CAPITOLO 2

Spazi vettoriali

Molto spesso in matematica si scoprono analogie in questioni diverse che, viste più astrattamente, ammettono di essere trattate allo stesso modo.

Dovrebbe ormai essere chiaro, per esempio, che ci sono somiglianze molto strette fra gli insiemi dei vettori colonna con due, tre o un qualunque numero di righe. Il concetto di spazio vettoriale permette di vederli dallo stesso punto di osservazione, cogliendo allo stesso tempo ciò che li distingue fra loro.

Lo sviluppo storico di un concetto matematico astratto è lungo e non lo seguiremo. Non si possono comunque tacere i nomi di Cayley, Sylvester e Steinitz che hanno contribuito a creare quello di spazio vettoriale.

L'idea di fondo nelle astrazioni è di cercare di isolare alcuni aspetti di una teoria, ignorandone altri. Nel nostro caso quelli che terremo sono:

- due vettori colonna con lo stesso numero di righe possono essere sommati;
- un vettore colonna può essere moltiplicato per uno scalare.

Naturalmente le operazioni che possiamo eseguire sui vettori hanno certe proprietà; anche fra queste ne isoleremo alcune.

1. Definizione di spazio vettoriale

Uno *spazio vettoriale* (*complesso*) V è un insieme sul quale siano definite due operazioni:

$$\begin{aligned} V \times V &\to V & \qquad \mathbb{C} \times V \\ (\mathbf{u}, \mathbf{v}) &\mapsto \mathbf{u} + \mathbf{v} & \qquad (\alpha, \mathbf{v}) \mapsto \alpha\mathbf{v} \end{aligned}$$

La prima si chiama *addizione*, la seconda *moltiplicazione per scalari.* Gli elementi di V saranno chiamati *vettori.* Richiediamo anche che le operazioni abbiano le seguenti proprietà:

(A1) per ogni $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w};$$

(A2) esiste un elemento $\mathbf{0} \in V$ tale che, per ogni $\mathbf{v} \in V$,

$$\mathbf{0} + \mathbf{v} = \mathbf{v}, \qquad \mathbf{v} + \mathbf{0} = \mathbf{v};$$

(A3) per ogni $\mathbf{v} \in V$ esiste $\mathbf{w} \in V$ tale che

$$\mathbf{v} + \mathbf{w} = \mathbf{0}, \qquad \mathbf{w} + \mathbf{v} = \mathbf{0}.$$

(M1) per ogni $\alpha, \beta \in \mathbb{C}$ e ogni $\mathbf{v} \in V$,

$$\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v};$$

(M2) per ogni $\alpha, \beta \in \mathbb{C}$ e ogni $\mathbf{v} \in V$,

$$(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v};$$

(M3) per ogni $\alpha \in \mathbb{C}$ e ogni $\mathbf{u}, \mathbf{v} \in V$,

$$\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v};$$

(M4) per ogni $\mathbf{v} \in V$,

$$1\mathbf{v} = \mathbf{v}.$$

È bene fare qualche commento sulle condizioni appena enunciate, che chiameremo *regole di calcolo* negli spazi vettoriali.

Nella (A1) stiamo imponendo che l'addizione sia *associativa*. Quindi, come si fa con le espressioni numeriche, potremo omettere le parentesi quando si sommano più elementi di uno spazio vettoriale.

Nella (A2) chiediamo l'esistenza di un elemento speciale, che denoteremo sempre con $\mathbf{0}$; come vedremo non c'è pericolo di confusione, perché questo elemento è unico (in ogni spazio vettoriale) e sarà chiamato anche *vettore nullo*.

Nella (A3) va notato l'ordine diverso delle parole 'per ogni' e 'esiste' rispetto alla precedente: l'elemento $\mathbf{w}$ dipende da $\mathbf{v}$ e si chiamerà *opposto* di $\mathbf{v}$, perché vedremo che è unico.

Nelle condizioni successive abbiamo implicitamente usato una convenzione: che la moltiplicazione per scalari abbia la precedenza sull'addizione.

La (M1) è una specie di proprietà associativa, la (M2) e la (M3) si chiamano proprietà *distributive*. Insieme alle altre dicono che le espressioni algebriche negli spazi vettoriali si trattano allo stesso modo in cui si è abituati con quelle numeriche.

Per finire, la (M4) impone una proprietà che sappiamo valere per i vettori colonna. Pur venendo per ultima è molto importante, come vedremo.

Elenchiamo ora alcuni risultati che dimostreremo solo usando le regole di calcolo appena esposte.

PROPOSIZIONE 1.1. *Sia V uno spazio vettoriale complesso.*

(1) *L'elemento neutro dell'addizione è unico.*
(2) *L'opposto di un elemento $\mathbf{v} \in V$ è unico.*
(3) *Per ogni $\mathbf{u}, \mathbf{v} \in V$, si ha*

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}.$$

DIMOSTRAZIONE. (1) Supponiamo esista un elemento $\mathbf{z} \in V$ tale che, per ogni $\mathbf{v} \in V$,

$$\mathbf{z} + \mathbf{v} = \mathbf{v}, \qquad \mathbf{v} + \mathbf{z} = \mathbf{v},$$

cioè un altro elemento neutro. La proprietà vale anche quando al posto di $\mathbf{v}$ scriviamo $\mathbf{0}$:

$$\mathbf{z} + \mathbf{0} = \mathbf{0}, \qquad \mathbf{0} + \mathbf{z} = \mathbf{0}.$$

Ma la proprietà (A2) dice che $\mathbf{z} + \mathbf{0} = \mathbf{z}$, perché $\mathbf{z} \in V$. Dunque $\mathbf{z} = \mathbf{0}$.

(2) Supponiamo che esistano due elementi $\mathbf{w}$ e $\mathbf{w}'$ che si comportino da opposto di $\mathbf{v}$, cioè:

$$\mathbf{v} + \mathbf{w} = \mathbf{0}, \qquad \mathbf{w} + \mathbf{v} = \mathbf{0},$$
$$\mathbf{v} + \mathbf{w}' = \mathbf{0}, \qquad \mathbf{w}' + \mathbf{v} = \mathbf{0}.$$

Consideriamo allora l'elemento

$$\mathbf{u} = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{w} + (\mathbf{v} + \mathbf{w}'),$$

dove la seconda uguaglianza segue dalla (A1). Applichiamo allora ciò che sappiamo: dalla (A3) e dalla (A2) segue

$$\mathbf{u} = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{0} + \mathbf{w}' = \mathbf{w}'.$$

Dall'ipotesi su $\mathbf{w}'$ e dalla (A2) segue

$$\mathbf{u} = \mathbf{w} + (\mathbf{v} + \mathbf{w}') = \mathbf{w} + \mathbf{0} = \mathbf{w}.$$

Dunque $\mathbf{w} = \mathbf{w}'$.

(3) Useremo tutte le proprietà tranne la (M1). Poniamo

$$\mathbf{w} = (1 + 1)(\mathbf{u} + \mathbf{v}).$$

Allora

$$\begin{aligned}
\mathbf{w} &= 1(\mathbf{u}+\mathbf{v}) + 1(\mathbf{u}+\mathbf{v}) && \text{(M2)}\\
&= (\mathbf{u}+\mathbf{v}) + (\mathbf{u}+\mathbf{v}) && \text{(M4)}\\
&= \mathbf{u}+\mathbf{v}+\mathbf{u}+\mathbf{v} && \text{(A1)}\\
\mathbf{w} &= (1+1)\mathbf{u} + (1+1)\mathbf{v} && \text{(M3)}\\
&= (1\mathbf{u}+1\mathbf{u}) + (1\mathbf{v}+1\mathbf{v}) && \text{(M2)}\\
&= (\mathbf{u}+\mathbf{u}) + (\mathbf{v}+\mathbf{v}) && \text{(M4)}\\
&= \mathbf{u}+\mathbf{u}+\mathbf{v}+\mathbf{v} && \text{(A1).}
\end{aligned}$$

Possiamo allora calcolare $(-\mathbf{u}) + \mathbf{w} + (-\mathbf{v})$ in due modi:

$$\begin{aligned}
(-\mathbf{u}) + \mathbf{w} + (-\mathbf{v}) &= (-\mathbf{u}) + \mathbf{u}+\mathbf{v}+\mathbf{u}+\mathbf{v} + (-\mathbf{v})\\
&= \mathbf{0}+\mathbf{v}+\mathbf{u}+\mathbf{0}\\
&= \mathbf{v}+\mathbf{u}\\
(-\mathbf{u}) + \mathbf{w} + (-\mathbf{v}) &= (-\mathbf{u}) + \mathbf{u}+\mathbf{u}+\mathbf{v}+\mathbf{v} + (-\mathbf{v})\\
&= \mathbf{0}+\mathbf{u}+\mathbf{v}+\mathbf{0}\\
&= \mathbf{u}+\mathbf{v}
\end{aligned}$$

dove abbiamo usato la (A3) e la (A2), oltre alla (A1). □

Useremo allora una notazione speciale per l'opposto, giustificata dal fatto che esso esiste, per la regola (A3), ed è unico per quanto visto in (2):

l'opposto di $\mathbf{v}$ si denota con $-\mathbf{v}$.

Il risultato (3) dice che l'addizione in uno spazio vettoriale è *commutativa*.

Ci sono altre proprietà che sono quasi ovvie negli spazi vettoriali già noti, ma che discendono dalle regole di calcolo. Invitiamo il lettore a esplicitare quali regole vengono usate in ogni passaggio. Notiamo anche la differenza fra 0 e $\mathbf{0}$: con il primo indichiamo il numero complesso zero, con il secondo il vettore nullo.

PROPOSIZIONE 1.2. *Sia V uno spazio vettoriale.*

(1) *Per ogni $\mathbf{v} \in V$, si ha*

$$0\mathbf{v} = \mathbf{0}.$$

(2) *Se $\alpha \in \mathbb{C}$, $\mathbf{v} \in V$ e $\alpha\mathbf{v} = \mathbf{0}$ si ha $\alpha = 0$ oppure $\mathbf{v} = \mathbf{0}$.*

(3) *Per ogni $\alpha \in \mathbb{C}$, si ha*

$$\alpha\mathbf{0} = \mathbf{0}.$$

(4) *Se $\mathbf{v}, \mathbf{w} \in V$, allora $\mathbf{w} = -\mathbf{v}$ se e solo se $\mathbf{v} + \mathbf{w} = \mathbf{0}$.*

(5) *Per ogni $\mathbf{v} \in V$ e ogni $\alpha \in \mathbb{C}$, si ha*

$$-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v} = \alpha(-\mathbf{v}).$$

DIMOSTRAZIONE. (1) Poniamo $\mathbf{u} = 0\mathbf{v}$; la proprietà fondamentale di 0 è che $0 + 0 = 0$; allora

$$\mathbf{u} = 0\mathbf{v} = (0+0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v} = \mathbf{u} + \mathbf{u}.$$

Perciò possiamo scrivere

$$\mathbf{0} = \mathbf{u} + (-\mathbf{u}) = (\mathbf{u}+\mathbf{u}) + (-\mathbf{u}) = \mathbf{u} + (\mathbf{u} + (-\mathbf{u})) = \mathbf{u} + \mathbf{0} = \mathbf{u},$$

e quindi $\mathbf{u} = \mathbf{0}$.

(2) Se $\alpha = 0$ non c'è nulla da dimostrare. Supponiamo allora $\alpha \neq 0$; allora

$$\mathbf{0} = \alpha^{-1}\mathbf{0} = \alpha^{-1}(\alpha\mathbf{v}) = (\alpha^{-1}\alpha)\mathbf{v} = 1\mathbf{v} = \mathbf{v}.$$

Lasciamo al lettore la giustificazione di ogni passaggio.

(3) Poniamo $\mathbf{u} = \alpha\mathbf{0}$; per la proprietà fondamentale di $\mathbf{0}$ che $\mathbf{0} + \mathbf{0} = \mathbf{0}$, abbiamo

$$\mathbf{u} = \alpha\mathbf{0} = \alpha(\mathbf{0}+\mathbf{0}) = \alpha\mathbf{0} + \alpha\mathbf{0} = \mathbf{u} + \mathbf{u}.$$

Come prima possiamo concludere che $\mathbf{u} = \mathbf{0}$.

(4) Se $\mathbf{w} = -\mathbf{v}$, che $\mathbf{v} + \mathbf{w} = \mathbf{0}$ fa parte della definizione. Se, viceversa, $\mathbf{0} = \mathbf{v} + \mathbf{w}$, possiamo sommare ad ambo i membri $-\mathbf{v}$; allora

$$-\mathbf{v} = -\mathbf{v} + \mathbf{0} = -\mathbf{v} + (\mathbf{v} + \mathbf{w}) = (-\mathbf{v} + \mathbf{v}) + \mathbf{w} = \mathbf{0} + \mathbf{w} = \mathbf{w}$$

e la tesi è provata.

(5) Dimostreremo dapprima che $-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v}$. Alla luce di (4) ci basta calcolare

$$\alpha\mathbf{v} + (-\alpha)\mathbf{v} = (\alpha + (-\alpha))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$$

e abbiamo la tesi. Analogamente,

$$\alpha\mathbf{v} + \alpha(-\mathbf{v}) = \alpha(\mathbf{v} + (-\mathbf{v})) = \alpha\mathbf{0} = \mathbf{0},$$

così la dimostrazione è completa. □

Lo scopo di queste proposizioni è di mostrare come dalle poche regole di calcolo che abbiamo imposto, ne seguano molte altre.

2. Spazi vettoriali reali

In alcuni casi la moltiplicazione di un vettore per un numero complesso non ha senso; ne ha invece quella per un numero reale. Si pensi, per esempio, ai vettori della fisica. Questi non formano uno spazio vettoriale nel senso specificato finora, ma solo perché le regole di calcolo (M1)–(M4) si applicano solo a numeri reali.

Tuttavia è facile verificare che ogni enunciato finora dimostrato è valido anche in questo contesto. Parleremo in questo caso di *spazi vettoriali reali*. Dunque uno spazio vettoriale reale è un insieme V con l'operazione di addizione e quella di moltiplicazione per numeri reali, nel quale valgano le regole (A1)–(A3) e (M1)–(M4) ristrette ai numeri reali.

Di fatto è possibile anche una maggiore astrazione, definendo spazi vettoriali con moltiplicazione per *scalari* che siano elementi di quelli che in algebra si chiamano *campi*: esempi di campi noti allo studente dovrebbero essere i numeri razionali, i numeri reali e i numeri complessi, ma ne esistono altri. In essi le proprietà rilevanti sono di poter eseguire operazioni di addizione e moltiplicazione con regole analoghe a quelle in $\mathbb{C}$, fra cui l'esistenza dell'inverso rispetto alla moltiplicazione di ogni elemento diverso da 0. Un esempio molto importante è quello degli interi modulo p, dove p è un numero primo. Non seguiremo questo tipo di astrazione. Quello che è rilevante notare è che i risultati che esporremo in questo Capitolo valgono per spazi vettoriali su campi qualunque.

Per parlare di spazi vettoriali in modo generico, chiameremo *scalari* i numeri per i quali è ammessa la moltiplicazione.

3. Esempi

Vettori colonna e matrici. Per ogni $m \geq 1$, l'insieme delle matrici $\mathrm{M}_{m\times 1}(\mathbb{C})$ è uno spazio vettoriale, prendendo come addizione quella fra matrici e come moltiplicazione per scalari quella già definita. Questi spazi vettoriali sono i più importanti, quindi useremo per essi la notazione speciale $\mathbb{C}^m$.

Più in generale, per ogni $m, n \geq 1$, l'insieme delle matrici $\mathrm{M}_{m\times n}(\mathbb{C})$ di forma $m \times n$ è uno spazio vettoriale, con le operazioni definite nel Capitolo 1 di addizione fra matrici e prodotto per numeri complessi.

Analogamente, $\mathrm{M}_{m\times n}(\mathbb{R})$ (cioè le matrici $m \times n$ a coefficienti reali) è uno spazio vettoriale reale. Porremo $\mathbb{R}^m = \mathrm{M}_{m\times 1}(\mathbb{R})$, come visto in precedenza.

Spazi di funzioni. Consideriamo l'insieme $C([a, b])$ delle funzioni reali continue sull'intervallo $[a, b]$. Possiamo renderlo uno spazio vettoriale reale definendo, per $f, g \in C([a, b])$ e $\alpha \in \mathbb{R}$,

$$(f + g)(t) = f(t) + g(t), \qquad (\alpha f)(t) = \alpha f(t), \quad t \in [a, b]$$

cioè definendo la funzione $f+g$ e la funzione αf sull'intervallo $[a, b]$. Verifiche note dimostrano che sia $f + g$ che αf sono ancora funzioni continue. Si lascia come esercizio al lettore la verifica che le regole di calcolo sono valide. Per esempio, il vettore nullo è la funzione che vale costantemente zero.

Polinomi. Un *polinomio reale a coefficienti reali* è un'espressione della forma

$$a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$$

per opportuni $a_0, a_1, \dots, a_n \in \mathbb{R}$ detti *coefficienti*. La X non indica un numero, ma è solo un simbolo che obbedisce alle regole $X^aX^b = X^{a+b}$. Anche queste espressioni formano uno spazio vettoriale reale con l'usuale addizione e la moltiplicazione per scalari definita come al solito (una costante è un particolare polinomio). Notiamo che in questo contesto non ci interessa la moltiplicazione generale fra polinomi, solo quella per le costanti.

Questo spazio vettoriale sarà indicato con $\mathcal{P}(\mathbb{R})$.

Il *grado* di un polinomio è l'esponente massimo che compare effettivamente; così il grado di $2 - X + X^2$ è due, quello di $3X + 2X^4$ è quattro. Il polinomio nullo ha, per convenzione, grado $-\infty$.

Possiamo notare che, dati i polinomi f e g, il grado di $f + g$ è non maggiore del grado di f e di g. Analogamente, il grado di αf è non maggiore del grado di f, se α è uno scalare.

Di conseguenza l'insieme $\mathcal{P}_n(\mathbb{R})$ dei polinomi di grado $< n$ è esso stesso uno spazio vettoriale. In particolare $\mathcal{P}_0(\mathbb{R}) = \{0\}$ contiene solo il polinomio nullo, mentre possiamo considerare $\mathcal{P}_1(\mathbb{R}) = \mathbb{R}$.

