1 Algebra lineare

1.1. *Struttura di uno spazio lineare*

In questo primo paragrafo vogliamo introdurre una nozione piuttosto astratta e generale, che riveste però notevole importanza in tutta l'Algebra lineare e nelle sue varie applicazioni: si tratta del concetto di *spazio lineare o spazio vettoriale* (o, più precisamente, *spazio lineare sul campo reale*). Si consideri un insieme S di elementi qualsiasi e per tali elementi siano definite due operazioni "binarie" (ossia con due "ingredienti"): un'operazione "interna", detta somma o addizione e indicata con l'usuale simbolo +, e un'operazione "esterna", detta prodotto di un elemento di S per un numero reale (o scalare). Gli elementi di S possiamo chiamarli "vettori", anche se nel seguito useremo tale terminologia soprattutto per i vettori di $\mathbb{R}^n$. L'insieme S è allora uno *spazio lineare o spazio vettoriale* (o, più precisamente, *spazio lineare sul campo dei numeri reali*) se sono soddisfatti i seguenti assiomi:

a) *Rispetto all'addizione*

A1. Se $x, y \in S$ anche $x + y \in S$. Ossia S è "chiuso" (in senso algebrico) rispetto all'operazione di addizione tra i suoi elementi. In altre parole, sommando due qualsiasi elementi di S, non "usciamo" da S.

A2. Se $x, y \in S$, allora $x + y = y + x$, ossia l'addizione è *commutativa.*

A3. Se $x, y, z \in S$, allora $(x + y) + z = x + (y + z)$, ossia vale la proprietà *associativa.*

A4. Esiste in S un (unico) elemento neutro, detto *vettore nullo* e indicato con [0], con $\mathbf{0}$, con $\underline{0}$ o in altri modi, tale che $x + [0] = x \quad \forall x \in S$.

A5. Per ogni $x \in S$ esiste un (unico) elemento, detto *opposto di* x, e indicato con $-x$, tale che $x + (-x) = [0]$. La somma $x + (-x)$ si scrive semplicemente $x - x$.

b) *Rispetto alla moltiplicazione per uno scalare (reale)*

M1. Se $x \in S$ e $\lambda \in \mathbb{R}$, allora $\lambda x \in S$, ossia S è "chiuso" (in senso algebrico) rispetto alla moltiplicazione per uno scalare. In altre parole, moltiplicando un qualsiasi elemento di S per un qualsiasi numero reale, non "usciamo" da S.

M2. La moltiplicazione è *distributiva rispetto all'addizione tra scalari*: $(\alpha + \beta)\, x = \alpha x + \beta x,\ \forall \alpha, \beta \in \mathbb{R},\ \forall x \in S$.

M3. La moltiplicazione è *distributiva rispetto all'addizione tra gli elementi di S*: $\lambda(x + y) = \lambda x + \lambda y,\ \forall \lambda \in \mathbb{R},\ \forall x, y \in S$.

M4. Vale la seguente *proprietà associativa*: $\alpha(\beta x) = (\alpha\beta)\, x$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall x \in S$.

M5. Esiste un (unico) *elemento neutro* rispetto alla moltiplicazione per scalari: è (ovviamente) il numero 1: $1x = x$, $\forall x \in S$.

I matematici dicono anche che uno spazio lineare è un *gruppo commutativo.* Dunque qualsiasi insieme di elementi ove sono soddisfatti gli assiomi **A1-A5** e **M1-M5**, dicesi spazio lineare o vettoriale (sopra il campo reale). È immediato vedere che da questi assiomi si possono dedurre altre proprietà, quali, ad esempio,

$-x = (-1)x$;
$0x = [0]$;
$\lambda[0] = [0]$.

