}(1 + i\sqrt{3}). \end{aligned}$$

Nota che $z_3 = \cos\dfrac{6\pi}{3} + i\sin\dfrac{6\pi}{3} = \cos(2\pi) + i\sin(2\pi) = 1 = z_0$, come previsto. Le radici terze dell'unità sono rappresentate nella Figura 4.3.

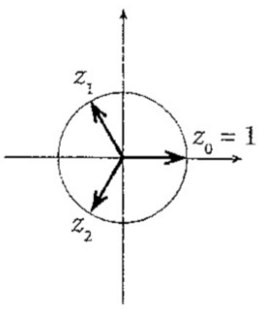

Figura 4.3 - Le radici terze dell'unità.

Complementi 4

4C.1 Il Teorema fondamentale dell'Algebra

Un altro modo di esprimere la Proposizione 4.23 è dire che il polinomio $z^n - c$ ammette sempre esattamente n radici qualunque sia $c \in \mathbb{C}^*$. La proprietà più importante dei numeri complessi è che questo vale per *qualsiasi* polinomio, e non solo per $z^n - c$: un polinomio di grado n ammette sempre esattamente n radici. Ma per spiegare bene cosa intendiamo, premettiamo un paio di definizioni e lemmi.

Definizione 4C.1

Se $p \in \mathbb{C}[z]$ è un polinomio in una variabile a coefficienti complessi, una *radice* (o *zero*) di p è un numero complesso $z_0 \in \mathbb{C}$ tale che $p(z_0) = 0$.

Lemma 4C.1

Sia $p \in \mathbb{C}[z]$ un polinomio in una variabile a coefficienti complessi di grado n, e $z_0 \in \mathbb{C}$. Allora z_0 è una radice di p se e solo se esiste un polinomio $q \in \mathbb{C}[t]$ di grado $n-1$ tale che $p(z) = (z - z_0)q(z)$.

Dimostrazione

Dividendo[4] $p(z)$ per $z - z_0$, si trova un quoziente $q(z)$ di grado $n - 1$ e un resto r di grado zero (cioè una costante) tali che $p(z) = (z - z_0)q(z) + r$. In particolare, $r = p(z_0)$, e quindi $p(z_0) = 0$ se e solo se $r = 0$. □

Definizione 4C.2

Sia $p \in \mathbb{C}[z]$ un polinomio a coefficienti complessi, e $z_0 \in \mathbb{C}$ una sua radice. Diremo che z_0 è una radice di *molteplicità* $m \in \mathbb{N}$ se il polinomio $(z - z_0)^m$ divide $p(z)$, mentre $(z - z_0)^{m+1}$ non lo divide.

Esempio 4C.1

Consideriamo il polinomio $p(z) = z^3 + (1 - 2i)z^2 - (1 + 2i)z - 1$. Siccome si ha $p(z) = (z - i)^2(z + 1)$, vediamo che i è radice di p di molteplicità 2, mentre -1 è radice di p di molteplicità 1.

[4] La divisione fra polinomi a coefficienti complessi si effettua esattamente come la divisione fra polinomi a coefficienti reali.

Definizione 4C.3

Diremo che un polinomio $p \in \mathbb{C}[z]$ ha *n radici, contate con la relativa molteplicità*, se le radici $z_1, \ldots, z_k$ di p, con molteplicità rispettivamente $m_1, \ldots, m_k$, sono tali che $m_1 + \cdots + m_k = n$.

Dunque il polinomio $p(z) = z^n - c$ ha n radici di molteplicità 1 se $c \neq 0$, e una sola radice di molteplicità n se $c = 0$; in entrambi i casi ha esattamente n radici, contate con la relativa molteplicità. Bene; il risultato basilare più importante della teoria dei numeri complessi è che questo accade per ogni polinomio di grado n.

Teorema 4C.2

Sia $p \in \mathbb{C}[z]$ un polinomio di grado n. Allora p ha esattamente n radici, contate con la relativa molteplicità.

Questo è il *Teorema fondamentale dell'Algebra*, che è il motivo per cui spesso i numeri complessi sono più utili dei numeri reali: tutti i polinomi hanno il numero giusto di radici complesse, non una di più, e soprattutto non una di meno.