Polinomi a coefficienti complessi. Tutto quanto detto per i polinomi a coefficienti reali vale ancora per polinomi a coefficienti complessi; useremo le notazioni $\mathcal{P}(\mathbb{C})$ e $\mathcal{P}_n(\mathbb{C})$.

Lo spazio vettoriale nullo. Un insieme con un solo elemento diventa in modo ovvio uno spazio vettoriale. Quell'unico elemento sarà $\mathbf{0}$, con l'unica possibile addizione e l'unica moltiplicazione per scalari, cioè $\mathbf{0} + \mathbf{0} = \mathbf{0}$ e $\alpha\mathbf{0} = \mathbf{0}$.

4. Sottospazi e combinazioni lineari

Consideriamo una matrice $\mathbf{A} \in \mathrm{M}_{m\times n}(\mathbb{C})$ e poniamo

$$\mathrm{N}(\mathbf{A}) = \{\, \mathbf{v} \in \mathbb{C}^n : \mathbf{A}\mathbf{v} = \mathbf{0} \,\}.$$

È ovvio che $\mathbf{0} \in \mathrm{N}(\mathbf{A})$. Se prendiamo due elementi $\mathbf{u}$ e $\mathbf{v}$ di $\mathrm{N}(\mathbf{A})$ e li sommiamo, otteniamo ancora un elemento di $\mathrm{N}(\mathbf{A})$: infatti

$$\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Analogamente, se $\alpha \in \mathbb{C}$,

$$\mathbf{A}(\alpha\mathbf{v}) = \alpha\mathbf{A}\mathbf{v} = \alpha\mathbf{0} = \mathbf{0},$$

quindi $\alpha\mathbf{v} \in \mathrm{N}(\mathbf{A})$. Non è difficile verificare allora che le regole di calcolo in $\mathrm{N}(\mathbf{A})$ valgono, dal momento che valgono in $\mathbb{C}^n$.

Questo esempio può essere generalizzato, notando che gli unici fatti usati sono che $\mathbf{0} \in \mathrm{N}(\mathbf{A})$, che dati $\mathbf{u}, \mathbf{v} \in \mathrm{N}(\mathbf{A})$ anche $\mathbf{u} + \mathbf{v} \in \mathrm{N}(\mathbf{A})$ e che dato $\mathbf{v} \in \mathrm{N}(\mathbf{A})$ e $\alpha \in \mathbb{C}$ anche $\alpha\mathbf{v} \in \mathrm{N}(\mathbf{A})$.

Definizione. Un sottoinsieme U dello spazio vettoriale V si dice un *sottospazio vettoriale* (o semplicemente un sottospazio) se

(S_1) $\mathbf{0} \in U$;
(S_2) per ogni $\mathbf{v}_1, \mathbf{v}_2 \in U$, si ha $\mathbf{v}_1 + \mathbf{v}_2 \in U$;
(S_3) per ogni $\mathbf{v} \in U$ e ogni $\alpha \in \mathbb{C}$, $\alpha\mathbf{v} \in U$.

Se U è un sottospazio vettoriale di V, le operazioni di V rendono U esso stesso uno spazio vettoriale. L'unica cosa non ovvia è l'esistenza degli opposti. Ma essa segue subito dal fatto che $-\mathbf{v} = (-1)\mathbf{v}$, per quanto visto prima.

Esempi 4.1. (1) Per ogni $n \geq 0$, $\mathcal{P}_n(\mathbb{C})$ è un sottospazio vettoriale di $\mathcal{P}(\mathbb{C})$.

(2) Se U è un sottospazio vettoriale di V e W è un sottospazio vettoriale di U, allora W è un sottospazio vettoriale di V.

(3) Se V è un qualunque spazio vettoriale, $\{\mathbf{0}\}$ e V sono sottospazi vettoriali di V. Un sottospazio U di V si dice *proprio* se $U \neq V$, cioè se esiste un elemento $\mathbf{v} \in V$ tale che $\mathbf{v} \notin U$.

(4) Sia V uno spazio vettoriale e sia $\mathbf{v} \in V$. Consideriamo

$$\langle \mathbf{v} \rangle = \{\, \alpha\mathbf{v} : \alpha \in \mathbb{C} \,\}$$

cioè l'insieme di tutti i *multipli scalari* di $\mathbf{v}$. Dimostriamo che $\langle \mathbf{v} \rangle$ è un sottospazio di V. Infatti $\mathbf{0} = 0\mathbf{v} \in \langle \mathbf{v} \rangle$. Se $\mathbf{v}_1, \mathbf{v}_2 \in \langle \mathbf{v} \rangle$, allora è $\mathbf{v}_1 = \alpha_1\mathbf{v}$ e $\mathbf{v}_2 = \alpha_2\mathbf{v}_2$, per opportuni scalari α_1 e α_2. Allora

$$\mathbf{v}_1 + \mathbf{v}_2 = \alpha_1\mathbf{v} + \alpha_2\mathbf{v}_2 = (\alpha_1 + \alpha_2)\mathbf{v} \in \langle \mathbf{v} \rangle.$$

Se poi $\alpha \in \mathbb{C}$ e $\mathbf{v}_1 \in \langle \mathbf{v} \rangle$, possiamo scrivere $\mathbf{v}_1 = \alpha_1\mathbf{v}$ e quindi

$$\alpha\mathbf{v}_1 = \alpha(\alpha_1\mathbf{v}) = (\alpha\alpha_1)\mathbf{v} \in \langle \mathbf{v} \rangle.$$

Diamo un modo diverso di verificare se un certo sottoinsieme U dello spazio vettoriale V è un sottospazio.

Proposizione 4.2. *Sia V uno spazio vettoriale e sia U un sottoinsieme di V. Allora U è un sottospazio di V se e solo se*

(V_1) $U \neq \emptyset$;

(V_2) *per ogni* $\mathbf{u}_1, \mathbf{u}_2 \in U$ *e ogni* $\beta_1, \beta_2 \in \mathbb{C}$, $\beta_1\mathbf{u}_1 + \beta\mathbf{u}_2 \in U$.

Dimostrazione. Verifichiamo che da (V_1) e (V_2) seguono (S_1), (S_2) e (S_3). Poiché $U \neq \emptyset$, esiste $\mathbf{u}_0 \in U$. Allora possiamo applicare (V_2) con $\beta_1 = 1$, $\beta_2 = -1$, $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{u}_0$, ottenendo che

$$\beta_1\mathbf{u}_1 + \beta\mathbf{u}_2 = 1\mathbf{u}_0 + (-1)\mathbf{u}_0 = (1-1)\mathbf{u}_0 = \mathbf{0} \in U.$$

Dunque abbiamo provato che vale (S_1). La (S_2) si prova scegliendo $\beta_1 = \beta_2 = 1$. La (S_3) segue prendendo $\beta_1 = \alpha$, $\beta_2 = 0$, $\mathbf{u}_1 = \mathbf{v}$ e $\mathbf{u}_2 = \mathbf{u}_0$.

Verifichiamo ora che da (S_1), (S_2) e (S_3) seguono (V_1) e (V_2). Il fatto che $U \neq \emptyset$ è evidente. Per (S_3), $\beta_1\mathbf{u}_1$ e $\beta_2\mathbf{u}_2$ sono elementi di U, quindi anche $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 \in U$, per ($\mathrm{S}_2$). □

Definizione. Un *insieme finito di vettori* dello spazio vettoriale V è una lista finita di elementi di V che denoteremo con lettere come $\mathcal{A}$, $\mathcal{B}$, eccetera. Scriveremo, per esempio

$$\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}.$$

La notazione non usuale vuole ricordare che stiamo usando un artificio, rispetto al solito concetto di insieme finito; infatti ammetteremo che fra i vettori che compaiono nell'elenco ci possano essere ripetizioni.

Quando nel seguito useremo la locuzione "insieme di vettori" sarà sempre in questo significato, sottintendendo "finito", a meno che non si dica esplicitamente il contrario.

Non è escluso il caso di $n = 0$, nel qual caso parleremo di *insieme vuoto*, indicato anche con $\{\}$.

Esempio 4.3. Una matrice $\mathbf{A} \in M_{m\times n}(\mathbb{C})$ definisce un insieme di vettori in $\mathbb{C}^m$; scrivendo infatti $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_p]$, possiamo associare ad $\mathbf{A}$ l'insieme di n vettori

$$\{\mathbf{a}_1; \mathbf{a}_2; \dots; \mathbf{a}_n\}.$$

Questo è proprio un caso nel quale si possono avere ripetizione, come accade considerando la matrice

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

A ogni insieme di vettori di V possiamo associare un sottospazio di V. Supponiamo per esempio di avere un insieme formato da un solo vettore, cioè $\mathcal{A} = \{\mathbf{v}\}$. A questo insieme possiamo associare allora il sottospazio $\langle\mathbf{v}\rangle$ definito prima.

Più in generale, dato un insieme di vettori $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$, siano $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$. Possiamo allora considerare l'elemento

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n \in V$$

che chiameremo *combinazione lineare* dei vettori dell'insieme $\mathcal{A}$, o semplicemente *combinazione lineare di* $\mathcal{A}$, con coefficienti $\alpha_1, \alpha_2, \dots, \alpha_n$.

DEFINIZIONE. Dato l'insieme di vettori $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ dello spazio vettoriale V, il *sottospazio generato* da $\mathcal{A}$ è l'insieme di tutte le combinazioni lineari dei vettori dell'insieme, che sarà indicato con

$$\langle\mathcal{A}\rangle \text{ oppure con } \langle\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\rangle.$$

Per convenzione, nel caso dell'insieme vuoto, si pone

$$\langle\{\}\rangle = \{\mathbf{0}\}.$$

Il nome dato all'insieme delle combinazioni lineari è giustificato.

PROPOSIZIONE 4.4. *Dato l'insieme di vettori $\mathcal{A}$ dello spazio vettoriale V, $\langle\mathcal{A}\rangle$ è un sottospazio di V.*

DIMOSTRAZIONE. Il fatto è evidente, per definizione, se $\mathcal{A}$ è l'insieme vuoto. Supponiamo allora che $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ e verifichiamo che $\langle\mathcal{A}\rangle$ è un sottospazio.

Possiamo scrivere $\mathbf{0}$ come combinazione lineare di questi vettori con coefficienti tutti nulli:

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_n = \mathbf{0}.$$

Se $\mathbf{u}_1, \mathbf{u}_2 \in \langle\mathcal{A}\rangle$, possiamo scrivere

$$\mathbf{u}_1 = \sum_{i=1}^{n} \alpha_i\mathbf{v}_i, \quad \mathbf{u}_2 = \sum_{i=1}^{n} \alpha'_i\mathbf{v}_i,$$

e, per $\beta_1, \beta_2 \in \mathbb{C}$, abbiamo

$$\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 = \beta_1\sum_{i=1}^{n} \alpha_i\mathbf{v}_i + \beta_2\sum_{i=1}^{n} \alpha'_i\mathbf{v}_i = \sum_{i=1}^{n}(\beta_1\alpha_i + \beta_2\alpha'_i)\mathbf{v}_i$$

e la verifica è completa. □

Il lettore esegua lo sviluppo della formula finale della dimostrazione, per esempio nel caso di $n = 3$, per rendersi conto delle regole di calcolo usate.

ESEMPIO 4.5. Consideriamo l'insieme di vettori in $\mathbb{C}^3$

$$\mathcal{A} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}; \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} \right\}.$$

Una combinazione lineare dei vettori di $\mathcal{A}$ sarà un vettore della forma

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \begin{bmatrix} \alpha_1 + 2\alpha_2 - \alpha_3 \\ -\alpha_1 + \alpha_2 + 4\alpha_3 \\ 0 \end{bmatrix}$$

e, come si può dimostrare per esercizio, ogni vettore della forma

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ 0 \end{bmatrix}$$

può essere scritto come combinazione lineare dei vettori di $\mathcal{A}$. Svolgendo i calcoli si può comprendere come il concetto di combinazione lineare e quello di sistema lineare siano strettamente legati.

Un insieme finito, nel senso usuale, non dipende dall'ordine con cui si elencano i suoi elementi. Lo stesso non accade nella nostra particolare accezione del termine; ma, per esempio, il sottospazio generato dall'insieme $\{\mathbf{u}; \mathbf{v}\}$ è lo stesso del sottospazio generato da $\{\mathbf{v}; \mathbf{u}\}$ a causa del fatto che l'addizione di vettori è commutativa. Ovviamente questo accade per qualunque numero di vettori nell'insieme.

Quando parleremo di *sottoinsieme* di un insieme finito di vettori, intenderemo che ogni vettore del sottoinsieme deve appartenere nell'insieme; inoltre supporremo sempre che, se un certo vettore compare k volte nel sottoinsieme, esso dovrà comparire *almeno* k volte nell'insieme.

Proposizione 4.6. *Se $\mathcal{A}$ è un sottoinsieme di $\mathcal{B}$, allora $\langle\mathcal{A}\rangle$ è un sottospazio di $\langle\mathcal{B}\rangle$.*

Dimostrazione. Di nuovo questo è ovvio nel caso in cui $\mathcal{A}$ sia l'insieme vuoto. Supponiamo perciò che $\mathcal{A}$ abbia m elementi, con $m > 0$. È allora evidente che possiamo scrivere $\mathcal{B}$ in modo che sia

$$\begin{aligned}\mathcal{A} &= \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_m\},\\ \mathcal{B} &= \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_m; \mathbf{v}_{m+1}; \dots; \mathbf{v}_n\}.\end{aligned}$$

Ci basta verificare che ogni vettore che sia combinazione lineare dei vettori di $\mathcal{A}$ è combinazione lineare dei vettori di $\mathcal{B}$. Ora,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_m\mathbf{v}_m = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_m\mathbf{v}_m + 0\mathbf{v}_{m+1} + \dots + 0\mathbf{v}_n$$

e la tesi è provata. □

Osserviamo che questa proposizione giustifica il fatto che si ponga $\langle\{\}\rangle = \{\mathbf{0}\}$, proprio perché l'insieme vuoto è sottoinsieme di ogni altro.

Possiamo anche mettere insieme due insiemi $\mathcal{A}$ e $\mathcal{B}$, semplicemente mettendo il secondo di seguito al primo. Indicheremo l'insieme così ottenuto con $\mathcal{A} \sqcup \mathcal{B}$. È ovvio che $\mathcal{A}$ e $\mathcal{B}$ sono entrambi sottoinsiemi di $\mathcal{A} \sqcup \mathcal{B}$. Si tratta della solita *unione* di insiemi, tenendo però conto delle eventuali ripetizioni di elementi; per questo useremo un simbolo diverso.

5. Insiemi di generatori

Consideriamo lo spazio vettoriale $\mathbb{C}^n$ e, in esso, un insieme finito di vettori particolare: quello ottenuta dalla matrice identità I_n. Porremo

$$\mathbf{I}_n = [\mathbf{e}_1\ \mathbf{e}_2\ \dots\ \mathbf{e}_n].$$

L'insieme $\mathcal{E} = \{\mathbf{e}_1; \mathbf{e}_2; \dots; \mathbf{e}_n\}$ ha una proprietà importante. Se infatti prendiamo $\mathbf{v} \in \mathbb{C}^n$, avremo

$$\mathbf{v} = \begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$$

e, di conseguenza, come si può facilmente verificare,

$$\mathbf{v} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \dots + \alpha_n\mathbf{e}_n$$

e quindi

$$\langle\mathbf{e}_1; \mathbf{e}_2; \dots; \mathbf{e}_n\rangle = \mathbb{C}^n.$$

dal momento che ogni vettore di $\mathbb{C}^n$ si può scrivere come combinazione lineare dei vettori dell'insieme $\mathcal{E}$.

Definizione. Un insieme finito di vettori $\mathcal{A}$ nello spazio vettoriale V si dice un *insieme di generatori* di V se $\langle\mathcal{A}\rangle = V$, cioè ogni vettore di V si può scrivere come combinazione lineare dei vettori dell'insieme $\mathcal{A}$.

ESEMPI 5.1. Lo spazio vettoriale nullo $\{\mathbf{0}\}$ ammette come insieme di generatori sia l'insieme vuoto che l'insieme $\{\mathbf{0}\}$.

Se $n > 0$, l'insieme $\{1; X; X^2; \dots; X^{n-1}\}$ è un insieme di generatori di $\mathcal{P}_n(\mathbb{C})$.

Viceversa, non esistono insiemi finiti di generatori di $\mathcal{P}(\mathbb{C})$. Sia infatti $\mathcal{A}$ un insieme finito in $\mathcal{P}(\mathbb{C})$; se ogni elemento di $\mathcal{A}$ è il polinomio nullo, il sottospazio generato è $\{0\}$. Altrimenti fra gli elementi di $\mathcal{A}$ ce ne sarà uno avente il grado massimo, diciamo $k \geq 0$. Una combinazione lineare degli elementi di $\mathcal{A}$ non può allora avere grado maggiore di k e, in particolare, $X^{k+1} \notin \langle\mathcal{A}\rangle$.

L'esempio seguente mostra che esiste una connessione fra insiemi di generatori e sistemi lineari.

ESEMPIO 5.2. Consideriamo i seguenti vettori di $\mathbb{C}^2$:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

e dimostriamo che $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3\}$ è un insieme di generatori di $\mathbb{C}^2$. Preso un qualunque vettore $\mathbf{v} = [r\ s]^T \in \mathbb{C}^2$, dobbiamo poterlo scrivere come combinazione lineare dei tre vettori dati, cioè trovare numeri complessi α_1, α_2 e α_3 in modo che

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3$$

o, esplicitamente

$$\begin{cases} 2\alpha_1 - \alpha_2 + \alpha_3 = r \\ \alpha_1 + 3\alpha_2 - 2\alpha_3 = s \end{cases}$$

e si verifica facilmente che tale sistema ha soluzione per ogni $r, s \in \mathbb{C}$.

Il lettore può osservare che anche l'insieme $\{\mathbf{v}_1; \mathbf{v}_2\}$ è un insieme di generatori di $\mathbb{C}^2$, mentre nessun insieme con un solo elemento può essere un insieme di generatori.

DEFINIZIONE. Diremo che lo spazio vettoriale V è *finitamente generato* se ammette un insieme finito di generatori.