In un qualsiasi spazio lineare l'elemento nullo [0] apparterrà quindi sempre a tale insieme. Un esempio di spazio lineare è l'insieme $\mathbb{R}^n$ di tutti i vettori reali a n componenti. Esistono però molti altri insiemi che godono delle proprietà **A1-A5** e **M1-M5** e che quindi sono spazi lineari: ad esempio, l'insieme $\mathbb{C}^n$ di tutti i vettori complessi a n componenti (si veda il § 1.5), l'insieme delle matrici (reali) di ordine (m, n), l'insieme di tutte le funzioni definite su uno stesso dominio $T \subset \mathbb{R}$, l'insieme dei polinomi, l'insieme dei polinomi di grado minore o uguale a n, ecc. *Non* è invece, ad esempio, uno spazio lineare l'insieme dei polinomi di un *assegnato* grado n. La ragione è semplice: se consideriamo, ad es., i due polinomi di 5° grado:

$$a)\ x^5 + 2x^2 + 1; \qquad b)\ -x^5 + 3x^3 - x$$

e ne facciamo la somma, otteniamo

$$3x^3 + 2x^2 - x - 1,$$

ossia un polinomio di 3° grado. Non rispettiamo dunque l'assioma **A1.**

***Definizione* 1.** Un sottoinsieme L (non vuoto) di uno spazio lineare S dicesi *sottospazio lineare di S* se gli elementi di L godono delle proprietà **A1-A5** e **M1-M5** (rispetto a L). ❑

Segue immediatamente che L, sottoinsieme di S, è sottospazio lineare di S se S è a sua volta uno spazio lineare e risulta

1. $x, y \in L \Rightarrow (x + y) \in L$;
2. $x \in L, \lambda \in \mathbb{R} \Rightarrow \lambda x \in L$.

Possiamo anche compendiare tali due proprietà nell'unica implicazione

$$x, y \in L, \quad \alpha_1, \alpha_2 \in \mathbb{R} \Rightarrow (\alpha_1 x + \alpha_2 y) \in L.$$

Ne segue che, ad esempio, in $\mathbb{R}^3$ si hanno i seguenti sottospazi lineari: l'origine, ogni retta passante per l'origine, ogni piano passante per l'origine, l'intero $\mathbb{R}^3$. Ad esempio, l'insieme

$$L_1 = \{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$$

non è un sottospazio lineare di $\mathbb{R}^2$, in quanto l'origine non vi appartiene. L'insieme

$$L_2 = \{x \in \mathbb{R}^2 : x_1 + x_2 = k,\ k \in \mathbb{R}\}$$

è un sottospazio lineare di $\mathbb{R}^2$ solo se è $k = 0$.

1.2. *Questioni sui vettori di* $\mathbb{R}^n$

Come già osservato, $\mathbb{R}^n$ costituisce uno spazio lineare. Questo particolare spazio lineare sarà l'oggetto di questo paragrafo e di parecchi paragrafi e capitoli successivi. Molte questioni sui *vettori reali*, ossia sugli elementi di $\mathbb{R}^n$, ci sono già note dal corso di Matematica Generale. Ricordiamo che con

$$[0] = [0, 0, 0, ..., 0]$$

indichiamo il *vettore nullo* di $\mathbb{R}^n$. Possiamo confrontare tra loro due vettori $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ nel seguente modo:

- $x = y$, se $x_i = y_i$, $\forall i = 1, 2, ..., n$ ($x \neq y$ in caso contrario);
- $x > y$, se $x_i > y_i$ $\forall i = 1, 2, ..., n$.
- $x \geqq y$ (leggere: x maggiore o uguale a y), se $x_i \geqq y_i$, $\forall i = 1, 2, ..., n$. Ad es. $[5, 2] \geqq$ $\geqq [1, -6]$ ma anche $[5, 2] \geqq [5, 2]$.
- $x \geq y$ (leggere: x quasi maggiore di y), se $x \geqq y$ ma $x \neq y$. In altre parole: $x_i \geqq y_i$, $\forall i =$ $= 1, ..., n$, ed $\exists i : x_i > y_i$.