Esistono molte dimostrazioni di questo teorema, ma nessuna completamente elementare. Nel seguito ne descriveremo una che cerca di mediare fra comprensibilità e numero di prerequisiti, richiedendo solo tre risultati di Analisi: il fatto che i polinomi e la radice quadrata sono funzioni continue, il fatto che una funzione continua su un intervallo chiuso e limitato ammette sempre massimo e minimo, e il Teorema di Weierstrass (una successione limitata di numeri reali ammette sempre una sottosuccessione convergente), che implica il seguente lemma.

Lemma 4C.3

Fissato $r \geq 0$, sia $D_r = \{z \in \mathbb{C} \mid |z| \leq r\}$ il disco di raggio r, e sia $f\colon D_r \to \mathbb{R}$ una funzione continua definita su D_r a valori reali. Allora esiste $z_0 \in D_r$ in cui la funzione f assume il suo valore minimo.

Dimostrazione

Se l'insieme $f(D_r) \subseteq \mathbb{R}$ non è limitato inferiormente, scegliamo una successione di punti $z_n = x_n + iy_n \in D_r$ tali che $f(z_n) < -n$. Se invece $f(D_r)$ è limitato inferiormente, sia $a = \inf f(D_r)$ e scegliamo una successione $z_n = x_n + iy_n \in D_r$ tali che $a \leq f(z_n) < a + 1/n$. In entrambi i casi per costruzione $\inf f(D_r) = \lim\limits_{k \to +\infty} f(z_{n_k})$. Inoltre le successioni $\{x_n\}$ e $\{y_n\}$ sono contenute nell'intervallo $[-r, r]$, grazie alla Proposizione 4.21.(3); quindi il Teorema di Weierstrass ci permette di estrarre due sottosuccessioni convergenti, $x_{n_k} \to x_0$ e $y_{n_k} \to y_0$. Dunque $z_{n_k} = x_{n_k} + iy_{n_k}$ tende a $z_0 = x_0 + iy_0$; siccome $|z_{n_k}| \leq r$ per ogni n_k, per definizione di limite (e per la continuità delle operazioni di elevamento al quadrato e di estrazione di radice, che implicano la continuità del modulo) si ha $|z_0| \leq r$, cioè $z_0 \in D_r$. Ma f è una funzione continua; quindi $f(z_0) = \lim\limits_{k \to +\infty} f(z_{n_k})$. Questo vuol dire che $\inf f(D_r)$ non può essere uguale a $-\infty$, per cui $f(z_0) = \lim\limits_{k \to +\infty} f(z_{n_k})$ è necessariamente uguale ad a, che è dunque il valore minimo assunto da f su D_r. □

Dimostrazione del Teorema 4C.2

Procediamo per induzione sul grado di p. Se p ha grado 1, cioè $p(z) = az + b$ con $a \neq 0$, allora p ha un'unica radice di molteplicità 1, e ci siamo.

Supponiamo ora che p abbia grado $n > 1$; vogliamo prima di tutto dimostrare che p ha almeno una radice. Supponiamo per assurdo che p non abbia radici; a meno di moltiplicare p per una costante non nulla (che non cambia la situazione), possiamo assumere che p sia della forma $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$. Sia $C > 1$ una costante più grande del modulo di tutti i coefficienti di p, cioè $C > \max\{1, |a_{n-1}|, \ldots, |a_0|\}$, e prendiamo $z \in \mathbb{C}$ tale che $|z| \geq 2nC$. In particolare, $|z| > 1$, e quindi $|z| \leq |z|^k$ per ogni $1 \leq k \leq n$. Dunque

$$\left|\frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}\right| \leq \frac{|a_{n-1}|}{|z|} + \cdots + \frac{|a_0|}{|z|^n} \leq \frac{nC}{|z|} \leq \frac{1}{2},$$

per cui

$$\begin{aligned}|p(z)| &= \left|z^n\left(1 + \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}\right)\right| \geq |z|^n\left|1 - \left|\frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}\right|\right| \\ &\geq \frac{1}{2}|z|^n \geq 2^{n-1}(nC)^n > |a_0|\end{aligned}$$

(in questi conti abbiamo ampiamente utilizzato la Proposizione 4.21). In altri termini, abbiamo fatto vedere che $|p(z)| > |a_0| = |p(0)|$ non appena $|z| \geq 2nC$.