Nel seguito, tranne esplicito avviso in contrario, considereremo solo spazi finitamente generati. Lo spazio vettoriale $\mathcal{P}(\mathbb{C})$ non è finitamente generato. La teoria si può generalizzare anche per spazi vettoriali non finitamente generati, che sono fra l'altro molto utili soprattutto in connessione con il concetto di prodotto scalare che si vedrà nel Capitolo 3 e per applicazioni alla fisica. Per i nostri scopi, che sono quelli legati ai sistemi lineari, gli spazi finitamente generati saranno sufficienti.

Per definizione, dato un insieme finito di vettori $\mathcal{A}$ in V, il sottospazio $\langle\mathcal{A}\rangle$ è finitamente generato e $\mathcal{A}$ ne è un insieme di generatori.

PROPOSIZIONE 5.3. *Lo spazio vettoriale delle matrici $M_{m\times n}(\mathbb{C})$ è finitamente generato.*

DIMOSTRAZIONE. Esercizio: si considerino le matrici che hanno un solo coefficiente uguale a 1 e tutti gli altri 0; esse sono mn. Si può scrivere ogni matrice come combinazione lineare di queste? □

PROPOSIZIONE 5.4. *Se $\mathcal{A}$ e $\mathcal{B}$ sono insiemi di vettori nello spazio vettoriale V tali che*

(i) *$\mathcal{A}$ è un sottoinsieme di $\mathcal{B}$;*
(ii) *$\mathcal{A}$ è un insieme di generatori di V.*

Allora anche $\mathcal{B}$ è un insieme di generatori.

DIMOSTRAZIONE. Dalle ipotesi segue che

$$V = \langle\mathcal{A}\rangle \subseteq \langle\mathcal{B}\rangle \subseteq V$$

e quindi la tesi. □

In certi casi è possibile anche eseguire un'operazione inversa.

PROPOSIZIONE 5.5. *Se* $\{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ *è un insieme di generatori di* V *e* $\mathbf{v}_n \in \langle \mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_{n-1} \rangle$, *allora anche* $\{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_{n-1}\}$ *è un insieme di generatori di* V.

DIMOSTRAZIONE. Sia $\mathbf{v} \in V$; sappiamo scrivere

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

per opportuni coefficienti. Per ipotesi,

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \beta_i \mathbf{v}_i$$

per opportuni coefficienti. Ci basta allora sostituire nella prima uguaglianza e ottenere

$$\mathbf{v} = \sum_{i=1}^{n-1} \alpha_i \mathbf{v}_i + \sum_{i=1}^{n-1} \alpha_n \beta_i \mathbf{v}_i = \sum_{i=1}^{n-1} (\alpha_i + \alpha_n \beta_i) \mathbf{v}_i$$

come richiesto. □

La proposizione precedente dice che, ogni volta che un vettore di un insieme è combinazione lineare degli altri, lo possiamo togliere dall'insieme senza cambiare lo spazio generato. Il fatto che nella proposizione abbiamo parlato dell'ultimo vettore è ovviamente non rilevante, perché il sottospazio generato da un insieme di vettori non dipende dall'ordine.

ESEMPIO 5.6. Consideriamo la matrice

$$\mathbf{U} = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3\ \mathbf{u}_4\ \mathbf{u}_5] = \begin{bmatrix} 1 & -2 & 2 & 1 & 3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

che è in forma ridotta. Possiamo scrivere

$$\mathbf{u}_4 = 3\mathbf{u}_1 + 0\mathbf{u}_2 - \mathbf{u}_3 + 0\mathbf{u}_5.$$

e, se un vettore $\mathbf{v} \in \mathbb{C}^3$ è combinazione lineare di $\{\mathbf{u}_1; \mathbf{u}_2; \mathbf{u}_3; \mathbf{u}_4; \mathbf{u}_5\}$, per esempio

$$\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3 + \alpha_4 \mathbf{u}_4 + \alpha_5 \mathbf{u}_5$$

abbiamo anche

$$\begin{aligned} \mathbf{v} &= \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3 + \alpha_4 (3\mathbf{u}_1 + 0\mathbf{u}_2 - \mathbf{u}_3 + 0\mathbf{u}_5) + \alpha_5 \mathbf{u}_5 \\ &= (\alpha_1 + 3\alpha_4)\mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + (\alpha_3 - \alpha_4)\mathbf{u}_3 + \alpha_5 \mathbf{u}_5. \end{aligned}$$

Abbiamo quindi che

$$V = \langle \mathbf{u}_1; \mathbf{u}_2; \mathbf{u}_3; \mathbf{u}_4; \mathbf{u}_5 \rangle = \langle \mathbf{u}_1; \mathbf{u}_2; \mathbf{u}_3; \mathbf{u}_5 \rangle.$$

Il lettore verifichi anche che

$$V = \langle \mathbf{u}_1; \mathbf{u}_3; \mathbf{u}_5 \rangle$$

ma che non è più possibile eliminare altri vettori dall'insieme senza cambiare il sottospazio generato (per esempio, $\mathbf{u}_5$ non è combinazione lineare di $\mathbf{u}_1$ e di $\mathbf{u}_3$).

6. Indipendenza lineare

La nozione di insieme di generatori corrisponde, come ora vedremo, a quella di sistema lineare risolubile. Il concetto collegato di sistema con al più una soluzione corrisponde, come vedremo in seguito, a quello di insieme linearmente indipendente.

Consideriamo il sistema lineare di m equazioni in n incognite $\mathbf{A}\mathbf{x} = \mathbf{b}$ e supponiamo che $[\alpha_1\ \alpha_2\ \dots\ \alpha_n]^T$ e $[\beta_1\ \beta_2\ \dots\ \beta_n]^T$ siano soluzioni distinte del sistema. Se scriviamo $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_n]$, questo si esprime con le relazioni

$$\begin{aligned} \mathbf{b} &= \alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \dots + \alpha_n \mathbf{a}_n \\ &= \beta_1 \mathbf{a}_1 + \beta_2 \mathbf{a}_2 + \dots + \beta_n \mathbf{a}_n \end{aligned}$$

e dunque il vettore $\mathbf{b}$ si può scrivere in almeno due modi come combinazione lineare dell'insieme delle colonne di $\mathbf{A}$. In effetti il sistema ha infinite soluzioni.

Se poniamo $\gamma_i = \alpha_i - \beta_i$ $(i = 1, 2, \dots, n)$, abbiamo che

$$\gamma_1\mathbf{a}_1 + \gamma_2\mathbf{a}_2 + \dots + \gamma_n\mathbf{a}_n = \mathbf{0}$$

e almeno uno fra i coefficienti γ_i è diverso da 0 (altrimenti le due soluzioni da cui siamo partiti non sarebbero distinte).

Definizione. Un insieme di vettori $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ si dice *linearmente dipendente* se esistono scalari α_1, α_2, …, α_n *non tutti nulli* tali che

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_n\mathbf{v}_n = \mathbf{0}.$$

Attenzione: basta che uno solo dei coefficienti sia diverso da 0. Per esempio, ogni insieme in cui compaia il vettore nullo è linearmente dipendente. Infatti, se $\mathcal{A} = \{\mathbf{v}_1 = \mathbf{0}; \mathbf{v}_2; \dots; \mathbf{v}_n\}$, possiamo prendere $\alpha_1 = 1$ e $\alpha_i = 0$, per $i = 2, \dots, n$. Evidentemente

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}.$$

Esempio 6.1. Consideriamo l'insieme di vettori in $\mathbb{C}^3$ già visto nell'esempio 4.5:

$$\mathcal{A} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}; \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} \right\}.$$

Una combinazione lineare dei vettori di $\mathcal{A}$ sarà un vettore della forma

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \begin{bmatrix} \alpha_1 + 2\alpha_2 - \alpha_3 \\ -\alpha_1 + \alpha_2 + 4\alpha_3 \\ 0 \end{bmatrix}$$

ed è facile verificare che, per $\alpha_1 = 3$, $\alpha_2 = -1$ e $\alpha_3 = 1$, la combinazione lineare corrispondente è il vettore nullo.

Proposizione 6.2. *Se $\mathcal{A}$ è un sottoinsieme non vuoto dell'insieme di vettori $\mathcal{B}$ e $\mathcal{A}$ è linearmente dipendente, allora anche $\mathcal{B}$ è linearmente dipendente.*

Dimostrazione. Possiamo supporre che

$$\mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_m\}$$
$$\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_m; \mathbf{v}_{m+1}; \dots; \mathbf{v}_n\}$$

e che

$$\sum_{i=1}^{m} \alpha_i\mathbf{v}_i = \mathbf{0},$$

dove i coefficienti α_i $(i = 1, \dots, m)$ non sono tutti nulli. Poniamo $\alpha_{m+1} = \dots = \alpha_n = 0$ e abbiamo

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \sum_{i=1}^{m} \alpha_i\mathbf{v}_i + \sum_{i=m+1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0},$$

e quindi la tesi. □

Esempio 6.3. Se nell'insieme $\mathcal{A}$ ci sono due vettori uguali, esso è linearmente dipendente. Vista la proposizione precedente, basta dimostrare che l'insieme $\{\mathbf{v}; \mathbf{v}\}$ è linearmente dipendente. In effetti

$$1\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}$$

e abbiamo la tesi.

Di fatto la nozione più importante è il contrario della dipendenza lineare.

DEFINIZIONE. Un insieme di vettori $\mathcal{A}$ è *linearmente indipendente* se non è linearmente dipendente. Per convenzione, anche l'insieme vuoto è linearmente indipendente.

Come possiamo allora verificare se un insieme di vettori è linearmente indipendente? La cosa è ovvia per l'insieme vuoto: un insieme linearmente dipendente ha, per definizione, almeno un elemento.

PROPOSIZIONE 6.4. *L'insieme $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ è linearmente indipendente se e solo se, per $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$, da*

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n = \mathbf{0}$$

segue che $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_n = 0$.

DIMOSTRAZIONE. Si tratta semplicemente di applicare la definizione di insieme linearmente dipendente. □

ESEMPIO 6.5. L'insieme $\mathcal{E}$ delle colonne dell'identità $\mathbf{I}_n$ è un insieme linearmente indipendente di vettori di $\mathbb{C}^n$. Infatti, da

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \dots + \alpha_n\mathbf{e}_n = \mathbf{0}$$

segue che $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_n = 0$. Il lettore scriva esplicitamente i calcoli per $n = 4$.

L'insieme di vettori in $\mathbb{C}^4$

$$\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 0 \end{bmatrix}; \mathbf{v}_3 = \begin{bmatrix} 0 \\ 3 \\ 2 \\ 0 \end{bmatrix} \right\}$$

è linearmente indipendente. Infatti, scrivere

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

equivale a scrivere il sistema lineare

$$\begin{cases} \alpha_1 + 2\alpha_2 = 0 \\ -\alpha_1 + \alpha_2 + 3\alpha_3 = 0 \\ 2\alpha_1 + 2\alpha_2 + 2\alpha_3 = 0 \end{cases}$$

il quale ammette solo la soluzione $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$.

Vediamo subito alcune conseguenze della definizione.

PROPOSIZIONE 6.6. *Se $\mathcal{A}$ è un insieme linearmente indipendente, ogni suo sottoinsieme è ancora linearmente indipendente.*

DIMOSTRAZIONE. Se un sottoinsieme di $\mathcal{A}$ fosse linearmente dipendente, anche $\mathcal{A}$ lo sarebbe. □

PROPOSIZIONE 6.7. *Se $\mathcal{A}$ è un insieme linearmente indipendente, esso consiste di vettori a due a due distinti e non nulli.*

ESEMPIO 6.8. Se $\mathbf{v} \in V$ e $\mathbf{v} \neq \mathbf{0}$, allora $\{\mathbf{v}\}$ è linearmente indipendente. Infatti, se $\alpha\mathbf{v} = \mathbf{0}$, sappiamo che o $\alpha = 0$ oppure $\mathbf{v} = \mathbf{0}$.

Sotto certe condizioni possiamo aggiungere vettori a un insieme linearmente indipendente ottenendo ancora un insieme linearmente indipendente. La dimostrazione che segue è un buon esempio per capire la tecnica che si usa per verificare l'indipendenza lineare.

TEOREMA 6.9. *Sia $\mathcal{A}$ un insieme linearmente indipendente nello spazio vettoriale V e sia $\mathbf{v} \in V$ tale che $\mathbf{v} \notin \langle\mathcal{A}\rangle$. Allora l'insieme $\mathcal{B} = \mathcal{A} \sqcup \{\mathbf{v}\}$ è linearmente indipendente.*

DIMOSTRAZIONE. La cosa è ovvia se $\mathcal{A}$ è l'insieme vuoto, per l'esempio precedente. Supponiamo allora $\mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$. Prendiamo scalari $\alpha_1, \dots, \alpha_n$ e α tali che

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n + \alpha\mathbf{v} = \mathbf{0}.$$

Supponiamo, per assurdo, che $\alpha \neq 0$. Allora

$$-\alpha\mathbf{v} = \sum_{i=1}^{n} \alpha_i\mathbf{v}_i$$

e, ponendo $\beta_i = -\alpha^{-1}\alpha_i$, otteniamo

$$\mathbf{v} = \sum_{i=1}^{n} \beta_i\mathbf{v}_i$$

da cui seguirebbe $\mathbf{v} \in \langle\mathcal{A}\rangle$, contraddizione.

Dunque $\alpha = 0$ e quindi anche

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}.$$

Per ipotesi l'insieme $\mathcal{A}$ è linearmente indipendente, quindi anche $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_n = 0$, e, come si voleva, tutti gli scalari sono nulli. □

ESEMPIO 6.10. Sia $\mathbf{U}$ una matrice $n \times p$ in forma ridotta e siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ le sue colonne dominanti. Allora l'insieme di vettori $\{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_k\}$ in $\mathbb{C}^n$ è linearmente indipendente.

Se infatti ricordiamo come risultano le colonne dominanti dopo la EG, possiamo scrivere

$$\mathbf{v}_i = \begin{bmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix}$$

dove gli asterischi indicano i coefficienti che effettivamente compaiono. Completeremo la dimostrazione per induzione.

Se c'è una sola colonna dominante, l'insieme è formato da un solo vettore non nullo e quindi è linearmente indipendente per l'esempio precedente. Supponiamo allora che $k > 1$; la matrice $\mathbf{U}'$ che ha come colonne quelle precedenti l'ultima colonna dominante di $\mathbf{U}$ è ancora in forma ridotta e ha esattamente come colonne dominanti le prime $k-1$ colonne dominanti di $\mathbf{U}$. Per ipotesi induttiva l'insieme di queste colonne è linearmente indipendente. Ci basta allora verificare che $\mathbf{v}_k \notin \langle\mathbf{v}_1; \dots; \mathbf{v}_{k-1}\rangle$ e applicare il teorema.

Ma una combinazione lineare delle prime $k-1$ colonne dominanti ha tutti i coefficienti dalla k-esima riga in poi nulli, e quindi non può essere uguale a $\mathbf{v}_k$ che sulla k-esima riga ha 1.

Per via della proposizione 6.6 ogni insieme formato da colonne dominanti (distinte) di una matrice in forma ridotta è linearmente indipendente.

ESEMPIO 6.11. Consideriamo una matrice reale $\mathbf{U}$ di forma $n \times p$ che sia in forma ridotta. Allora l'insieme delle colonne non nulle di $\mathbf{U}^T$ è linearmente indipendenti in $\mathbb{R}^p$ (esercizio).

Per esempio, si consideri la matrice

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

Scriviamo $\mathbf{U}^T = [\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3\ \mathbf{r}_4]$ e

$$\mathbf{0} = \alpha_1\mathbf{r}_1 + \alpha_2\mathbf{r}_2 + \alpha_3\mathbf{r}_3$$

che equivale al sistema lineare

$$\begin{cases} \alpha_1 = 0 \\ -\alpha_1 = 0 \\ 2\alpha_1 + \alpha_2 = 0 \\ 3\alpha_1 + \alpha_2 = 0 \\ 5\alpha_1 + 2\alpha_2 + \alpha_3 = 0 \\ 4\alpha_1 - 2\alpha_2 = 0 \end{cases}$$

il quale ha come unica soluzione $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$.

7. Basi

Ora metteremo insieme i due concetti visti prima. Osserviamo che la procedura esposta nella proposizione 5.5 permette di eliminare da un insieme di generatori vettori che siano combinazione lineare degli altri; quando questa non può più essere eseguita, rimaniamo con un insieme linearmente indipendente.

Proposizione 7.1. *Un insieme di vettori $\mathcal{A}$ è linearmente indipendente se e solo se nessuno dei vettori di $\mathcal{A}$ è combinazione lineare degli altri.*

Dimostrazione. Supponiamo che $\mathcal{A}$ sia linearmente indipendente; se è l'insieme vuoto, evidentemente non ci sono vettori che possano essere combinazione lineare degli altri (perché non ce ne sono). Sia allora $\mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ e supponiamo per assurdo che $\mathbf{v}_n$ sia combinazione lineare dei rimanenti (non è restrittivo, cambiando eventualmente l'ordine dell'elenco). Dunque possiamo scrivere

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \alpha_i\mathbf{v}_i,$$

da cui, ponendo $\alpha_n = -1$, otteniamo

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}$$

che, essendo l'insieme linearmente indipendente, implica in particolare $\alpha_n = 0$, che è assurdo.

Supponiamo allora che nessuno dei vettori di $\mathcal{A}$ sia combinazione lineare dei rimanenti e dimostriamo che $\mathcal{A}$ è linearmente indipendente. Se $\alpha_i \in \mathbb{C}$ $(i = 1, \dots, n)$ e

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}$$

con $\alpha_n \neq 0$, possiamo scrivere anche

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \beta_i\mathbf{v}_i$$

dove $\beta_i = -\alpha_n^{-1}\alpha_i$. Questo non è possibile, per ipotesi e dunque $\alpha_n = 0$. Allo stesso modo possiamo dimostrare che $\alpha_i = 0$ $(i = 1, \dots, n-1)$. □

Da questa proposizione discende dunque che la procedura di 5.5 a un certo punto termina e precisamente quando l'insieme ottenuto è linearmente indipendente. Ricordiamo che uno spazio vettoriale V si dice finitamente generato quando ha un insieme (finito) di generatori.

Definizione. Un insieme $\mathcal{A}$ di vettori di V si dice una *base* di V se è un insieme di generatori ed è linearmente indipendente.