Le relazioni $x \leqq y$, $x \leq y$, $x < y$ sono introdotte in modo analogo. Ovviamente sussistono le implicazioni

$$x > y \Rightarrow x \geq y \Rightarrow x \geqq y.$$

Se risulta, con $[0] \in \mathbb{R}^n$,

- $x \geqq [0]$, si dice che x è vettore *non negativo*;
- $x \geq [0]$, si dice che x è vettore *semipositivo*;
- $x > [0]$, si dice che x è vettore *positivo.*

Analogamente si definiscono i vettori *non positivi* ($x \leqq [0]$) *seminegativi* ($x \leq [0]$) *negativi* ($x < [0]$).

Ricordiamo che i *vettori fondamentali di* $\mathbb{R}^n$ (o *di base* o *elementari* o *unitari*) sono:

$e^1 = [1, 0, 0, ..., 0]$;
$e^2 = [0, 1, 0, ..., 0]$;
............................;
$e^n = [0, 0, ..., 0, 1]$.

Il *vettore somma* (di $\mathbb{R}^n$) sarà indicato con $e = (1, 1, 1, ..., 1)$. Il vettore $x^\top$ è il *trasposto* del vettore x, ossia il vettore che ha le stesse componenti di x, nello stesso ordine, però disposte su colonna se x è riga, su riga se x è colonna.

Sia ora $L \subseteq \mathbb{R}^n$ per quanto detto nel paragrafo precedente, L è un *sottospazio lineare o vettoriale* di $\mathbb{R}^n$ se:

a) $x \in L, y \in L \Rightarrow (x+y) \in L$;
b) $\lambda \in \mathbb{R}, x \in L \Rightarrow \lambda x \in L$.

Sia ora L un sottospazio lineare di $\mathbb{R}^n$ e siano $x^1, x^2, ..., x^m$ vettori di L e $\lambda_1, \lambda_2, ..., \lambda_m$ scalari reali. L'espressione

$$\lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_m x^m = \sum_{i=1}^{m} \lambda_i x^i = \bar{x}$$

è chiamata *combinazione lineare (c.l.)* di $x^1, x^2, ..., x^m$, con *pesi o coefficienti* $\lambda_1, \lambda_2, ..., \lambda_m$. È chiaro che $\bar{x} \in L$.

Un modo per costruire un sottospazio lineare di $\mathbb{R}^n$ (eventualmente coincidente con tutto $\mathbb{R}^n$) è di considerare l'insieme X di *tutte le combinazioni lineari* di certi k vettori dati:

$$X = \left\{ x : x = \sum_{i=1}^{k} \lambda_i x^i \right\}.$$

È facile dimostrare che X è un sottospazio lineare di $\mathbb{R}^n$:

1. Vale la proprietà di "chiusura" rispetto all'addizione. Prendiamo due vettori di X,

$$x = \sum_{i=1}^{k} \lambda_i x^i; \quad y = \sum_{i=1}^{k} \alpha_i x^i.$$

La loro somma è

$$x + y = \sum_{i=1}^{k} \lambda_i x^i + \sum_{i=1}^{k} \alpha_i x^i = \sum_{i=1}^{k} (\lambda_i + \alpha i) x^i$$

ossia è ancora una c. l. dei vettori assegnati.

2. Vale la proprietà di "chiusura" rispetto alla moltiplicazione per scalari. Da

$$\alpha x = \alpha \left(\sum_{i=1}^{k} \lambda_i x^i \right) = \left(\sum_{i=1}^{k} (\alpha \lambda_i \right) x^i$$

si deduce l'analoga conclusione.

Si dice allora che $x^1, x^2, ..., x^k$ sono *vettori generatori o di sostegno* per X e che X è il sottospazio lineare di $\mathbb{R}^n$ *generato* da tali vettori. Nei testi in lingua inglese (ma anche nei testi italiani) si trova scritto

$$X = \textit{span}\ \{x^1, x^2, ..., x^k\}$$

(*to span*: espandere). Si dice anche che $x^1, x^2, ..., x^k$ sono un *sostegno* per X o un *sistema di generatori* per X. Ad esempio, possiamo esprimere qualunque $x \in \mathbb{R}^n$ (ossia tutto $\mathbb{R}^n$) come c. l. dei vettori base $e^1, e^2, ..., e^n$. Difatti è $\lambda_1 e^1 + \lambda_2 e^2 + ... + \lambda_n e^n = [\lambda_1, \lambda_2, ..., \lambda_n]$.