Ora, per il Lemma 4C.3 la funzione $|p(z)|$ (che è una funzione continua, in quanto modulo di un polinomio) assume un valore minimo nel disco $|z| \leq 2nC$. Il punto z_0 in cui $|p(z)|$ assume il minimo è necessariamente dentro al disco (non può essere sulla circonferenza), in quanto $|p(z_0)|$ dev'essere minore o uguale di $|p(0)|$, mentre per ipotesi $|p(z)| > |p(0)| = |a_0|$ sulla circonferenza.

A meno di sostituire a p il polinomio $p_1(z) = p(z + z_0)$, che ovviamente ha esattamente tante radici quante ne ha p, possiamo supporre che il punto di minimo sia $z_0 = 0$. Siccome abbiamo supposto per assurdo che p_1 non abbia radici, si ha $p_1(0) = b_0 \neq 0$, per cui possiamo scrivere $p_1(z) = b_0 + b_1z^k + z^{k+1}q(z)$ per qualche $k \geq 1$ e $b_1 \neq 0$, dove $q(z) \in \mathbb{C}[z]$ è un polinomio opportuno.

Sia $z_1 \neq 0$ tale che $z_1^k = -b_0/b_1$ (un tale z_1 esiste per la Proposizione 4.23), e prendiamo $t \in [0, 1]$ reale. Allora

$$p_1(tz_1) = b_0 - t^kb_0 + t^{k+1}z_1^{k+1}q(tz_1) = b_0\left[1 - t^k - t^{k+1}z_1q(tz_1)/b_1\right].$$

Ora, la funzione $t \mapsto |z_1q(tz_1)/b_1|$ è continua, per cui ammette massimo nell'intervallo chiuso e limitato $[0, 1]$. Pertanto esiste $C_1 > 1$ tale che per ogni $t \in [0, 1]$ si ha $|z_1q(tz_1)/b_1| \leq C_1$, e quindi $|p_1(tz_1)| \leq |b_0|(1 - t^k + C_1t^{k+1})$ per ogni $t \in [0, 1]$. Prendiamo allora $0 < t_0 < 1/C_1$; si ha (esercizio) $0 < 1 - t_0^k + C_1t_0^{k+1} < 1$, e quindi $|p_1(t_0z_1)| < |b_0| = |p_1(0)|$, contraddicendo così l'ipotesi che 0 fosse il punto di minimo del modulo di p_1.

La contraddizione deriva dall'aver supposto $b_0 \neq 0$, cioè dall'aver supposto che il nostro polinomio p non abbia radici. Questo si è rivelato impossibile, e quindi p deve avere almeno una radice z_0. Ma allora possiamo scrivere $p(z) = (z - z_0)q(z)$, dove q è un polinomio di grado $n - 1$. Per ipotesi induttiva, q ha esattamente $n - 1$ radici, contate con la relativa molteplicità. Siccome ognuna di queste è anche radice di p, abbiamo fatto vedere che p ha esattamente n radici, contate con la relativa molteplicità, come richiesto. □

Esercizi 4

4.1 Sia V uno spazio vettoriale. Dimostra che $0v = O$ per ogni $v \in V$ utilizzando soltanto le altre proprietà della definizione di spazio vettoriale. (*Suggerimento:* parti da $0 + 0 = 0$.)

4.2 Sia V uno spazio vettoriale. Dimostra che $(-1)v + v = O$ e che $(-1)v = -v$ per ogni $v \in V$. Come conseguenza, dimostra che se v, v_1, $v_2 \in V$ sono tali che $v_1 + v = O = v_2 + v$, allora $v_1 = v_2$; in altre parole, l'opposto di un vettore è univocamente determinato.