Teorema 7.2. *Sia V uno spazio vettoriale finitamente generato. Allora V ha una base.*

Dimostrazione. Sia $\mathcal{A}$ un insieme di generatori di V. Applichiamo la procedura di 5.5 fino a che è possibile, ottenendo un insieme $\mathcal{B}$ che è ancora un insieme di generatori ed è anche linearmente indipendente. □

Esempio 7.3. Lo spazio vettoriale $\mathbb{C}^n$ ha come base l'insieme $\mathcal{E}$ dei vettori coordinati

$$\mathcal{E} = \{\mathbf{e}_1; \mathbf{e}_2; \dots; \mathbf{e}_n\}.$$

La verifica è già stata fatta in esempi precedenti, ma la ripetiamo.

Ogni elemento di $\mathbb{C}^n$ è combinazione lineare di questi vettori: se $\mathbf{v} = [\alpha_1\ \alpha_2\ \dots\ \alpha_n]^T$, possiamo scrivere

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i$$

e dunque $\mathcal{E}$ è un insieme di generatori.

Inoltre,

$$\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = [\alpha_1\ \alpha_2\ \dots\ \alpha_n]^T$$

e perciò, da $\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = \mathbf{0}$ segue che $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_n = 0$, cioè che $\mathcal{E}$ è anche linearmente indipendente.

Chiameremo $\mathcal{E}$ la *base canonica* di $\mathbb{C}^n$.

Esempio 7.4. Lo spazio vettoriale nullo $\{\mathbf{0}\}$ ha come base l'insieme vuoto. Infatti l'insieme vuoto è un insieme di generatori di $\{\mathbf{0}\}$, per definizione, ed è linearmente indipendente.

Esempio 7.5. L'insieme $\{p_1 = 1 - X; p_2 = 1 + X; p_3 = 1 - X^2\}$ è una base di $\mathcal{P}_3(\mathbb{C})$. Per verificarlo dimostriamo che è sia un insieme di generatori che un insieme linearmente indipendente. Come vedremo, si tratta di analizzare sistemi lineari. Cerchiamo di capire quello che occorre vedere: intanto che ogni polinomio $a_0 + a_1X + a_2X^2$ è combinazione lineare di $\{p_1; p_2; p_3\}$, cioè che esistono α_1, α_2 e α_3 tali che

$$a_0 + a_1X + a_2X^2 = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3.$$

Questo si traduce nel sistema lineare

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = a_0 \\ -\alpha_1 + \alpha_2 = a_1 \\ -\alpha_3 = a_2 \end{cases}$$

che ha come matrice

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & a_0 \\ -1 & 1 & 0 & a_1 \\ 0 & 0 & -1 & a_2 \end{array}\right]$$

e una semplice EG porta alla forma ridotta

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & a_0 \\ 0 & 1 & 1/2 & (a_0 + a_1)/2 \\ 0 & 0 & 1 & -a_2 \end{array}\right]$$

da cui si vede che ogni sistema del genere, al variare di a_0, a_1 e a_2 ha un'unica soluzione. Dunque la soluzione esiste e l'insieme è di generatori; inoltre, quando $a_0 = 0$, $a_1 = 0$ e $a_2 = 0$, cioè il sistema è omogeneo, anch'esso ha un'unica soluzione, quella con $\alpha_1 = 0$, $\alpha_2 = 0$ e $\alpha_3 = 0$, perciò l'insieme dato è anche linearmente indipendente.

Ovviamente anche l'insieme $\{1; X; X^2\}$ è una base di $\mathcal{P}_3(\mathbb{C})$, come si verifica immediatamente: il sistema lineare costruito come il precedente ha come matrice dei coefficienti l'identità. Di fatto ogni spazio vettoriale di dimensione finita ha infinite basi, con l'eccezione dello spazio vettoriale nullo che ha come unica base l'insieme vuoto.

ESEMPIO 7.6. Se $\mathcal{B} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$ è una base di V, per ogni scalare $\gamma \neq 0$, anche

$$\mathcal{B}' = \{\mathbf{v}'_1 = \gamma\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$$

è una base di V.

Dimostriamo infatti che $\mathcal{B}'$ è un insieme di generatori. Se $\mathbf{v} \in V$, sappiamo che esistono scalari α_1, α_2, ..., α_n tali che

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n$$

e quindi

$$\mathbf{v} = (\alpha_1\gamma^{-1})\mathbf{v}'_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n$$

cioè $\mathbf{v} \in \langle\mathcal{B}'\rangle$.

Lasciamo come esercizio la dimostrazione che $\mathcal{B}'$ è linearmente indipendente.

Negli esempi che abbiamo illustrato abbiamo visto che basi diverse dello stesso spazio vettoriale hanno qualcosa in comune: il numero di elementi. Questo non è casuale, anzi è un fatto fondamentale della teoria degli spazi vettoriali. Il teorema che segue è dovuto a Steinitz, un suo corollario è appunto che due basi di uno stesso spazio vettoriale hanno lo stesso numero di elementi.

La dimostrazione di questo teorema è importantissima, perché usa una tecnica ingegnosa: se $\mathcal{B}$ è un insieme linearmente indipendente e $\mathcal{A}$ un insieme di generatori, entrambi nello spazio vettoriale V, allora possiamo mettere ogni vettore di $\mathcal{B}$ al posto di uno di $\mathcal{A}$ ottenendo ancora un insieme di generatori.

Ricordiamo prima due fatti. Primo: quando diciamo che un insieme finito di vettori ha n elementi, intendiamo contare anche gli elementi eventualmente ripetuti. Secondo: in un insieme linearmente indipendente non ci sono ripetizioni.

TEOREMA 7.7. *Sia V uno spazio vettoriale finitamente generato. Siano $\mathcal{A}$ rispettivamente un insieme di generatori di V e $\mathcal{B}$ un insieme linearmente indipendente di vettori di V. Se $\mathcal{A}$ ha n elementi e $\mathcal{B}$ ha m elementi, allora $m \leq n$.*

DIMOSTRAZIONE. Se $\mathcal{B}$ è l'insieme vuoto, non c'è niente da dimostrare; quindi assumiamo che $m > 0$. Allora, necessariamente, $n \geq 1$: altrimenti $\mathcal{A}$ sarebbe l'insieme vuoto che genera lo spazio vettoriale nullo, nel quale non potrebbero esistere m vettori non nulli distinti (si veda 6.7).

Supponiamo perciò

$$\mathcal{B} = \{\mathbf{w}_1; \dots; \mathbf{w}_m\}, \qquad \mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}.$$

Procederemo per induzione, dimostrando la seguente affermazione: *Se $0 \leq k \leq m$, esiste un insieme di generatori $\mathcal{C}_k$ formato da k vettori dell'insieme $\mathcal{B}$ e da $n - k$ vettori dell'insieme $\mathcal{A}$.*

Chiaramente questo prova che $m \leq n$, altrimenti non potremmo costruire l'insieme $\mathcal{C}_m$.

Il passo base, $k = 0$, è ovvio: basta prendere $\mathcal{C}_0 = \mathcal{A}$. Supponiamo allora di avere a disposizione l'insieme $\mathcal{C}_k$ e che $k + 1 \leq m$. Non è restrittivo supporre che

$$\mathcal{C}_k = \{\mathbf{w}_1; \dots; \mathbf{w}_k; \mathbf{v}_{k+1}; \dots; \mathbf{v}_n\}$$

eventualmente riordinando gli insiemi da cui siamo partiti. Per ipotesi induttiva, $\mathcal{C}_k$ è un insieme di generatori, quindi possiamo scrivere

$$(*) \qquad \mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i\mathbf{w}_i + \sum_{j=k+1}^{n} \beta_j\mathbf{v}_j.$$

Non è possibile che $\beta_{k+1} = \dots = \beta_n = 0$, perché in tal caso

$$\mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i\mathbf{w}_i$$

e l'insieme $\mathcal{B}$ non sarebbe linearmente indipendente. Dunque possiamo supporre che $\beta_{k+1} \neq 0$, eventualmente riordinando $\mathcal{A}$. Consideriamo l'insieme

$$\mathcal{C}'_k = \{\mathbf{w}_1; \dots; \mathbf{w}_k; \mathbf{w}_{k+1}; \mathbf{v}_{k+1}; \dots; \mathbf{v}_n\}.$$

Poiché esso contiene un insieme di generatori, anche $\mathcal{C}'_k$ è un insieme di generatori. La relazione (*) si può anche scrivere

$$\mathbf{v}_{k+1} = \sum_{i=1}^{k+1} \gamma_i \mathbf{w}_i + \sum_{j=k+2}^{n} \delta_j \mathbf{v}_j$$

(si calcolino i coefficienti per esercizio) e quindi l'insieme che si ottiene eliminando da $\mathcal{C}'_k$ il vettore $\mathbf{v}_{k+1}$ è ancora un insieme di generatori (per 5.5). Chiamiamo $\mathcal{C}_{k+1}$ questo insieme, perché ha le proprietà richieste. □

La tecnica della dimostrazione è importante di per sé e quindi la isoliamo in un risultato a parte.

Teorema 7.8. *Siano $\mathcal{A}$ è un insieme di generatori dello spazio vettoriale V e $\mathcal{B}$ una insieme linearmente indipendente di vettori di V. Supponiamo che $\mathcal{A}$ abbia n elementi e che $\mathcal{B}$ abbia m elementi. Allora esiste un insieme di generatori $\mathcal{C}$ formato da tutti i vettori di $\mathcal{B}$ e da $n - m$ vettori di $\mathcal{A}$.*

Siccome ogni spazio che consideriamo ha un insieme di generatori, possiamo usare questo per ottenere che *ogni insieme linearmente indipendente può essere completato a una base.*

Esempio 7.9. Per illustrare il risultato precedente, applichiamo la tecnica della dimostrazione del Teorema 7.7 agli insiemi $\mathcal{A} = \{\mathbf{e}_1; \mathbf{e}_2; \mathbf{e}_3; \mathbf{e}_4\}$ e $\mathcal{B} = \{\mathbf{u}; \mathbf{v}\}$ in $\mathbb{C}^4$, dove

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Si esegue l'EG sulla matrice [$\mathbf{u}\ \mathbf{v}\ \mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3\ \mathbf{e}_4$:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 \end{bmatrix}$$
$$\to \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1/2 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}.$$

Le colonne dominanti sono le prime quattro e quindi l'insieme $\{\mathbf{u}; \mathbf{v}; \mathbf{e}_1; \mathbf{e}_2\}$ è linearmente indipendente, quindi una base di $\mathbb{C}^4$.

Il teorema appena dimostrato ha una conseguenza importantissima. Di solito lo si riassume dicendo che *ogni insieme linearmente indipendente in V può essere completato a una base.* Ci basta infatti considerare come insieme di generatori quello che esiste per ipotesi, visto che lo spazio vettoriale in cui lavoriamo è finitamente generato.

Teorema 7.10. *Sia V uno spazio vettoriale di dimensione finita; siano $\mathcal{A}$ e $\mathcal{B}$ basi di V. Allora $\mathcal{A}$ e $\mathcal{B}$ hanno lo stesso numero di elementi.*

Dimostrazione. Sia n il numero di elementi di $\mathcal{A}$ e sia m il numero di elementi di $\mathcal{B}$.

- Siccome $\mathcal{A}$ è un insieme di generatori e $\mathcal{B}$ è linearmente indipendente, il teorema dice che $m \leq n$.
- Siccome $\mathcal{B}$ è un insieme di generatori e $\mathcal{A}$ è linearmente indipendente, il teorema dice che $n \leq m$.

Dunque deve essere $m = n$. □

Abbiamo già visto che ogni spazio vettoriale V finitamente generato ha una base. Ciò che abbiamo dimostrato è che due basi di V devono avere lo stesso numero di elementi. Si noti che V può avere molte basi (di fatto ne ha infinite, se non è lo spazio nullo); tuttavia due basi diverse hanno in comune il numero di elementi.

Definizione. Se V è uno spazio vettoriale di dimensione finita, indicheremo con $\dim V$ il numero di elementi di una sua base; $\dim V$ si chiama *dimensione* di V.

Esempi 7.11. (1) Lo spazio $\mathbb{C}^n$ ha dimensione n, perché la base canonica $\mathcal{E}$ ha n elementi.
(2) Sia $\mathbf{U}$ una matrice $m \times n$ in forma ridotta. Scriviamola come

$$\mathbf{U} = [\mathbf{u}_1\ \mathbf{u}_2\ \dots\ \mathbf{u}_n]$$

e consideriamo $V = \langle \mathbf{u}_1; \mathbf{u}_2; \dots; \mathbf{u}_n \rangle$, che è un sottospazio di $\mathbb{C}^n$. Per quanto visto nell'esempio 6.10, l'insieme delle colonne dominanti è linearmente indipendente. Ogni altra colonna è combinazione lineare delle colonne dominanti che la precedono, come indica l'Esempio 5.6 forma una base di V. Quindi, se k è il numero delle colonne dominanti (che è uguale al numero di righe non nulle), $\dim V = k$.
(3) Lo spazio $\{\mathbf{0}\}$ ha dimensione 0; una sua base è la l'insieme vuoto.

8. Calcolo della dimensione

Esistono vari modi di determinare o almeno valutare, nel senso di maggiorare o minorare, la dimensione di uno spazio vettoriale. Uno è di trovarne una base, ma in certe situazioni si deve trattare il problema in modo diverso.

Cominciamo con un lemma che dimostra un fatto intuitivamente chiaro, ma che non è del tutto ovvio, come si vedrà.

Lemma 8.1. *Se V è uno spazio vettoriale di dimensione finita e U è un sottospazio di V, allora U ha dimensione finita e* $\dim U \le \dim V$.

Dimostrazione. Procederemo per assurdo, assumendo quindi che U non abbia dimensione finita. Ciò significa che nessun insieme di vettori di U è un insieme di generatori. In particolare $U \ne \{\mathbf{0}\}$, perché altrimenti l'insieme vuoto sarebbe un insieme di generatori. Dunque esiste $\mathbf{u}_1 \in U$, $\mathbf{u}_1 \ne \mathbf{0}$.

L'insieme $\mathcal{A}_1 = \{\mathbf{u}_1\}$ non è di generatori di U, quindi esiste un vettore $\mathbf{u}_2 \in U$ tale che $\mathbf{u}_2 \notin \langle \mathcal{A}_1 \rangle$. Ma allora (si veda 6.9) l'insieme $\mathcal{A}_2 = \{\mathbf{u}_1; \mathbf{u}_2\}$ è linearmente indipendente.

Possiamo andare avanti per induzione allo stesso modo: se abbiamo ottenuto l'insieme linearmente indipendente $\mathcal{A}_k$ in U, questo non è di generatori e quindi, prendendo

$$\mathbf{u}_{k+1} \in U \setminus \langle \mathcal{A}_k \rangle$$

l'insieme $\mathcal{A}_{k+1} = \mathcal{A}_k \sqcup \{\mathbf{v}_{k+1}\}$ è linearmente indipendente.

Lo spazio V abbia dimensione n: allora esiste un insieme di generatori di lunghezza n. Ma abbiamo costruito con il procedimento induttivo l'insieme $\mathcal{A}_{n+1}$ che è linearmente indipendente in U, quindi anche in V. Questo contraddice il teorema 7.7.

Dunque U è finitamente generato e perciò ha una base che avrà m elementi, con $m \le \dim V$, sempre per il teorema 7.7. □

Teorema 8.2. *Sia V uno spazio vettoriale di dimensione finita. Allora* $\dim V$ *è il numero minimo di elementi di un insieme di generatori di V.*

Dimostrazione. Sia $\mathcal{A}$ un insieme di generatori con numero minimo n di elementi. Ciò significa che ogni altro insieme di generatori ha numero di elementi $\ge n$. Se $\mathcal{A}$ non è linearmente indipendente, possiamo ottenerne un insieme di generatori togliendo un vettore (5.5). Assurdo. □

Teorema 8.3. *Sia V uno spazio vettoriale finitamente generato. Allora* $\dim V$ *è il numero massimo di elementi di un insieme linearmente indipendente in V.*

Dimostrazione. Il numero massimo di elementi esiste, perché un insieme linearmente indipendente ha al più tanti elementi quanti un insieme di generatori (7.7). Se $\mathcal{A}$ è un insieme linearmente indipendente con numero massimo di vettori e $\langle\mathcal{A}\rangle \neq V$, potremmo aggiungere ad $\mathcal{A}$ un vettore ottenendo ancora un insieme linearmente indipendente (si veda 6.9). Assurdo. $\square$

Teorema 8.4. *Se* $\dim V = n$ *e* $\mathcal{A}$ *è un insieme linearmente indipendente di vettori di* V *con* n *elementi, allora* $\mathcal{A}$ *è una base di* V.

Dimostrazione. Se $\mathcal{A}$ non fosse una base di V, il teorema 7.8 ci permette di completarla a una base che avrebbe $m > n$ elementi. Assurdo. $\square$

Come conseguenza di questo fatto otteniamo una tecnica per dimostrare che un sottospazio U di V coincide con V: infatti $U = V$ se e solo se $\dim U = \dim V$. Infatti un'implicazione è ovvia. Se invece $\dim U = \dim V$, troviamo in U una base il cui numero di elementi è uguale al numero massimo di elementi di un insieme linearmente indipendente in V. Essa perciò è una base di V e ogni vettore di V è combinazione lineare di vettori di U, quindi appartiene a U.

Esempio 8.5. Sia $\mathbf{A}$ una matrice $m \times n$, $\mathbf{A} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$. Definiamo

$$\mathrm{C}(\mathbf{A}) = \langle \mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n \rangle$$

che chiameremo *spazio delle colonne* di $\mathbf{A}$. Esso è un sottospazio di $\mathbb{C}^m$ e quindi $\dim \mathrm{C}(\mathbf{A}) \leq m$. Inoltre $\mathrm{C}(\mathbf{A})$ ha un insieme di generatori con n elementi, quindi $\dim \mathrm{C}(\mathbf{A}) \leq n$.

Nel caso $\mathbf{A}$ sia in forma ridotta, sappiamo calcolare la dimensione di $\mathrm{C}(\mathbf{A})$: è precisamente il numero di colonne dominanti. Per il caso generale avremo bisogno di un concetto più astratto, che è quello di applicazione lineare.