Ricordiamo ora un altro concetto fondamentale: quello di *dipendenza e indipendenza lineare* tra vettori di $\mathbb{R}^n$. Siano $x^1, x^2, ..., x^m$ vettori di un sottospazio lineare L di

$\mathbb{R}^n$. Costruiamo la c. l. $z = \lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_m x^m = \sum_{i=1}^{m} \lambda_i x^i$. È evidente che se scegliamo pesi $\lambda_1 = \lambda_2 = ... = \lambda_m = 0$ otteniamo $z = [0]$. Se tale scelta dei pesi (ossia pesi tutti nulli) è *l'unica* che conduce al vettore [0] allora diciamo che gli m vettori dati sono *linearmente indipendenti (l. i.)* o che il loro insieme è linearmente indipendente o che sono tra loro l. i. Se invece esiste una scelta di pesi $\lambda_1, \lambda_2, ..., \lambda_m$, *non tutti nulli*, tali che

$$\sum_{i=1}^{m} \lambda_i x^i = [0]$$

allora diciamo che quei vettori sono *linearmente dipendenti (l. d.)* o che il loro insieme è linearmente dipendente o che sono tra loro l. d.

Ad esempio, due vettori proporzionali sono l. d.: da x e αx si ha $(-\alpha)x + 1(\alpha x) = [0]$, con 1 e α non nulli. È poi immediato verificare che se tra i vettori dati c'è il vettore nullo [0] allora quei vettori sono l. d. Di contro, i vettori base $e^1, e^2, ..., e^n$ sono l. i. Risulta infatti è $\lambda_1 e^1 + \lambda_2 e^2 + ... + \lambda_n e^n = [\lambda_1, \lambda_2, ..., \lambda_n]$ vettore che è nullo se e solo se $\lambda_1 = \lambda_2 = ... = \lambda_n = 0$.

Valgono poi i seguenti risultati.

Teorema 1. *Se e solo se m* vettori di un sottospazio lineare $L \subseteq \mathbb{R}^n$ sono l. d. uno almeno di essi può essere espresso come c. l. dei rimanenti $(m - 1)$ vettori. Quindi tali vettori sono l. i. se e solo se nessuno di essi è c. l. dei rimanenti vettori.

Dimostrazione. Siano $x^1, x^2, ..., x^m$ vettori l. d. di $L \subseteq \mathbb{R}^n$. Allora esiste almeno un $\lambda_h \neq 0$ $(1 \leqq h \leqq m)$ tale che da

$$\lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_m x^m = [0]$$

è possibile dedurre

$$\sum_{i=1}^{m} \frac{\lambda_i}{\lambda_h} x^i = \frac{\lambda_1}{\lambda_h} x^1 + ... + x^h + ... + \frac{\lambda_m}{\lambda_h} x^m = [0]$$

da cui

$$x^h = \sum_{\substack{i=1 \\ i \neq h}}^{m} \left(-\frac{\lambda_i}{\lambda_h} \right) x^i .$$

Viceversa, se $x^h = \sum_{\substack{i=1 \\ i \neq h}}^{m} \lambda_i x^i$, allora

$$\left(\sum_{\substack{i=1 \\ i \neq h}}^{m} \lambda_i x^i \right) - x^h = [0]$$

pertanto, con $\lambda_h = -1$, i vettori $x^1, x^2, ..., x^m$ sono l. d. ❑

Sfruttando il Teorema 1, possiamo dunque affermare che un insieme di vettori è l. d.:

1. se uno almeno di tali vettori è nullo;
2. se un vettore è multiplo di un altro (ad esempio, se due vettori sono uguali);
3. se un vettore è somma (o differenza) di altri vettori.