4.3 Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e prendiamo $v \in V$ e $\lambda \in \mathbb{K}$. Dimostra che $\lambda v = O$ se e solo se $\lambda = 0$ oppure $v = O$.

4.4 Fissiamo un riferimento affine in $\mathcal{A}^3$, in modo da identificare lo spazio euclideo con $\mathbb{R}^3$. Dimostra che $\{O\}$, le rette per l'origine, i piani per l'origine ed $\mathbb{R}^3$ sono tutti e soli i sottospazi di $\mathbb{R}^3$.

4.5 Dimostra che $\mathbb{R}_n[t]$, l'insieme dei polinomi a coefficienti reali in una variabile di grado minore o uguale a n, è un sottospazio vettoriale di $\mathbb{R}[t]$.

4.6 Dimostra che l'insieme delle matrici diagonali (rispettivamente, triangolari superiori o inferiori) è un sottospazio vettoriale dello spazio delle matrici quadrate.

4.7 Per quali polinomi $p \in \mathbb{R}[t]$ il grafico $\Gamma = \{(t, p(t)) \mid t \in \mathbb{R}\}$ è un sottospazio vettoriale di $\mathbb{R}^2$?

4.8 Sia $V_c = \{p \in \mathbb{R}_7[t] \mid p(1) = c\}$. Per quali valori di $c \in \mathbb{R}$ l'insieme V_c è un sottospazio vettoriale di $\mathbb{R}_7[t]$?

4.9 Sia $S \subseteq V$ un sottoinsieme qualunque di uno spazio vettoriale. Dimostra che $\mathrm{Span}(S)$ è il più piccolo sottospazio di V che contiene S, nel senso che se W è un sottospazio di V che contiene S allora $W \supseteq \mathrm{Span}(S)$.

4.10 Dimostra che $\left\{\begin{vmatrix} 2 \\ 1 \end{vmatrix}, \begin{vmatrix} 1 \\ 1 \end{vmatrix}\right\}$ $\left\{\begin{vmatrix} 1 \\ 3 \end{vmatrix}, \begin{vmatrix} 2 \\ 3 \end{vmatrix}\right\}$, $\left\{\begin{vmatrix} -1 \\ 1 \end{vmatrix}, \begin{vmatrix} -1 \\ -1 \end{vmatrix}\right\}$ sono basi di $\mathbb{R}^2$.

4.11 Trova le coordinate dei vettori $\begin{vmatrix} 1 \\ 3 \end{vmatrix}$, $\begin{vmatrix} 2 \\ -1 \end{vmatrix}$, $\begin{vmatrix} 7 \\ -7 \end{vmatrix}$ e $\begin{vmatrix} 1 \\ 1 \end{vmatrix}$ rispetto alle basi di $\mathbb{R}^2$ del precedente esercizio.

4.12 Dati $\mathcal{B}_1 = \{e_1, e_1 + e_2, e_1 + e_3\}$, $\mathcal{B}_2 = \{e_1 + e_2, e_1 + 2e_2 + e_3, 2e_1 + 3e_2 + e_3\} \subset \mathbb{R}^3$, dimostra che $\mathcal{B}_1$ è una base di $\mathbb{R}^3$ mentre $\mathcal{B}_2$ non lo è.

4.13 Trova le coordinate di $2e_1 + e_2 + 7e_3$ rispetto alla base $\mathcal{B}_1$ del precedente esercizio.

4.14 Per quali valori di $t \in \mathbb{R}$ l'insieme $\left\{ \begin{vmatrix} 2 \\ t \end{vmatrix}, \begin{vmatrix} t \\ 2 \end{vmatrix} \right\}$ è una base di $\mathbb{R}^2$?

4.15 Dimostra che $\left\{ \begin{vmatrix} a \\ c \end{vmatrix}, \begin{vmatrix} b \\ d \end{vmatrix} \right\}$ è una base di $\mathbb{R}^2$ se e solo se $ad - bc \neq 0$.

4.16 Prendiamo i polinomi $p_1(t) = 1 + t$, $p_2(t) = 1 + 2t + t^2$, $p_3(t) = t - t^2 \in \mathbb{R}_2[t]$. Dimostra che $\{p_1, p_2, p_3\}$ è una base di $\mathbb{R}_2[t]$, e trova le coordinate di $q_1(t) = 2 - t + t^2$ e $q_2(t) = 3 + t^2$ rispetto a questa base.