9. Applicazioni lineari

Una tecnica fondamentale in algebra è di confrontare fra loro due strutture diverse: nel nostro caso trattiamo di spazi vettoriali. Per il confronto si impiegano applicazioni (che è quasi sinonimo di funzione).

Definizione. Siano V e W spazi vettoriali e sia $f\colon V \to W$ un'applicazione di dominio V e codominio W. Diremo che f è un'*applicazione lineare* se

(L_1) per ogni $\mathbf{v}_1, \mathbf{v}_2 \in V$,

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2);$$

(L_2) per ogni $\mathbf{v} \in V$ e ogni scalare α,

$$f(\alpha\mathbf{v}) = \alpha f(\mathbf{v}).$$

Notiamo che nelle formule della definizione si usano, a sinistra del segno di uguaglianza, le operazioni in V e, a destra, le operazioni di W.

Esempi 9.1. (1) Se λ è uno scalare, l'applicazione $V \to V$ definita da $\mathbf{v} \mapsto \lambda\mathbf{v}$ è lineare.

(2) Se $f\colon V \to W$ è un'applicazione lineare e U è un sottospazio di V, allora $f|U\colon U \to W$ definita tramite $\mathbf{u} \mapsto f(\mathbf{u})$ è lineare. Si chiama *restrizione* di f al sottospazio U.

(3) Scegliamo $1 \leq i_1 < i_2 < \cdots < i_d \leq n$ e per un vettore $\mathbf{v} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_n]^T \in \mathbb{C}^n$ poniamo

$$f(\mathbf{v}) = \begin{bmatrix} \alpha_{i_1} \\ \alpha_{i_2} \\ \vdots \\ \alpha_{i_d} \end{bmatrix}.$$

L'applicazione $f\colon \mathbb{C}^n \to \mathbb{C}^d$ così ottenuta è lineare. Possiamo chiamarla *scelta delle componenti* $i_1, i_2, \dots, i_d$.

Elenchiamo alcune semplici conseguenze della definizione.

PROPOSIZIONE 9.2. *Se* $f\colon V \to W$ *è un'applicazione lineare, allora* $f(\mathbf{0}) = \mathbf{0}$.

DIMOSTRAZIONE. Poniamo $\mathbf{w} = f(\mathbf{0})$. Siccome $\mathbf{0} = \mathbf{0} + \mathbf{0}$ (stiamo usando il vettore nullo di V), abbiamo per la (L1)

$$\mathbf{w} = f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}) + f(\mathbf{0}) = \mathbf{w} + \mathbf{w}$$

da cui $\mathbf{w} = \mathbf{0}$ (il vettore nullo di W). □

Come per i sottospazi, esiste una verifica "breve" che un'applicazione è lineare.

PROPOSIZIONE 9.3. *Siano* V *e* W *spazi vettoriali e sia* $f\colon V \to W$ *un'applicazione di dominio* V *e codominio* W*. Allora* f *è lineare se e solo se, per ogni* $\mathbf{v}_1, \mathbf{v}_2 \in V$ *e per gli scalari* α_1, α_2,

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2).$$

DIMOSTRAZIONE. Se f è lineare, basta applicare le condizioni (L_1) e (L_2). Supponiamo che per f valga la condizione data. Se prendiamo $\alpha_1 = \alpha_2 = 1$, abbiamo

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(1\mathbf{v}_1 + 1\mathbf{v}_2) = 1f(\mathbf{v}_1) + 1f(\mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$$

che è la (L_1). Se prendiamo $\mathbf{v}_1 = \mathbf{v}$, $\mathbf{v}_2 = \mathbf{0}$, $\alpha_1 = \alpha$ e $\alpha_2 = 0$, abbiamo

$$f(\alpha\mathbf{v}) = f(\alpha\mathbf{v} + 0\mathbf{0}) = \alpha f(\mathbf{v}) + 0f(\mathbf{0}) = \alpha f(\mathbf{v})$$

che è la (L_2). □

Con una semplice induzione (esercizio) si dimostra il corollario seguente che useremo spesso in futuro.

COROLLARIO 9.4. *Sia* $f\colon V \to W$ *un'applicazione lineare tra spazi vettoriali; siano* $\mathbf{v}_i \in V$ *e* α_i *scalari* $(i = 1, \dots, n)$*. Allora*

$$f(\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \cdots + \alpha_n f(\mathbf{v}_n).$$

ESEMPIO 9.5. Sia $\mathbf{A}$ una matrice $m \times n$. Dato $\mathbf{v} \in \mathbb{C}^n$ possiamo considerare $\mathbf{A}\mathbf{v} \in \mathbb{C}^m$; quindi possiamo definire un'applicazione

$$\begin{aligned} f_{\mathbf{A}}\colon \mathbb{C}^n &\to \mathbb{C}^m \\ \mathbf{v} &\mapsto \mathbf{A}\mathbf{v} \end{aligned}$$

che è lineare. Questo segue facilmente dalle proprietà del prodotto di matrici. L'applicazione $f_{\mathbf{A}}$ si dice *indotta da* $\mathbf{A}$.

Questo esempio è fondamentale. Vedremo più avanti che, in un certo senso, ogni applicazione lineare tra spazi vettoriali di dimensione finita è di questo tipo. Renderemo precisa questa affermazione quando parleremo di coordinate.

ESEMPI 9.6. Possiamo naturalmente considerare applicazioni lineari anche fra spazi vettoriali reali, con le analoghe definizioni. (1) Se $\mathbf{A}$ è la matrice

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

allora $f_{\mathbf{A}}\colon \mathbb{R}^2 \to \mathbb{R}^2$ è l'applicazione lineare

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(2) Se $\mathbf{B}$ è la matrice

$$\mathbf{B} = \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

allora

$$f_{\mathbf{B}}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1\cos\alpha - x_2\operatorname{sen}\alpha \\ x_1\operatorname{sen}\alpha + x_2\cos\alpha \end{bmatrix}.$$

Talvolta è utile considerare più applicazioni lineari; la composizione di due, qualora sia definita, è ancora lineare.

PROPOSIZIONE 9.7. *Siano U, V e W spazi vettoriali; siano $f\colon U \to V$ e $g\colon V \to W$ applicazioni lineari. Allora $g \circ f\colon U \to W$ è un'applicazione lineare.*

DIMOSTRAZIONE. Basta eseguire le verifiche. Abbiamo

$$\begin{aligned} g \circ f(\mathbf{u}_1 + \mathbf{u}_2) &= g(f(\mathbf{u}_1 + \mathbf{u}_2)) = g(f(\mathbf{u}_1) + f(\mathbf{u}_2)) \\ &= g(f(\mathbf{u}_1)) + g(f(\mathbf{u}_2)) = g \circ f(\mathbf{u}_1) + g \circ f(\mathbf{u}_2) \end{aligned}$$

e questo prova la (L_1). Analogamente

$$g \circ f(\alpha\mathbf{u}) = g(f(\alpha\mathbf{u})) = g(\alpha f(\mathbf{u})) = \alpha(g(f(\mathbf{u}))) = \alpha(g \circ f(\mathbf{u}))$$

prova la (L_2). □

ESEMPIO 9.8. Per essere in questa situazione quando si hanno applicazioni indotte da matrici, consideriamo una matrice $\mathbf{A}$ che sia $m \times n$ e una matrice $\mathbf{B}$ che sia $n \times p$. Le applicazioni indotte sono

$$f_{\mathbf{A}}\colon \mathbb{C}^n \to \mathbb{C}^m, \qquad f_{\mathbf{B}}\colon \mathbb{C}^p \to \mathbb{C}^n.$$

La composizione $f_{\mathbf{A}} \circ f_{\mathbf{B}}$ ha senso ed è lineare. Questo si vede anche direttamente perché, per $\mathbf{v} \in \mathbb{C}^p$,

$$f_{\mathbf{A}} \circ f_{\mathbf{B}}(\mathbf{v}) = f_{\mathbf{A}}(f_{\mathbf{B}}(\mathbf{v})) = f_{\mathbf{A}}(\mathbf{B}\mathbf{v}) = \mathbf{A}(\mathbf{B}\mathbf{v}) = (\mathbf{A}\mathbf{B})\mathbf{v} = f_{\mathbf{A}\mathbf{B}}(\mathbf{v}).$$

Dunque, nel caso di applicazioni lineari indotte da matrici, la composizione corrisponde al prodotto di matrici. Questo è la motivazione profonda della definizione di prodotto di matrici che abbiamo usato.

Ogni applicazione lineare definisce un sottospazio del dominio e un sottospazio del codominio.

DEFINIZIONE. Se $f\colon V \to W$ è un'applicazione lineare tra spazi vettoriali, allora

$$\mathrm{N}(f) = \{\, \mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0} \,\}$$

è un sottospazio di V, detto *spazio nullo di f*. Analogamente

$$\mathrm{Im}(f) = \{\, f(\mathbf{v}) : \mathbf{v} \in V \,\}$$

è un sottospazio di W, detto *immagine di f*.

ESEMPI 9.9. (1) Consideriamo l'applicazione lineare definita nell'esempio 9.1 (3) di scelta delle componenti. Per fissare le idee, consideriamo la scelta della seconda, terza e quinta componente in $\mathbb{C}^5$, quindi l'applicazione lineare $f\colon \mathbb{C}^5 \to \mathbb{C}^3$ definita da

$$f\left(\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_5\end{bmatrix}\right) = \begin{bmatrix}\alpha_2\\ \alpha_3\\ \alpha_5\end{bmatrix}.$$

Lo spazio nullo $\mathrm{N}(f)$ è:

$$\mathrm{N}(f) = \{\, [\alpha_1\ 0\ 0\ \alpha_4\ 0]^T \mid \alpha_1, \alpha_4 \in \mathbb{C} \,\} = \langle \mathbf{e}_1; \mathbf{e}_4 \rangle.$$

L'immagine, come si verifica facilmente, è: $\mathrm{Im}(f) = \mathbb{C}^3$.

(2) Sia $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare definita da:

$$f\left(\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}\right) = \begin{bmatrix}x_1 + 2x_2 - x_3\\ -2x_1 - 4x_2 + x_3\end{bmatrix}.$$

Poniamo $\mathbf{w} = [1\ {-2}]^T$: allora ogni elemento di $\mathrm{Im}(f)$ è un multiplo scalare di $\mathbf{w}$, perché

$$\begin{bmatrix}x_1 + 2x_2 - x_3\\ -2x_1 - 4x_2 + x_3\end{bmatrix} = (x_1 + 2x_2 - x_3)\begin{bmatrix}1\\ -2\end{bmatrix}.$$

Inoltre $\mathbf{w} = f(\mathbf{e}_1)$, quindi $\operatorname{Im}(f) = \langle \mathbf{w} \rangle$.

Un vettore $\mathbf{v} = [\alpha_1\ \alpha_2\ \alpha_3]^T$ appartiene a $\mathrm{N}(f)$ se e solo se $x_1 + 2x_2 - x_3 = 0$.

Ovviamente va dimostrato che $\mathrm{N}(f)$ è un sottospazio di V e che $\operatorname{Im}(f)$ è un sottospazio di W.

Siano $\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{N}(f)$ e siano α_1 e α_2 scalari. Allora

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) = \alpha_1\mathbf{0} + \alpha_2\mathbf{0} = \mathbf{0},$$

dunque $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \in \mathrm{N}(f)$. Abbiamo già mostrato che $\mathbf{0} \in \mathrm{N}(f)$, infatti $f(\mathbf{0}) = \mathbf{0}$.

È poi ovvio che $\operatorname{Im}(f)$ non sia vuota; siano $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{Im}(f)$ e siano α_1 e α_2 scalari. Allora, per definizione, possiamo scrivere $\mathbf{w}_1 = f(\mathbf{v}_1)$ e $\mathbf{w}_2 = f(\mathbf{v}_2)$ per opportuni $\mathbf{v}_1, \mathbf{v}_2 \in V$. Dunque

$$\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) = f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) \in \operatorname{Im}(f).$$

Notiamo che, per dimostrare che $\mathbf{w} \in \operatorname{Im}(f)$, basta riuscire a scrivere $\mathbf{w} = f(\mathbf{v})$ per un $\mathbf{v} \in V$.

È particolarmente importante il caso in cui $\mathrm{N}(f) = \{\mathbf{0}\}$.

PROPOSIZIONE 9.10. *Un'applicazione lineare $f\colon V \to W$ è iniettiva se e solo se* $\mathrm{N}(f) = \{\mathbf{0}\}$.

DIMOSTRAZIONE. Supponiamo che f sia iniettiva e sia $\mathbf{v} \in \mathrm{N}(f)$. Allora $f(\mathbf{v}) = \mathbf{0} = f(\mathbf{0})$ e quindi $\mathbf{v} = \mathbf{0}$, cioè l'unico elemento di $\mathrm{N}(f)$ è il vettore nullo.

Viceversa, supponiamo che $\mathrm{N}(f) = \{\mathbf{0}\}$ e supponiamo che $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Allora

$$\mathbf{0} = f(\mathbf{v}_1) - f(\mathbf{v}_2) = f(\mathbf{v}_1 - \mathbf{v}_2)$$

cioè $\mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{N}(f)$. Per ipotesi, allora, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, quindi $\mathbf{v}_1 = \mathbf{v}_2$ e f è iniettiva. □

Data un'applicazione lineare $f\colon V \to W$ e un insieme finito $\mathcal{A}$ di vettori in V possiamo considerare l'insieme $f[\mathcal{A}]$ ottenuto valutando f nei vettori di $\mathcal{A}$. Più esplicitamente, se $\mathcal{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\}$, poniamo

$$f[\mathcal{A}] = \{f(\mathbf{v}_1); f(\mathbf{v}_2); \dots; f(\mathbf{v}_n)\}.$$

Per convenzione associamo all'insieme vuoto l'insieme vuoto.

PROPOSIZIONE 9.11. *Sia $f\colon V \to W$ lineare e sia $\mathcal{A}$ un insieme finito di vettori di V.*

(1) Se $f[\mathcal{A}]$ è linearmente indipendente, allora $\mathcal{A}$ è linearmente indipendente.

(2) Se f è iniettiva e $\mathcal{A}$ è linearmente indipendente, allora $f[\mathcal{A}]$ è linearmente indipendente.

DIMOSTRAZIONE. Entrambi i casi sono ovvi se $\mathcal{A}$ è vuoto. Supponiamo allora $\mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$.

(1) Supponiamo che $f[\mathcal{A}]$ sia linearmente indipendente e sia

$$\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0};$$

abbiamo, in W,

$$\mathbf{0} = f(\mathbf{0}) = f\biggl(\sum_{i=1}^{n} \alpha_i\mathbf{v}_i\biggr) = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i)$$

e, per l'ipotesi che $f[\mathcal{A}]$ è linearmente indipendente, $\alpha_1 = \dots = \alpha_n = 0$.

(2) Supponiamo di avere

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i).$$

Allora

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = f\biggl(\sum_{i=1}^{n} \alpha_i\mathbf{v}_i\biggr)$$

e quindi $\sum_{i=1}^{n} \alpha_i\mathbf{v}_i \in \mathrm{N}(f)$. Per ipotesi allora $\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}$ e quindi $\alpha_1 = \dots = \alpha_n = 0$. □

Ci riferiremo al teorema seguente con il nome di *nullità + rango.* Se $f\colon V \to W$ è lineare e V ha dimensione finita, diremo *nullità di* f la dimensione di $\mathrm{N}(f)$ e *rango di* f la dimensione di $\mathrm{Im}(f)$.

TEOREMA 9.12. *Sia* $f\colon V \to W$ *un'applicazione lineare tra spazi vettoriali e supponiamo che* V *sia finitamente generato. Allora* $\mathrm{Im}(f)$ *è finitamente generato e*

$$\dim V = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f).$$

DIMOSTRAZIONE. Sia $\mathcal{A} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ un insieme di generatori di V (nel caso in cui $\mathcal{A}$ sia l'insieme vuoto il teorema è di facile verifica). Poniamo $\mathbf{w}_i = f(\mathbf{v}_i)$ $(i = 1, \dots, n)$. Se $\mathbf{w} \in \mathrm{Im}(f)$, esiste $\mathbf{v} \in V$ tale che $\mathbf{w} = f(\mathbf{v})$. Ora $\mathbf{v}$ si può scrivere come combinazione lineare dei vettori di $\mathcal{A}$:

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.$$

Per il corollario 9.4, abbiamo:

$$\mathbf{w} = f(\mathbf{v}) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = \sum_{i=1}^{n} \alpha_i \mathbf{w}_i$$

e dunque $f[\mathcal{A}] = \{\mathbf{w}_1; \dots; \mathbf{w}_n\}$ è un insieme di generatori di $\mathrm{Im}(f)$ che quindi è finitamente generato.

Completeremo la dimostrazione facendo vedere che, se $\dim \mathrm{N}(f) = n$ e $\dim \mathrm{Im}(f) = k$, allora V ha una base con $n+k$ elementi. Le notazioni del seguito non hanno nessuna relazione con le precedenti.

Se $\dim \mathrm{Im}(f) = 0$, allora $\mathrm{Im}(f) = \{\mathbf{0}\}$ e quindi $\mathrm{N}(f) = V$. La relazione da dimostrare è, in questo caso, valida.

Sia $\{\mathbf{w}_1; \dots; \mathbf{w}_k\}$ una base di $\mathrm{Im}(f)$. Possiamo allora scrivere $\mathbf{w}_i = f(\mathbf{v}_i)$ $(i = 1, \dots, k)$ per opportuni vettori di V e poniamo $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_k\}$. Fissiamo anche una base $\mathcal{A}$ di $\mathrm{N}(f)$.

Sistemiamo il caso in cui $\mathcal{A}$ è vuota, cioè $\mathrm{N}(f) = \{\mathbf{0}\}$, quindi f è iniettiva. Per 9.10, $\dim V \le k$ perché insiemi linearmente indipendenti vanno in insiemi linearmente indipendenti. Ora l'insieme $\mathcal{B}$ è linearmente indipendente, per 9.4, perciò $k \le \dim V$. In definitiva $\dim V = k$.

Possiamo perciò supporre che $\mathcal{A}$ non sia vuoto: $\mathcal{A} = \{\mathbf{u}_1; \dots; \mathbf{u}_n\}$. Dimostreremo che $\mathcal{C} = \mathcal{A} \sqcup \mathcal{B}$ è una base di V.