Teorema 2. Se tra i vettori $x^1, x^2, \dots, x^m$ di L ve ne sono $k < m$ linearmente dipendenti, allora anche gli m vettori assegnati sono l. d.

Dimostrazione. Supponiamo, senza perdere di generalità, che i vettori l. d. siano i primi k tra gli m considerati. Allora esiste almeno un $\lambda_h \neq 0$, con $1 \leqq h \leqq m$, tale che $\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_h x^h + \dots + \lambda_k x^k = [0]$ con $\lambda_h \neq 0$. Se ora si pone $\lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_m = 0$, si ha

$$\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_m x^m = [0],$$

senza che gli scalari λ_i, $i = 1, \dots, m$, siano tutti nulli. ❑

Teorema 3. Se $x^1, x^2, \dots, x^m$ di L sono vettori l. i., allora ogni loro sottoinsieme composto da $k < m$ vettori, è costituito da vettori l. i.

Dimostrazione. Per assurdo si supponga che tra gli m vettori dati, ve ne siano k l. d. Allora, per il teorema precedente qualsiasi insieme che li contiene è costituito da vettori l. d., contro l'ipotesi che gli m vettori dati sono l. i. ❑

Abbiamo ricordato la nozione di *sostegno* per un sottospazio lineare L di $\mathbb{R}^n$. Va subito osservato che un sottospazio lineare L ha quanti sostegni vogliamo. Non è difficile convincersi di tale fatto. Abbiamo visto che, ad esempio, i due vettori fondamentali di $\mathbb{R}^2$, e^1, e^2, generano l'intero $\mathbb{R}^2$. È ovvio allora che anche la terna

$$e^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ e^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix},\ \alpha, \beta \in \mathbb{R}$$

genera ancora $\mathbb{R}^2$: basta considerare la c. l. $\lambda_1 e^1 + \lambda_2 e^2 + \lambda_3 x$, con $\lambda_3 = 0$. Consideriamo il seguente esempio, forse un po' meno banale del precedente. Mostriamo che i vettori

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \end{bmatrix},\ \alpha, \beta \in \mathbb{R}, \alpha, \beta \neq 0,$$

generano, per c. l., tutto $\mathbb{R}^2$ (data l'arbitrarietà di α e β avremo così *infiniti* sostegni per $\mathbb{R}^2$). Deve essere

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \lambda_1 + \alpha\lambda_3 \\ \lambda_1 + \lambda_2 + \beta\lambda_3 \end{bmatrix}.$$

Il sistema

$$\begin{cases} x_1 = \lambda_1 + \alpha\lambda_3 \\ x_2 = \lambda_1 + \lambda_2 + \beta\lambda_3 \end{cases}$$

ha manifestamente infinite soluzioni. Si può prendere, ad esempio, $\lambda_1 = t \in \mathbb{R}$ qualsiasi. Risulta (è $\alpha \neq 0$ per ipotesi)

$$\lambda_3 = \frac{x_1 - t}{\alpha}; \quad \lambda_2 = x_2 - t - \beta\frac{x_1 - t}{\alpha}.$$

Controlliamo:

$$t\begin{bmatrix}1\\1\end{bmatrix} + \left(x_2 - t - \frac{\beta}{\alpha}(x_1 - t)\right)\begin{bmatrix}0\\1\end{bmatrix} + \frac{x_1 - t}{\alpha}\begin{bmatrix}\alpha\\\beta\end{bmatrix} =$$

$$= \begin{bmatrix} t + x_1 - t \\ t + x_2 - t - \frac{\beta}{\alpha}(x_1 - t) + \frac{\beta}{\alpha}(x_1 - t)\end{bmatrix} = \begin{bmatrix}x_1\\x_2\end{bmatrix}.$$

Abbiamo trovato quindi infiniti sostegni per $\mathbb{R}^2$. Desideriamo ora richiamare un'altra fondamentale nozione, diretta ad individuare quegli speciali sostegni che hanno il *numero minimo possibile* di elementi. Proviamo a "limare" il precedente sistema di generatori per $\mathbb{R}^2$ eliminando l'ultimo vettore e considerando quindi solo i primi due vettori