4.17 Sia $E_{ij} \in M_{m,n}(\mathbb{R})$ la matrice che ha 1 al posto (i,j) e 0 altrove. Dimostra che $\{E_{11}, \dots, E_{mn}\}$ è una base di $M_{m,n}(\mathbb{R})$; in particolare, $\dim M_{m,n}(\mathbb{R}) = mn$.

4.18 Dimostra che due vettori di $\mathcal{V}_O^2$ sono linearmente indipendenti se e solo se non sono proporzionali, e che tre vettori di $\mathcal{V}_O^3$ sono linearmente indipendenti se e solo se non sono complanari.

4.19 Sia $\mathcal{B}$ un sottoinsieme finito massimale di vettori linearmente indipendenti in uno spazio vettoriale V. Dimostra che $\mathcal{B}$ è una base di V. Analogamente, sia $\mathcal{C}$ un sistema finito di generatori minimale (nel senso che non è più un sistema di generatori se gli si toglie anche solo un elemento); dimostra che anche $\mathcal{C}$ è una base.

4.20 Sia V uno spazio vettoriale di dimensione n, e $\mathcal{B} = \{v_1, \dots, v_n\}$ un sistema di generatori di V. Dimostra che $\mathcal{B}$ è una base di V.

4.21 Sia V uno spazio vettoriale di dimensione n, e prendiamo $w_1, \dots, w_p \in V$. Dimostra che se $p < n$ allora $\{w_1, \dots, w_p\}$ non è un sistema di generatori di V.

4.22 Dimostra che un sistema lineare omogeneo con più incognite che equazioni ha sempre soluzioni non banali (cioè diverse dalla soluzione nulla).

4.23 Sia V uno spazio vettoriale di dimensione n, e $\mathcal{A} = \{v_1, \dots, v_k\} \subset V$. Dimostra che $\mathcal{A}$ è un sistema di generatori di V se e solo se $k \geq n$ e $\mathcal{A}$ contiene n vettori linearmente indipendenti.

4.24 Sia $\mathbb{K}$ un campo. Dimostra che $\mathbb{K}$ è uno spazio vettoriale su se stesso di dimensione 1.

4.25 Dimostra che $\mathbb{R}[t]$ non ha un sistema di generatori finito.

4.26 Siano V e W due spazi vettoriali. Definisci una somma e un prodotto per scalari sul prodotto cartesiano $V \times W$ in modo da renderlo uno spazio vettoriale. Se $\dim V = m$ e $\dim W = n$, calcola la dimensione di $V \times W$.

4.27 Dimostra che $U + W$ è il più piccolo sottospazio di V contenente U e W.

4.28 Sia V uno spazio vettoriale, e U, W sottospazi. Dimostra che $U \cup W$ è un sottospazio se e solo se $U \subseteq W$ oppure $W \subseteq U$. Nel primo caso, $U + W = U \cup W = W$, mentre nel secondo $U + W = U \cup W = U$.

4.29 Trova due sottoinsiemi di $\mathbb{R}^2$, il primo chiuso rispetto alla somma ma non rispetto al prodotto per scalari, e il secondo chiuso rispetto al prodotto per scalari ma non rispetto alla somma.

4.30 Siano U e W sottospazi di uno spazio vettoriale V con $V = U + W$, e siano $\mathcal{B}$ una base di U e $\mathcal{C}$ una base di W. Mostra che se $U \cap V = \{O\}$ allora $\mathcal{B} \cup \mathcal{C}$ è una base di V, e trova un esempio in cui $U \cap W \neq \{O\}$ e $\mathcal{B} \cup \mathcal{C}$ non è una base di V.

4.31 Sia $v_0 = \begin{vmatrix} 1 \\ 1 \end{vmatrix} \in \mathbb{R}^2$; trova quattro supplementari distinti di $\mathbb{R}v_0 \subset \mathbb{R}^2$.