L'insieme $\mathcal{C}$ *è un insieme di generatori di* V. Sia $\mathbf{v} \in V$. Siccome $f(\mathbf{v}) \in \mathrm{Im}(f)$, abbiamo

$$f(\mathbf{v}) = \sum_{j=1}^{k} \beta_j \mathbf{w}_j.$$

Consideriamo il vettore

$$\mathbf{v}' = \sum_{j=1}^{k} \beta_j \mathbf{v}_j.$$

In generale non potremo aspettarci che $\mathbf{v}' = \mathbf{v}$. Tuttavia

$$f(\mathbf{v} - \mathbf{v}') = f(\mathbf{v}) - f(\mathbf{v}') = \sum_{j=1}^{k} \beta_j \mathbf{w}_j - f\left(\sum_{j=1}^{k} \beta_j \mathbf{v}_j\right) = \sum_{j=1}^{k} \beta_j \mathbf{w}_j - \sum_{j=1}^{k} \beta_j f(\mathbf{v}_j) = \mathbf{0}.$$

Dunque $\mathbf{v} - \mathbf{v}' \in \mathrm{N}(f)$ e possiamo scrivere

$$\mathbf{v} - \mathbf{v}' = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i,$$

da cui

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \mathbf{v}' = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j$$

e quindi $\mathbf{v} \in \langle \mathcal{C} \rangle$.

L'insieme $\mathcal{C}$ è linearmente indipendente. Supponiamo di avere

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j.$$

Allora

$$\mathbf{0} = f(\mathbf{0}) = f\Bigl(\sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j\Bigr) = \sum_{i=1}^{n} \alpha_i f(\mathbf{u}_i) + \sum_{j=1}^{k} \beta_j f(\mathbf{v}_j) = \sum_{j=1}^{k} \beta_j \mathbf{w}_j$$

in quanto $\mathbf{u}_i \in \mathrm{N}(f)$ $(i = 1, \dots, n)$. Per l'indipendenza lineare di $f[\mathcal{B}]$ possiamo concludere che $\beta_1 = \dots = \beta_k = 0$. Ma allora

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$$

e, per l'indipendenza lineare di $\mathcal{A}$, abbiamo che $\alpha_1 = \dots = \alpha_n = 0$. □

Diamo subito una conseguenza rilevante di questo teorema. Ricordiamo che un'applicazione $f\colon V \to W$ è *suriettiva* se $\mathrm{Im}(f) = W$.

Corollario 9.13. *Se V è uno spazio vettoriale finitamente generato e $f\colon V \to V$ è lineare, allora f è iniettiva se e solo se f è suriettiva.*

Dimostrazione. Dire che f è iniettiva equivale a dire che $\dim \mathrm{N}(f) = 0$. Il teorema nullità+rango dice allora che $\dim V = 0+\dim \mathrm{Im}(f)$, cioè che $\dim \mathrm{Im}(f) = \dim V$. L'osservazione dopo 8.4 dice che $\mathrm{Im}(f) = V$.

Viceversa, se f è suriettiva abbiamo $\mathrm{Im}(f) = V$ e quindi $\dim \mathrm{Im}(f) = \dim V$. Per il teorema nullità + rango, $\dim V = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f)$ e quindi $\dim \mathrm{N}(f) = 0$. □

Attenzione: il corollario vale solo quando dominio e codominio di f *coincidono* o, più in generale, quando hanno la stessa dimensione (esercizio). Non vale se dominio e codominio hanno dimensioni diverse perché in tal caso nessuna applicazione lineare iniettiva può essere suriettiva e nessuna applicazione lineare suriettiva può essere iniettiva.

Non vale nemmeno se V non è finitamente generato. Per esempio, l'applicazione $f\colon \mathcal{P}(\mathbb{C}) \to \mathcal{P}(\mathbb{C})$ definita da

$$f(p) = pq,$$

dove q è un fissato polinomio di grado > 0, è iniettiva ma non suriettiva. Ne lasciamo la verifica per esercizio: il polinomio costante 1 non può appartenere a $\mathrm{Im}(f)$.

Come già accennato, il caso di applicazioni lineari indotte da matrici è molto importante. Calcoliamo anzitutto $\mathrm{Im}(f_{\mathbf{A}})$, dove $\mathbf{A}$ è una matrice $m \times n$. Se scriviamo $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \dots\ \mathbf{a}_n]$ e prendiamo $\mathbf{v} = [\alpha_1\ \alpha_2\ \dots\ \alpha_n]^T \in \mathbb{C}^n$, abbiamo

$$f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v} = \alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n$$

e dunque $f_{\mathbf{A}}(\mathbf{v}) \in \mathrm{C}(\mathbf{A})$. Ricordiamo che lo spazio delle colonne $\mathrm{C}(\mathbf{A})$ della matrice $\mathbf{A}$ è il sottospazio di $\mathbb{C}^m$ generato dall'insieme delle colonne di $\mathbf{A}$. Viceversa, se un vettore appartiene a $\mathrm{C}(\mathbf{A})$, esso si può scrivere come combinazione lineare dell'insieme delle colonne di $\mathbf{A}$, $\mathbf{w} = \alpha_1\mathbf{a}_1+\dots+\alpha_n\mathbf{a}_n$, ed è evidente che $\mathbf{w} = \mathbf{A}\mathbf{v} = f_{\mathbf{A}}(\mathbf{v})$, dove $\mathbf{v} = [\alpha_1\ \alpha_2\ \dots\ \alpha_n]^T \in \mathbb{C}^n$. Quindi $\mathrm{C}(\mathbf{A}) \subseteq \mathrm{Im}(f_{\mathbf{A}})$.

In definitiva $\mathrm{C}(\mathbf{A})$ coincide con l'immagine di $f_{\mathbf{A}}$. Se ne può anche dare un'altra descrizione: $\mathrm{C}(\mathbf{A})$ è l'insieme dei vettori $\mathbf{b} \in \mathbb{C}^m$ per i quali il sistema lineare $\mathbf{A}\mathbf{x} = \mathbf{b}$ ha soluzione.

Lo spazio nullo di $f_{\mathbf{A}}$ è invece l'insieme dei vettori $\mathbf{v} \in \mathbb{C}^n$ tali che $\mathbf{A}\mathbf{v} = \mathbf{0}$. In questo caso si chiama anche *spazio nullo di* $\mathbf{A}$ e si indica con $\mathrm{N}(\mathbf{A})$.

Definizione. La dimensione dello spazio delle colonne di una matrice $\mathbf{A}$ si chiama *rango* di $\mathbf{A}$ e si denota con $\operatorname{rk} \mathbf{A}$. Vedremo in seguito che questa nozione di rango di una matrice coincide con quella data nel Capitolo 1.

Vediamo come si comporta il rango rispetto alla composizione di applicazioni. Siano $f\colon U \to V$ e $g\colon V \to W$ lineari. Possiamo applicare il teorema nullità + rango a $g \circ f$:

$$\dim U = \dim \mathrm{N}(g \circ f) + \dim \mathrm{Im}(g \circ f).$$

È facile vedere che $\mathrm{N}(f) \subseteq \mathrm{N}(g \circ f)$: infatti, se $f(\mathbf{u}) = \mathbf{0}$, è anche $g \circ f(\mathbf{u}) = g(\mathbf{0}) = \mathbf{0}$. Dunque $\dim \mathrm{N}(f) \leq \dim \mathrm{N}(g \circ f)$. Sappiamo anche che

$$\dim U = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f);$$

mettendo assieme con la disuguaglianza precedente, otteniamo che

$$\dim \mathrm{Im}(g \circ f) \leq \dim \mathrm{Im}(f).$$

È anche evidente che $\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g)$, quindi che

$$\dim \mathrm{Im}(g \circ f) \leq \dim \mathrm{Im}(g).$$

Nel caso in cui le applicazioni lineari siano indotte da matrici, sappiamo che $f_{\mathbf{B}} \circ f_{\mathbf{A}} = f_{\mathbf{BA}}$, quindi ne ricaviamo

$$\operatorname{rk} \mathbf{BA} \leq \operatorname{rk} \mathbf{A}, \qquad \operatorname{rk} \mathbf{BA} \leq \operatorname{rk} \mathbf{B}.$$

Un ultimo fatto importante sulle applicazioni lineari.

Teorema 9.14. *Se $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ è una base dello spazio vettoriale V e $\{\mathbf{w}_1; \dots; \mathbf{w}_n\}$ è un insieme di vettori nello spazio vettoriale W, allora esiste una e una sola applicazione lineare $f\colon V \to W$ tale che $f(\mathbf{v}_i) = \mathbf{w}_i$ $(i = 1, 2, \dots, n)$.*

Dimostrazione. Diamo solamente un'idea, tralasciando i dettagli. Se un'applicazione lineare f del tipo voluto esiste, essa è unica. Infatti, supponiamo che anche $g\colon V \to W$ abbia la proprietà richiesta, cioè che $g(\mathbf{v}_i) = \mathbf{w}_i$ $(i = 1, 2, \dots, n)$. Se $\mathbf{v} \in V$, possiamo scrivere $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$ e dunque

$$g(\mathbf{v}) = \sum_{i=1}^n \alpha_i g(\mathbf{v}_i) = \sum_{i=1}^n \alpha_i \mathbf{w}_i = \sum_{i=1}^n \alpha_i f(\mathbf{v}_i) = f(\mathbf{v}).$$

Come si definisce allora f? Proprio scrivendo

$$f(\mathbf{v}) = \sum_{i=1}^n \alpha_i \mathbf{w}_i$$

e l'aspetto complicato della verifica è appunto di controllare che questa posizione definisca un'applicazione lineare. Si tratta di un noioso controllo dei dettagli e lo omettiamo. □

10. Inverse e rango

Abbiamo legato lo spazio delle colonne di una matrice $\mathbf{A}$ alla soluzione di sistemi lineari. Vediamo allora come caratterizzare, attraverso il rango, che è la dimensione dello spazio delle colonne, l'esistenza di inverse di $\mathbf{A}$; riotterremo risultati già visti nel Capitolo 1 in altro modo. Diamo qui per buono quanto vedremo più avanti, cioè che le definizioni di rango di una matrice date nel Capitolo 1 e qui sopra coincidono.

L'esistenza di un'inversa destra della matrice $\mathbf{A}$, di forma $m \times n$, equivale alla risolubilità dei sistemi lineari $\mathbf{Ax} = \mathbf{e}_i$ $(i = 1, 2, \dots, m)$. Questo a sua volta significa che i vettori $\mathbf{e}_i$ appartengono allo spazio delle colonne di $\mathbf{A}$, quindi che $\mathrm{C}(\mathbf{A})$ contiene una base di $\mathbb{C}^m$: ciò accade se e solo se $\mathrm{C}(\mathbf{A}) = \mathbb{C}^m$. Questa uguaglianza si può anche scrivere $\operatorname{rk} \mathbf{A} = m$.

Abbiamo però visto che $\operatorname{rk} \mathbf{A} \leq m$ e $\operatorname{rk} \mathbf{A} \leq n$. Se $m > n$, l'uguaglianza $\operatorname{rk} \mathbf{A} = m$ non può verificarsi. Perciò abbiamo dimostrato i risultati che seguono; per il secondo basta ricordare che $\mathbf{A}$ ha un'inversa sinistra se e solo se $\mathbf{A}^H$ ha un'inversa destra.

Proposizione 10.1. *Sia $\mathbf{A}$ una matrice $m \times n$.*

(1) $\mathbf{A}$ ha un'inversa destra se e solo se $\operatorname{rk} \mathbf{A} = m \leq n$.
(2) $\mathbf{A}$ ha un'inversa sinistra se e solo se $\operatorname{rk} \mathbf{A} = n \leq m$.

Da questi due fatti discende che una matrice non quadrata non può avere inversa destra *e* inversa sinistra e giustifica la definizione di matrice invertibile per le sole matrici quadrate, come già abbiamo osservato nel Capitolo 1.

COROLLARIO 10.2. *Una matrice quadrata* $\mathbf{A}$ *di ordine* n *è invertibile se e solo se* $\operatorname{rk}\mathbf{A} = n$.

11. Spazio delle colonne e spazio delle righe

Rimane il problema di calcolare il rango di una matrice; già nel Capitolo 1 questo era stato definito come il numero di colonne dominanti (in una forma ridotta); quello che mancava allora era la dimostrazione che questo numero è lo stesso, qualunque sia il procedimento di eliminazione seguito. Sappiamo comunque che per ogni matrice $\mathbf{A}$ esistono una matrice in forma ridotta $\mathbf{U}$ e una matrice invertibile $\mathbf{F}$ tali che $\mathbf{FA} = \mathbf{U}$.

Affronteremo il problema più in generale, considerando due matrici $\mathbf{A}$ e $\mathbf{B}$ di forma $m \times n$ tali che esista $\mathbf{F}$ invertibile di ordine m tale che $\mathbf{FA} = \mathbf{B}$. Fino al teorema 11.2 ci metteremo in questa ipotesi.

Scriviamo $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n]$ e $\mathbf{B} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_n]$. Gli insiemi di vettori $\mathcal{A} = \{\mathbf{a}_1; \dots; \mathbf{a}_n\}$ e $\mathcal{B} = \{\mathbf{b}_1; \dots; \mathbf{b}_n\}$ sono insiemi di generatori di $\mathrm{C}(\mathbf{A})$ e $\mathrm{C}(\mathbf{B})$ rispettivamente. Da essi perciò possiamo estrarre sottoinsiemi che sono basi. Dato un sottoinsieme di uno dei due è chiaro chi sia il sottoinsieme corrispondente nell'altro; per esempio, il sottoinsieme corrispondente a $\{\mathbf{a}_1; \mathbf{a}_3\}$ è $\{\mathbf{b}_1; \mathbf{b}_3\}$.

Ricordiamo infine che $\mathbf{b}_i = \mathbf{F}\mathbf{a}_i$.

PROPOSIZIONE 11.1. *Un vettore* $\mathbf{v} \in \mathbb{C}^m$ *è combinazione lineare dell'insieme delle colonne di* $\mathbf{A}$ *se e solo se* $\mathbf{Fv}$ *è combinazione lineare dell'insieme delle colonne di* $\mathbf{B}$.

DIMOSTRAZIONE. Se $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{a}_i$, allora $\mathbf{Fv} = \sum_{i=1}^n \alpha_i \mathbf{F}\mathbf{a}_i = \sum_{i=1}^n \alpha_i \mathbf{b}_i$.
Viceversa, se $\mathbf{Fv} = \sum_{i=1}^n \alpha_i \mathbf{b}_i$, abbiamo anche

$$\mathbf{v} = \mathbf{F}^{-1}(\mathbf{Fv}) = \mathbf{F}^{-1}\left(\sum_{i=1}^n \alpha_i \mathbf{b}_i\right) = \sum_{i=1}^n \alpha_i \mathbf{F}^{-1}\mathbf{b}_i = \sum_{i=1}^n \alpha_i \mathbf{a}_i$$

e la dimostrazione è completa. □

Ciò che abbiamo visto è che la moltiplicazione per $\mathbf{F}$ induce un'applicazione $f\colon \mathrm{C}(\mathbf{A}) \to \mathrm{C}(\mathbf{B})$ che è lineare proprio perché è la moltiplicazione tra matrici. Inoltre f è biiettiva: infatti, per simmetria, la moltiplicazione per $\mathbf{F}^{-1}$ induce un'applicazione lineare $g\colon \mathrm{C}(\mathbf{B}) \to \mathrm{C}(\mathbf{A})$. È immediato verificare che g è l'inversa di f.

Un'altra facile osservazione è che $\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{B})$. Infatti, se $\mathbf{Av} = \mathbf{0}$, si ha $\mathbf{Bv} = \mathbf{FAv} = \mathbf{F0} = \mathbf{0}$; analogamente il viceversa, usando che $\mathbf{A} = \mathbf{F}^{-1}\mathbf{B}$.

Da questo discende che $\operatorname{rk}\mathbf{A} = \operatorname{rk}\mathbf{B}$, usando il teorema nullità + rango. Infatti

$$\begin{aligned} n &= \dim \mathrm{N}(\mathbf{A}) + \operatorname{rk}\mathbf{A} \\ &= \dim \mathrm{N}(\mathbf{B}) + \operatorname{rk}\mathbf{B} \end{aligned}$$

Abbiamo anche un'altra conseguenza, che riguarda come trovare sottoinsiemi linearmente indipendenti dell'insieme delle colonne di $\mathbf{A}$ noti sottoinsiemi linearmente indipendenti dell'insieme delle colonne di $\mathbf{B}$.

TEOREMA 11.2. *Sia* $\mathbf{B} = \mathbf{FA}$*, con* $\mathbf{F}$ *matrice invertibile. Allora un sottoinsieme di colonne di* $\mathbf{A}$ *è linearmente indipendente se e solo se il corrispondente sottoinsieme di colonne di* $\mathbf{B}$ *è linearmente indipendente.*

Una sottoinsieme di colonne di $\mathbf{A}$ *è una base di* $\mathrm{C}(\mathbf{A})$ *se e solo se il sottoinsieme corrispondente di colonne di* $\mathbf{B}$ *è una base di* $\mathrm{C}(\mathbf{B})$.

DIMOSTRAZIONE. Se $\mathcal{A}$ è un sottoinsieme di colonne di $\mathbf{A}$, allora il sottoinsieme di colonne di $\mathbf{B}$ corrispondente è precisamente $f[\mathcal{A}]$, dove f è l'applicazione lineare $f\colon \mathrm{C}(\mathbf{A}) \to \mathrm{C}(\mathbf{B})$ indotta dalla moltiplicazione per $\mathbf{F}$. Siccome f è iniettiva, supponendo che $\mathcal{A}$ sia linearmente indipendente, abbiamo che $f[\mathcal{A}]$ è linearmente indipendente, per 9.11.

Il viceversa si ottiene considerando l'applicazione lineare $g = f^{-1}$.

Se $\mathcal{A}$ è una base di $\mathrm{C}(\mathbf{A})$, allora ha $k = \operatorname{rk}\mathbf{A}$ elementi. Siccome il sottoinsieme corrispondente è linearmente indipendente e ha $k = \operatorname{rk}\mathbf{B}$ elementi, esso è una base di $\mathrm{C}(\mathbf{B})$ per 8.4. $\square$

Da questa discussione ricaviamo dunque una procedura per determinare una base di $\mathrm{C}(\mathbf{A})$.