4.32 Siano $U = \mathrm{Span}(e_1 + e_3)$ e $W = \mathrm{Span}(2e_1 + e_2 + e_3, e_2)$ due sottospazi di $\mathbb{R}^3$. Dimostra che $\mathbb{R}^3 = U \oplus W$.

4.33 Siano U, W, Z sottospazi di uno spazio vettoriale V di dimensione finita tali che $U \cap W = \{O\}$, $U + Z = V$ e $\dim W = \dim Z = \dim V - \dim U$. Dimostra che sia W sia Z sono supplementari di U.

4.34 Sia V uno spazio vettoriale di dimensione finita. Una *bandiera* di V è una successione $V_0, \dots, V_n$ di sottospazi di V tali che $V_0 \subset V_1 \subset \cdots \subset V_n$. Dimostra che $\dim V_j = j$ per $j = 1, \dots, n$ (per cui $V_0 = \{O\}$ e $V_n = V$), e che ogni sottospazio di V è contenuto in una bandiera.

4.35 Calcola le coordinate polari dei seguenti punti: $(-3, 0)$; $(0, 2\pi)$; $(2\sqrt{3}, 2)$; $(6, -6)$; $(-2, 2\sqrt{3})$.

4.36 Esprimi nella forma $a + bi$ i seguenti numeri complessi (ovvero, fai i conti):

$$\left(\frac{1 + i\sqrt{3}}{2}\right)^3; \quad (-i)^4; \quad (2 + 2i)(-2 + 2i); \quad \frac{1 + 3i}{1 - 3i}.$$

4.37 Dimostra la formula del "modulo al quadrato del binomio":

$$|z + w|^2 = |z|^2 + 2\mathrm{Re}(z\overline{w}) + |w|^2.$$

4.38 Utilizza (4.16) per ritrovare le formule trigonometriche che esprimono seno e coseno del doppio e del triplo di un angolo.

4.39 Trova le quattro radici quarte di 1, e le due radici quadrate di $1 - 4i\sqrt{3}$.

4.40 Determina modulo e argomento delle radici cubiche di $(1 + i)(1 - i)^{-1}$.

4.41 Determina i numeri complessi $z \in \mathbb{C}$ tali che $z^4 = \overline{z}$.

4.42 Determina, al variare di $k \in \mathbb{R}$, le soluzioni dell'equazione $ikz^3 = |z|^3$.

4.43 Trova le radici complesse del polinomio $p_1(z) = z^3 + i \in \mathbb{C}[z]$. Trova, se esiste, un polinomio $p_2 \in \mathbb{C}[z]$ tale che $p_1 p_2$ abbia tutti i coefficienti reali.

4.44 Se $z = x + iy \in \mathbb{C}$, si definisce l'*esponenziale complesso* $e^z \in \mathbb{C}$ con la formula $e^z = e^x(\cos y + i \sin y)$. Dimostra che si ha $e^z \cdot e^w = e^{z+w}$ per ogni z, $w \in \mathbb{C}$, e che $e^{i\pi} + 1 = 0$.

Esercizi relativi ai Complementi

4C.1 Determina per quali $\lambda \in \mathbb{C}$ l'equazione $(\lambda^8 - 1)z^3 + (\lambda^4 - 1)z + \lambda^2 - i = 0$ non ammette soluzioni.

4C.2 Trova un polinomio a coefficienti reali di quinto grado con solo una radice reale, e un polinomio a coefficienti reali di quarto grado senza nessuna radice reale.

4C.3 Dato un polinomio $p(t) = a_n t^n + \cdots + a_0 \in \mathbb{Z}[t]$ a coefficienti interi, prendiamo una radice razionale $r = a/b \in \mathbb{Q}$ di p, dove $a, b \in \mathbb{Z}$ sono relativamente primi. Dimostra che a divide a_0, e che b divide a_n.

4C.4 Trova le radici dei polinomi a coefficienti interi $p_1(t) = t^4 - 2t^3 - 5t^2 + 4t + 6$ e $p_2(t) = t^5 - 5t^3 - 2t^2 + 6t + 4$. (*Suggerimento*: usa l'esercizio precedente.)