(1) Eseguiamo l'eliminazione di Gauss su $\mathbf{A}$, ottenendo una matrice in forma ridotta $\mathbf{U}$ e una matrice invertibile $\mathbf{F}$ tali che $\mathbf{U} = \mathbf{FA}$.

(2) Le colonne dominanti di $\mathbf{U}$ formano una base di $\mathrm{C}(\mathbf{U})$.

(3) Le colonne di $\mathbf{A}$ corrispondenti alle colonne dominanti di $\mathbf{U}$ formano una base di $\mathrm{C}(\mathbf{A})$.

È bene ricordare che, in generale, gli spazi delle colonne di $\mathbf{A}$ e di $\mathbf{U}$ sono *diversi*. Essi hanno però la stessa dimensione.

Un caso diverso è quando $\mathbf{B} = \mathbf{AG}$, con $\mathbf{G}$ matrice invertibile. In tal caso, infatti, $\mathrm{C}(\mathbf{B}) = \mathrm{C}(\mathbf{A})$. Sia $\mathbf{G} = [\gamma_{ij}]$ e scriviamo $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n]$, $\mathbf{B} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_n]$. Allora

$$\mathbf{b}_i = \sum_{j=1}^{n} \gamma_{ij}\mathbf{a}_j \qquad (i = 1, \dots, n)$$

da cui segue che ogni colonna di $\mathbf{B}$ è combinazione lineare dell'insieme delle colonne di $\mathbf{A}$, perciò che $\mathrm{C}(\mathbf{B}) \subseteq \mathrm{C}(\mathbf{A})$. L'inclusione inversa discende dall'uguaglianza $\mathbf{A} = \mathbf{BG}^{-1}$.

Esiste un altro importante sottospazio associato a una matrice $\mathbf{A}$ di forma $m \times n$. La sua importanza risulterà da quanto si vedrà nel Capitolo 3. La sua definizione può sembrare contorta, tuttavia è necessaria nell'ottica di considerare, come faremo noi, sempre vettori colonna.

Definizione. Lo *spazio delle righe* della matrice $\mathbf{A}$ è lo spazio delle colonne di $\mathbf{A}^H$ e si indica con $\mathrm{R}(\mathbf{A})$.

Supponiamo ancora una volta che $\mathbf{B} = \mathbf{FA}$, con $\mathbf{F}$ invertibile. In questo caso possiamo asserire che $\mathrm{R}(\mathbf{A}) = \mathrm{R}(\mathbf{B})$. Infatti abbiamo

$$\mathbf{B}^H = (\mathbf{FA})^H = \mathbf{A}^H\mathbf{F}^H$$

e che $\mathbf{F}^H$ è invertibile. Perciò, per quanto visto prima,

$$\mathrm{R}(\mathbf{B}) = \mathrm{C}(\mathbf{B}^H) = \mathrm{C}(\mathbf{A}^H) = \mathrm{R}(\mathbf{A}).$$

Dunque, se sappiamo trovare una base dello spazio delle righe della matrice in forma ridotta ottenuta da $\mathbf{A}$ con l'eliminazione, questa sarà una base di $\mathrm{R}(\mathbf{A})$.

Proposizione 11.3. *Sia $\mathbf{U}$ una matrice in forma ridotta. Una base di $\mathrm{R}(\mathbf{U})$ è data dall'insieme delle righe non nulle di $\mathbf{U}$.*

Dimostrazione. Quando parliamo di righe di $\mathbf{U}$, in realtà stiamo considerando le colonne di $\mathbf{U}^H$. Ci basta allora dimostrare che le righe non nulle di $\mathbf{U}$ sono linearmente indipendenti: da un insieme di generatori possiamo estrarre una base e, ovviamente, tutte le righe nulle possono essere eliminate.

La verifica è analoga a quanto fatto nell'esempio 6.10 e la lasciamo per esercizio nel caso generale, dando solo un esempio con la matrice

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

Scriviamo una combinazione lineare nulla delle colonne non nulle di $\mathbf{U}^H$:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \alpha_1 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} \alpha_3 \\ \alpha_2 + 2\alpha_3 \\ \alpha_2 \\ \alpha_1 - \alpha_3 \\ -\alpha_1 + 2\alpha_2 - \alpha_3 \end{bmatrix}.$$

Guardando solo le equazioni corrispondenti alle posizioni dei *pivot*, in questo caso la prima, la seconda e la quarta, abbiamo subito che $\alpha_1 = 0$, $\alpha_2 = 0$ e $\alpha_3 = 0$, come si voleva. □

Salta subito agli occhi un fatto notevole: lo spazio delle colonne di $\mathbf{U}$ e lo spazio delle righe di $\mathbf{U}$ hanno la stessa dimensione. Questo vale allora per ogni matrice.

TEOREMA 11.4. *Per ogni matrice* $\mathbf{A}$*, si ha* $\operatorname{rk}\mathbf{A} = \operatorname{rk}\mathbf{A}^H$.

DIMOSTRAZIONE. Sia $\mathbf{U}$ una forma ridotta di $\mathbf{A}$, $\mathbf{U} = \mathbf{FA}$ con $\mathbf{F}$ invertibile. Allora

$$\operatorname{rk}\mathbf{A} = \dim \mathrm{C}(\mathbf{A}) = \dim \mathrm{C}(\mathbf{U}) = \dim \mathrm{R}(\mathbf{U}) = \dim \mathrm{R}(\mathbf{A}) = \dim \mathrm{C}(\mathbf{A}^H) = \operatorname{rk}\mathbf{A}^H.$$ □

Anche un quarto sottospazio associato a una matrice è talvolta utile.

DEFINIZIONE. Lo *spazio nullo sinistro* della matrice $\mathbf{A}$ è lo spazio nullo della matrice $\mathbf{A}^H$.

Se $\mathbf{A}$ ha forma $m \times n$, lo spazio nullo sinistro è perciò dato dai vettori $\mathbf{w} \in \mathbb{C}^m$ tali che $\mathbf{A}^H\mathbf{w} = 0$ o, equivalentemente, $\mathbf{w}^H\mathbf{A} = 0$. Ciò ne giustifica il nome. Non useremo notazioni particolari per indicare lo spazio nullo sinistro. Ci interessa solo valutarne la dimensione.

PROPOSIZIONE 11.5. *La dimensione dello spazio nullo di una matrice* $\mathbf{A}$ *di forma* $m \times n$ *è* $n - \operatorname{rk}\mathbf{A}$*. La dimensione dello spazio nullo sinistro di* $\mathbf{A}$ *è* $m - \operatorname{rk}\mathbf{A}$.

DIMOSTRAZIONE. Si tratta di applicare il teorema nullità+rango. Per definizione, $\mathrm{N}(\mathbf{A}) = \mathrm{N}(f_{\mathbf{A}})$ dove $f_{\mathbf{A}} \colon \mathbb{C}^n \to \mathbb{C}^m$. Allora

$$\dim \mathrm{N}(\mathbf{A}) = n - \dim \mathrm{C}(\mathbf{A}) = n - \operatorname{rk}\mathbf{A}.$$

Analogamente, $\mathrm{N}(\mathbf{A}^H) = \mathrm{N}(f_{\mathbf{A}^H})$, dove $f_{\mathbf{A}^H} \colon \mathbb{C}^m \to \mathbb{C}^n$. Allora

$$\dim \mathrm{N}(\mathbf{A}^H) = m - \dim \mathrm{C}(\mathbf{A}^H) = m - \operatorname{rk}\mathbf{A}.$$ □

Rimane solo da descrivere una procedura per calcolare una base dello spazio nullo di una matrice $m \times n$. Sappiamo che lo spazio nullo di $\mathbf{A}$ coincide con quello di una sua forma ridotta con eliminazione in avanti e all'indietro $\mathbf{U}$. I vettori dello spazio nullo di $\mathbf{A}$ sono le soluzioni del sistema omogeneo $\mathbf{Ux} = \mathbf{0}$ e ci basterà trovare fra esse $n - \operatorname{rk}\mathbf{A}$ vettori linearmente indipendenti.

Una soluzione del sistema omogeneo si ottiene dando valori arbitrari alle incognite non dominanti, ottenendo in corrispondenza i valori delle incognite dominanti. Abbiamo un modo per scrivere $d = n - \operatorname{rk}\mathbf{A}$ vettori distinti: se le incognite non dominanti sono x_{i_1}, x_{i_2}, …, x_{i_d}, possiamo dare a una di esse il valore 1 e alle altre il valore 0. Questa assegnazione si può fare in d modi diversi, ottenendo i vettori $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_d$.

Questo insieme di vettori è linearmente indipendente: infatti a essi corrispondono, nell'applicazione che sceglie le componenti i_1, i_2, …, i_d (si veda l'esempio 9.1 (3)), i vettori della base canonica di $\mathbb{C}^d$, che formano un insieme linearmente indipendente.

12. Il teorema di Rouché-Capelli

I concetti di rango e di spazio delle colonne di una matrice permettono di dare una veste precisa alle questioni sulla risolubilità di un sistema lineare. Il teorema che enunceremo è in realtà già stato dimostrato nel Capitolo 1, in forma leggermente diversa. Invitiamo il lettore a indagare sulle somiglianze e sulle differenze.

Un sistema lineare di m equazioni in n incognite si può scrivere nella forma $\mathbf{Ax} = \mathbf{b}$ e sappiamo che esso è risolubile se e solo se $\mathbf{b} \in \mathrm{C}(\mathbf{A})$. La matrice $\mathbf{A}$ è la *matrice dei coefficienti* del sistema, la matrice $[\mathbf{A} \mid \mathbf{b}]$ è la *matrice aumentata*.

Il sistema è allora risolubile se e solo se $\mathbf{b}$ è combinazione lineare dell'insieme delle prime n colonne della matrice aumentata. Quanto abbiamo visto nel teorema 11.2 mostra allora quello che già sappiamo: supponendo che $[\mathbf{U} \mid \mathbf{c}]$ sia una forma ridotta della matrice aumentata del sistema, il sistema è risolubile se e solo se $\mathbf{c}$ è non dominante. Infatti ciò accade se e solo se l'ultima colonna di $[\mathbf{A} \mid \mathbf{b}]$ è combinazione lineare dell'insieme delle colonne precedenti.

Un altro modo di vedere la faccenda usa il concetto di rango; si tratta di un aspetto utile nella teoria, anche perché finora l'unico modo che abbiamo visto per calcolare il rango di una matrice è di eseguire l'eliminazione.

Osserviamo che, chiaramente, $\mathrm{C}(\mathbf{A}) \subseteq \mathrm{C}([\mathbf{A} \mid \mathbf{b}])$. Dunque $\operatorname{rk}\mathbf{A} \leq \operatorname{rk}[\mathbf{A} \mid \mathbf{b}]$. Di quanto può aumentare il rango della matrice aumentata? Poniamo $k = \operatorname{rk}\mathbf{A}$; ciò significa che fra le colonne di $\mathbf{A}$ possiamo sceglierne k a formare una base. Ma allora $\mathrm{C}([\mathbf{A} \mid \mathbf{b}])$ ha un insieme di generatori con $k+1$ elementi (la base di $\mathrm{C}(\mathbf{A})$ a cui aggiungiamo $\mathbf{b}$). Quindi $\dim \mathrm{C}([\mathbf{A} \mid \mathbf{b}]) \leq k+1$.

Il sistema omogeneo associato a quello dato, cioè $\mathbf{Ax} = \mathbf{0}$, fornisce un altro dato. Supponiamo che $\mathbf{v}_0$ sia una soluzione di $\mathbf{Ax} = \mathbf{b}$ e sia $\mathbf{u}$ una soluzione del sistema omogeneo associato, in altre parole un elemento di $\mathrm{N}(\mathbf{A})$. Abbiamo allora

$$\mathbf{A}(\mathbf{u} + \mathbf{v}_0) = \mathbf{Au} + \mathbf{Av}_0 = \mathbf{0} + \mathbf{b} = \mathbf{b}$$

e quindi $\mathbf{u} + \mathbf{v}_0$ è una soluzione del sistema dato. Viceversa, se $\mathbf{v}$ è una soluzione del sistema dato e poniamo $\mathbf{u} = \mathbf{v} - \mathbf{v}_0$, abbiamo

$$\mathbf{Au} = \mathbf{A}(\mathbf{v} - \mathbf{v}_0) = \mathbf{Av} - \mathbf{Av}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

cioè $\mathbf{u} \in \mathrm{N}(\mathbf{A})$ ed è quindi soluzione del sistema omogeneo.

Abbiamo quindi visto che la differenza fra due soluzioni del sistema dato è un elemento dello spazio nullo della matrice dei coefficienti. Quindi una volta nota *una soluzione particolare* del sistema, tutte le altre si ottengono sommando a essa un elemento dello spazio nullo. Perciò, per conoscere tutte le soluzioni, basta conoscerne una e avere a disposizione una base dello spazio nullo, la cui dimensione è precisamente $n - \operatorname{rk}\mathbf{A}$.

È a questo che ci si riferisce quando si dice che le soluzioni di un sistema lineare, se esistono, dipendono da un numero di parametri pari al numero delle incognite meno il numero di incognite dominanti: questi parametri sono i coefficienti che possiamo assegnare agli elementi di una base dello spazio nullo per costruire una loro combinazione lineare.

Teorema 12.1 (Rouché-Capelli). *Sia $\mathbf{Ax} = \mathbf{b}$ un sistema lineare di m equazioni in n incognite. Il sistema è risolubile se e solo se*

$$\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A} \mid \mathbf{b}]$$

e, in tal caso, le soluzioni del sistema dipendono da $n - \operatorname{rk}\mathbf{A}$ parametri. In particolare il sistema ammette una e una sola soluzione se e solo se

$$\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A} \mid \mathbf{b}] = n.$$

13. Coordinate

In tutta questa sezione supporremo che gli spazi vettoriali che menzioneremo siano non nulli (cioè di dimensione > 0), oltre che finitamente generati.

Nel contesto che tratteremo, a differenza di quanto fatto in precedenza, l'ordine in cui si considerano gli elementi di una base è importante. Perciò quando tratteremo di coordinate e, in seguito, di matrici associate ad applicazioni lineari, supporremo che una base sia definita sia dandone gli elementi che fissandone l'ordine in cui sono considerati. Non useremo notazioni diverse da prima, ma daremo risalto a questo fatto parlando di *basi ordinate*.

Se $\mathcal{B}$ è una base dello spazio vettoriale V, ogni elemento di V si può scrivere *in modo unico* come combinazione lineare dei vettori di $\mathcal{B}$.

Infatti, se $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ e

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$$

allora

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i - \sum_{i=1}^{n} \beta_i \mathbf{v}_i = \sum_{i=1}^{n} (\alpha_i - \beta_i) \mathbf{v}_i$$

e, per l'indipendenza lineare, $\alpha_i = \beta_i$ $(i = 1, \dots, n)$.

Se poi consideriamo $\mathcal{B}$ come base *ordinata*, possiamo definire un'applicazione $\mathrm{C}_{\mathcal{B}}\colon V \to \mathbb{C}^n$ ponendo

$$\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{se e solo se} \quad \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.$$

È proprio il fatto di aver fissato l'ordine dei vettori di $\mathcal{B}$ a permetterci di dare questa definizione.

Questa applicazione si dice *applicazione delle coordinate rispetto alla base ordinata* $\mathcal{B}$. Diremo che $\mathrm{C}_{\mathcal{B}}(\mathbf{v})$ è il *vettore delle coordinate di* $\mathbf{v}$ *rispetto alla base ordinata* $\mathcal{B}$.

PROPOSIZIONE 13.1. *L'applicazione delle coordinate* $\mathrm{C}_{\mathcal{B}}\colon V \to \mathbb{C}^n$ *è lineare e biiettiva.*

DIMOSTRAZIONE. Sia $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$. Siano $\mathbf{u}, \mathbf{v} \in V$ e poniamo

$$\mathrm{C}_{\mathcal{B}}(\mathbf{u}) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \qquad \mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Questo significa che

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i, \qquad \mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i,$$

da cui

$$\mathbf{u} + \mathbf{v} = \sum_{i=1}^{n} (\alpha_i + \beta_i) \mathbf{v}_i$$

e quindi che

$$\mathrm{C}_{\mathcal{B}}(\mathbf{u} + \mathbf{u}) = \begin{bmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{bmatrix} = \mathrm{C}_{\mathcal{B}}(\mathbf{u}) + \mathrm{C}_{\mathcal{B}}(\mathbf{v}).$$

Lasciamo al lettore la verifica che $\mathrm{C}_{\mathcal{B}}(\alpha\mathbf{u}) = \alpha\,\mathrm{C}_{\mathcal{B}}(\mathbf{u})$ per ogni $\mathbf{u} \in V$ e ogni scalare α.

Inoltre $\mathrm{C}_{\mathcal{B}}$ è chiaramente suriettiva: per definizione

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \mathrm{C}_{\mathcal{B}}\biggl(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\biggr).$$

Per il teorema nullità + rango, la dimensione dello spazio nullo di $\mathrm{C}_{\mathcal{B}}$ è zero. □

L'applicazione delle coordinate rispetto a una base è lo strumento che rende possibile trattare ogni questione su spazi vettoriali attraverso matrici, proprio perché è lineare e biiettiva. Infatti un insieme di vettori $\mathcal{A}$ in V è linearmente indipendente se e solo se l'insieme $\mathrm{C}_{\mathcal{B}}[\mathcal{A}]$ è linearmente indipendente e quest'ultimo è un insieme di vettori in $\mathbb{C}^n$ e dunque si può trattare con i metodi matriciali.

Analogamente un vettore $\mathbf{v} \in V$ è combinazione lineare dei vettori di $\mathcal{A}$ se e solo se $\mathrm{C}_{\mathcal{B}}(\mathbf{v})$ è combinazione lineare dei vettori di $\mathrm{C}_{\mathcal{B}}[\mathcal{A}]$ (esercizio).

Questa situazione è analoga a quella discussa quando abbiamo costruito un'applicazione lineare biiettiva tra $\mathrm{C}(\mathbf{A})$ e $\mathrm{C}(\mathbf{B})$, dove $\mathbf{B} = \mathbf{FA}$ per una matrice $\mathbf{F}$ invertibile.

ESEMPI 13.2. (1) Nelle notazioni precedenti, abbiamo che

$$\mathrm{C}_{\mathcal{B}}(\mathbf{v_i}) = \mathbf{e}_i \qquad (i = 1, 2, \dots, n).$$

Infatti possiamo scrivere

$$\mathbf{v}_i = 0\mathbf{v}_1 + \dots + 1\mathbf{v}_i + \dots + 0\mathbf{v}_n$$

e dunque

$$\mathrm{C}_{\mathcal{B}}(\mathbf{v_i}) = \begin{array}{c} \\ [0\ \dots\ 0 \end{array} \begin{array}{c} i \\ \downarrow \\ 1 \end{array} \begin{array}{c} \\ 0\ \dots\ 0]^T = \mathbf{e}_i. \end{array}$$

Dunque il vettore delle coordinate di un vettore della base $\mathcal{B}$ è il corrispondente vettore della base canonica di $\mathbb{C}^n$.

(2) L'insieme $\mathcal{B} = \{2; 1-X; 1+X^2\}$ è una base di $\mathcal{P}_3(\mathbb{C})$ (lo si verifichi e si verifichi anche che $\dim \mathcal{P}_n(\mathbb{C}) = n$). Vogliamo calcolare $\mathrm{C}_{\mathcal{B}}(4-2X+3X^2)$. Per fare questo ci occorre scrivere

$$4 - 2X + 3X^2 = \alpha \cdot 2 + \beta(1-X) + \gamma(1+X^2)$$

che equivale al sistema lineare

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$$

e scrivendo la soluzione otteniamo

$$\mathrm{C}_{\mathcal{B}}(4 - 2X + 3X^2) = [-1/2\ 2\ 3]^T.$$

(3) Se lo spazio vettoriale V è $\mathbb{C}^n$, possiamo considerarne la base canonica $\mathcal{E}$. Secondo la definizione $\mathrm{C}_{\mathcal{E}}(\mathbf{v})$ è la colonna formata dai coefficienti che servono a scrivere $\mathbf{v}$ come combinazione lineare di $\mathcal{E}$. Ma

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \dots + \alpha_n \mathbf{e}_n$$

e quindi $\mathrm{C}_{\mathcal{E}}(\mathbf{v}) = \mathbf{v}$. In altre parole $\mathrm{C}_{\mathcal{E}}$ è l'applicazione *identità*.

È ovvio che se $\mathcal{B}$ e $\mathcal{D}$ sono basi ordinate distinte di V, le applicazioni delle coordinate saranno diverse (perché?). È tuttavia altrettanto chiaro che ci deve essere un modo per calcolare le coordinate di un vettore rispetto a una base conoscendo quelle rispetto a un'altra base.

Assumiamo che questa trasformazione delle coordinate si possa ottenere tramite la moltiplicazione per una matrice, cioè che esista una matrice quadrata $\mathbf{M}$ di ordine $n = \dim V$ tale che, per ogni $\mathbf{v} \in V$,

$$\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{M}\,\mathrm{C}_{\mathcal{D}}(\mathbf{v}).$$

Abbiamo dati sufficienti per determinare $\mathbf{M}$? Sappiamo calcolare $\mathrm{C}_{\mathcal{D}}(\mathbf{w}_i)$, dove $\mathcal{D} = \{\mathbf{w}_1; \dots; \mathbf{w}_n\}$, infatti

$$\mathrm{C}_{\mathcal{D}}(\mathbf{w}_i) = \mathbf{e}_i.$$

Ma allora, la nostra ipotesi impone che

$$\mathrm{C}_{\mathcal{B}}(\mathbf{w}_i) = \mathbf{M}\,\mathrm{C}_{\mathcal{D}}(\mathbf{w}_i) = \mathbf{M}\mathbf{e}_i = i\text{-esima colonna di } \mathbf{M}.$$

Dunque possiamo scrivere $\mathbf{M}$ nel modo seguente:

$$\mathbf{M} = \begin{bmatrix} \mathrm{C}_{\mathcal{B}}(\mathbf{w}_1) & \mathrm{C}_{\mathcal{B}}(\mathbf{w}_2) & \dots & \mathrm{C}_{\mathcal{B}}(\mathbf{w}_n) \end{bmatrix}$$

e ci resta solo da vedere che effettivamente la matrice $\mathbf{M}$ così definita è adeguata alla nostra ipotesi.

TEOREMA 13.3. *Siano $\mathcal{B}$ e $\mathcal{D}$ basi dello spazio vettoriale V. Allora esiste una e una sola matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$ tale che, per ogni $\mathbf{v} \in V$,*

$$\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\,\mathrm{C}_{\mathcal{D}}(\mathbf{v}).$$

La matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$ è invertibile e la sua inversa è $\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}}$.

DIMOSTRAZIONE. Nella discussione precedente abbiamo mostrato l'unicità della matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$: se una tale matrice esiste essa è proprio

$$\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}} = \big[\mathrm{C}_{\mathcal{B}}(\mathbf{w}_1)\ \ \mathrm{C}_{\mathcal{B}}(\mathbf{w}_2)\ \ \dots\ \ \mathrm{C}_{\mathcal{B}}(\mathbf{w}_n)\big].$$

Facciamo vedere che questa matrice funziona come voluto. Sia $\mathbf{v} \in V$ e poniamo

$$\mathrm{C}_{\mathcal{D}}(\mathbf{v}) = [\beta_1\ \beta_2\ \dots\ \beta_n]^T.$$

Allora $\mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{w}_i$; possiamo usare la linearità di $\mathrm{C}_{\mathcal{B}}$ e il fatto che, *per definizione*, $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{e}_i = \mathrm{C}_{\mathcal{B}}(\mathbf{w}_i)$:

$$\begin{aligned}
\mathrm{C}_{\mathcal{B}}(\mathbf{v}) &= \sum_{i=1}^{n} \beta_i\, \mathrm{C}_{\mathcal{B}}(\mathbf{w}_i) = \sum_{i=1}^{n} \beta_i \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{e}_i \\
&= \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\left(\sum_{i=1}^{n} \beta_i \mathbf{e}_i\right) = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{bmatrix} \\
&= \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\, \mathrm{C}_{\mathcal{D}}(\mathbf{v}).
\end{aligned}$$

Rimane da mostrare che $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$ è invertibile. Osserviamo alcune questioni importanti su quanto abbiamo visto.

(1) Ciò che abbiamo dimostrato vale per due basi qualunque. In particolare la matrice $\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}}$ esiste ed è unica.
(2) Per verificare che, data una certa matrice $\mathbf{M}$, si ha $\mathbf{M} = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$, basta dimostrare che, per ogni $\mathbf{v} \in V$, si ha $\mathbf{M}\,\mathrm{C}_{\mathcal{D}}(\mathbf{v}) = \mathrm{C}_{\mathcal{B}}(\mathbf{v})$.
(3) Non abbiamo in nessun punto usato che $\mathcal{B}$ e $\mathcal{D}$ siano diverse; quindi la matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}}$ esiste ed è unica.

Quale sarà la matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}}$? Siccome $\mathbf{I}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \mathrm{C}_{\mathcal{B}}(\mathbf{v})$ l'osservazione (2) dice che $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}} = \mathbf{I}$. Ora proviamo a eseguire il prodotto $(\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}})\mathbf{v}$:

$$(\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}})\mathbf{v} = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}(\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}}\mathbf{v}) = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\,\mathrm{C}_{\mathcal{D}}(\mathbf{v}) = \mathrm{C}_{\mathcal{B}}(\mathbf{v}).$$

Dunque la matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}}$ coincide con la matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}} = \mathbf{I}$. Perciò abbiamo dimostrato che $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}} = \mathbf{I}$. Analogamente $\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{B}}\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}} = \mathbf{I}$. □

La matrice $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$ si chiama *matrice del cambiamento di base*. Consigliamo al lettore di non cercare di imparare a memoria le formule, ma di ricordare il procedimento con cui si è costruita $\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}$: si vuole che $\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\,\mathrm{C}_{\mathcal{D}}(\mathbf{v})$ ed è facile calcolare le coordinate dei vettori della base $\mathcal{D}$ rispetto a $\mathcal{D}$, perché sono proprio i vettori della base canonica di $\mathbb{C}^n$. Inoltre $\mathbf{A}\mathbf{e}_i$ è la i-esima colonna della matrice $\mathbf{A}$.

Un caso particolare di queste matrici si ha quando $V = \mathbb{C}^n$ e una delle basi è la base canonica. Vogliamo dunque calcolare $\mathbf{M}_{\mathcal{E}\leftarrow\mathcal{B}}$, dove $\mathcal{B}$ è un'altra base di $\mathbb{C}^n$. Secondo la definizione, posto $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$, dobbiamo calcolare $\mathrm{C}_{\mathcal{E}}(\mathbf{v}_i)$ e questo è facile, per quanto visto nell'esempio 13.2 (3): $\mathrm{C}_{\mathcal{E}}(\mathbf{v}_i) = \mathbf{v}_i$. Dunque

$$\mathbf{M}_{\mathcal{E}\leftarrow\mathcal{B}} = [\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_n].$$

Per esercizio si verifichi che, data un'altra base $\mathcal{D}$ di $\mathbb{C}^n$, si ha

$$\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}} = \mathbf{M}_{\mathcal{E}\leftarrow\mathcal{B}}^{-1}\mathbf{M}_{\mathcal{E}\leftarrow\mathcal{D}}.$$

Più in generale, si dimostri che, se $\mathcal{B}$, $\mathcal{D}$ e $\mathcal{F}$ sono basi dello spazio vettoriale V, si ha

$$\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{F}} = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{D}}\mathbf{M}_{\mathcal{D}\leftarrow\mathcal{F}}.$$

14. Matrici associate alle applicazioni lineari

Come abbiamo visto nella sezione precedente, le coordinate permettono di trasferire questioni sugli spazi vettoriali a questioni su matrici. Il caso principale è quello delle applicazioni lineari. Con una tecnica analoga a quella per le matrici del cambiamento di base, assoceremo a un'applicazione lineare una matrice.

TEOREMA 14.1. *Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali; siano $\mathcal{B}$ una base di V e $\mathcal{D}$ una base di W. Se $n = \dim V$ e $m = \dim W$, esiste una e una sola matrice $\mathbf{A}$ di forma $m \times n$ tale che, per ogni $\mathbf{v} \in V$,*

$$\mathrm{C}_{\mathcal{D}}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}).$$

DIMOSTRAZIONE. Come nel caso delle matrici del cambiamento di base vediamo prima l'unicità di $\mathbf{A}$. Poniamo $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ e ricordiamo che $\mathrm{C}_{\mathcal{B}}(\mathbf{v}_i) = \mathbf{e}_i$. Dunque dobbiamo avere

$$\mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_i)) = \mathbf{A}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}_i) = \mathbf{A}\mathbf{e}_i = i\text{-esima colonna di } \mathbf{A}.$$

Perciò, se la matrice $\mathbf{A}$ esiste, deve essere

$$\mathbf{A} = \begin{bmatrix}\mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_1)) & \mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_2)) & \dots & \mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_n))\end{bmatrix}.$$

Facciamo vedere che questa matrice fa proprio al caso nostro. Sia $\mathbf{v} \in V$ e scriviamo $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$, cioè $\mathrm{C}_{\mathcal{B}}(\mathbf{v}) = [\alpha_1 \ \dots \ \alpha_n]^T$. Allora

$$\begin{aligned}
\mathrm{C}_{\mathcal{D}}(f(\mathbf{v})) &= \mathrm{C}_{\mathcal{D}}\biggl(f\biggl(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\biggr)\biggr) = \sum_{i=1}^{n} \alpha_i\,\mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_i)) \\
&= \sum_{i=1}^{n} \alpha_i \mathbf{A}\mathbf{e}_i = \mathbf{A}\biggl(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\biggr) \\
&= \mathbf{A}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}),
\end{aligned}$$

come si desiderava. □

La matrice $\mathbf{A}$ così trovata si chiama *matrice associata a f rispetto alla base $\mathcal{B}$ sul dominio e alla base $\mathcal{D}$ sul codominio.* L'unicità dice che, ogni volta che si trovi una matrice che svolge il compito della matrice associata, allora essa *è* la matrice associata.

ESEMPIO 14.2. Prendiamo un'applicazione lineare $f\colon \mathbb{C}^n \to \mathbb{C}^m$. Se su dominio e codominio prendiamo le basi canoniche $\mathcal{E}_n$ e $\mathcal{E}_m$, possiamo calcolare la matrice associata a f. Sarà la matrice $\mathbf{A}$ di forma $m \times n$ tale che, per ogni $\mathbf{v} \in \mathbb{C}^n$,

$$\mathrm{C}_{\mathcal{E}_n}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathcal{E}_p}(\mathbf{v}) \qquad \text{cioè} \qquad f(\mathbf{v})) = \mathbf{A}\mathbf{v}$$

per le proprietà della base canonica. Dunque ogni applicazione lineare $\mathbb{C}^n \to \mathbb{C}^m$ è la moltiplicazione per una matrice.

Vediamo un primo uso della matrice associata: il rango di un'applicazione lineare si può calcolare tramite il rango della sua matrice associata.

PROPOSIZIONE 14.3. *Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali; siano $\mathcal{B}$ una base di V e $\mathcal{D}$ una base di W. Se $\mathbf{A}$ è la matrice associata a f rispetto a queste basi, allora*

$$\dim \mathrm{Im}(f) = \mathrm{rk}\,\mathbf{A}.$$

Di conseguenza $\dim \mathrm{N}(f) = \dim V - \mathrm{rk}\,\mathbf{A}$.

DIMOSTRAZIONE. Poniamo $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$. Sia $k = \mathrm{rk}\,\mathbf{A}$; allora sappiamo trovare un insieme di$k$ colonne di $\mathbf{A}$ che sia linearmente indipendenti. Per non complicare le notazioni supponiamo che l'insieme sia quello delle prime k colonne. Allora l'insieme

$$\{\mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_1)); \mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_2)); \dots; \mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_k))\}$$

è linearmente indipendente e quindi anche l'insieme

$$\{f(\mathbf{v}_1); f(\mathbf{v}_2); \dots; f(\mathbf{v}_k)\}$$

è linearmente indipendente. Siccome è un insieme in $\operatorname{Im}(f)$ ne segue che $\dim \operatorname{Im}(f) \le k$.

La disuguaglianza inversa si dimostra in modo analogo. Siccome $f[\mathcal{B}]$ è un insieme di generatori di $\operatorname{Im}(f)$, da essa possiamo estrarre una base. Ancora, per non complicare le notazioni, assumiamo che questa sia $\{f(\mathbf{v}_1); \dots; f(\mathbf{v}_{k'})\}$.

Allora $\{\mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_1)); \dots; \mathrm{C}_{\mathcal{D}}(f(\mathbf{v}_{k'}))\}$ è un insieme linearmente indipendente in $\mathbb{C}^n$. Questi vettori sono k' colonne di $\mathbf{A}$, quindi $\dim \operatorname{Im}(f) = k' \le \operatorname{rk} \mathbf{A}$, come si voleva. □

Nella dimostrazione abbiamo usato il fatto che un'applicazione lineare biiettiva, in particolare l'applicazione delle coordinate o la sua inversa, trasforma insiemi linearmente indipendenti in insiemi linearmente indipendenti.

Talvolta è data la matrice associata a un'applicazione lineare rispetto a basi che non sono quelle che ci interessano; quale sarà la relazione fra due matrici associate?

Fissiamo le notazioni: avremo l'applicazione lineare $f\colon V \to W$; le basi $\mathcal{B}$ e $\mathcal{B}'$ di V; le basi $\mathcal{D}$ e $\mathcal{D}'$ di W; la matrice $\mathbf{A}$ associata a f rispetto alle basi $\mathcal{B}$ e $\mathcal{D}$; la matrice $\mathbf{A}'$ associata a f rispetto alle basi $\mathcal{B}'$ e $\mathcal{D}'$.

Assumiamo di conoscere $\mathbf{A}'$ e cerchiamo di calcolare $\mathbf{A}$, usando ovviamente le matrici dei cambiamenti di base. Ci serve una matrice $\mathbf{A}$ con la proprietà che, per ogni $\mathbf{v} \in V$,

$$\mathrm{C}_{\mathcal{D}}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}).$$

Abbiamo a disposizione le seguenti uguaglianze, valide per ogni $\mathbf{v} \in V$ e ogni $\mathbf{w} \in W$:

$$\begin{aligned} &(1) \qquad \mathrm{C}_{\mathcal{D}'}(f(\mathbf{v})) = \mathbf{A}'\,\mathrm{C}_{\mathcal{B}'}(\mathbf{v}) \\ &(2) \qquad \mathrm{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}'}\,\mathrm{C}_{\mathcal{B}'}(\mathbf{v}) \\ &(3) \qquad \mathrm{C}_{\mathcal{D}}(\mathbf{w}) = \mathbf{M}_{\mathcal{D}\leftarrow\mathcal{D}'}\,\mathrm{C}_{\mathcal{D}'}(\mathbf{w}) \end{aligned}$$

Abbiamo allora, indicando sopra il segno di uguaglianza quale relazione stiamo impiegando,

$$\mathrm{C}_{\mathcal{D}}(f(\mathbf{v})) \overset{(3)}{=} \mathbf{M}_{\mathcal{D}\leftarrow\mathcal{D}'}\,\mathrm{C}_{\mathcal{D}'}(f(\mathbf{v})) \overset{(2)}{=} \mathbf{M}_{\mathcal{D}\leftarrow\mathcal{D}'}\mathbf{A}'\,\mathrm{C}_{\mathcal{B}'}(\mathbf{v}) \overset{(1)}{=} \mathbf{M}_{\mathcal{D}\leftarrow\mathcal{D}'}\mathbf{A}'\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}'}^{-1}\,\mathrm{C}_{\mathcal{B}}(\mathbf{v}).$$

Per l'unicità della matrice associata, avremo

$$\mathbf{A} = \mathbf{M}_{\mathcal{D}\leftarrow\mathcal{D}'}\mathbf{A}'\mathbf{M}_{\mathcal{B}\leftarrow\mathcal{B}'}^{-1}.$$

Ancora una volta raccomandiamo di non imparare a memoria la formula, ma di impadronirsi del procedimento, il quale non dipende dalle notazioni usate.