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

Ci si rende conto che tali vettori sono ancora in grado di generare, per c. l., tutto $\mathbb{R}^2$. Cerchiamo λ_1 e λ_2 tali che

$$\begin{bmatrix}x_1\\x_2\end{bmatrix} = \lambda_1\begin{bmatrix}1\\1\end{bmatrix} + \lambda_2\begin{bmatrix}0\\1\end{bmatrix}.$$

Si ha

$$\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}\lambda_1\\\lambda_1 + \lambda_2\end{bmatrix},$$

da cui l'unica soluzione $\lambda_1 = x_1$, $\lambda_2 = x_2 - x_1$.

En passant, facciamo notare che mentre vi erano infiniti modi per rappresentare i vettori di $\mathbb{R}^2$ con i tre generatori considerati precedentemente, qui, fissata la coppia x_1, x_2 vi è un solo modo per ottenere tale coppia a partire da

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

Se partiamo invece dalla coppia

$$\begin{bmatrix}1\\1\end{bmatrix}, \quad \begin{bmatrix}2\\2\end{bmatrix}.$$

sempre ottenuta dai tre vettori generatori considerati precedentemente, eliminando stavolta il secondo vettore e ponendo $\alpha = \beta = 2$ nel terzo, si ha

$$\lambda_1\begin{bmatrix}1\\1\end{bmatrix}+\lambda_2\begin{bmatrix}2\\2\end{bmatrix}=\begin{bmatrix}\lambda_1+2\lambda_2\\\lambda_1+2\lambda_2\end{bmatrix}=\begin{bmatrix}\gamma\\\gamma\end{bmatrix},\ \gamma\in\mathbb{R}.$$

Quindi in questo caso (da notare che i due vettori di partenza sono l. d.) riusciamo a generare soltanto quei vettori di $\mathbb{R}^2$ che hanno le due componenti *uguali* (ossia che giacciono sulla bisettrice del *I* e *III* quadrante del piano cartesiano). Generiamo quindi solo un sottospazio lineare (proprio) di $\mathbb{R}^2$.

Sia dato un sottospazio lineare $L \subseteq \mathbb{R}^n$ e certi vettori $x^1, x^2, ..., x^m$ che lo generano, ossia che sono un sostegno per L. Interessa, a questo punto, individuare un sostegno che contenga il *minimo* numero possibile di elementi con i quali generare tutto L. È subito visto che tali elementi debbono essere l. i. poiché, in caso contrario, uno almeno di essi si potrebbe esprimere come c. l. dei rimanenti $(m-1)$ vettori e il numero dei generatori di L scenderebbe di almeno un'unità (Teorema 1), in quanto ogni altro vettore di L potrebbe essere espresso come c. l. dei rimanenti $(m-1)$ vettori considerati. Difatti, se il vettore "scartato", fosse, per fissare le idee, x^k, $1 \leqq k \leqq m$, risulta

$$x^k = \sum_{i\neq k}\alpha_i x^i$$

e ogni vettore $\bar{x}$ di L che era espresso come c. l. dei vettori di partenza, potrà essere scritto come

$$\bar{x} = \sum_{i\neq k}\lambda_i x^i + \lambda_k x^k = \sum_{i\neq k}\lambda_i x^i + \lambda_k \sum_{i\neq k}\alpha_i x^i = \sum_{i\neq k}(\lambda_i + \lambda_k\alpha_i)x^i.$$

Per contro, sempre grazie al Teorema 1, *nessuno* degli elementi l. i. potrebbe essere tolto dall'insieme dei generatori, senza che l'insieme ottenuto perda la proprietà di generare L: difatti, in particolare, l'elemento tolto non potrebbe esprimersi come c. l. dei rimanenti elementi. *Quindi il numero minimo di generatori di $L \subseteq \mathbb{R}^n$ coincide con il numero massimo di vettori l. i. che è possibile trovare in L.* Sia k tale numero: tale numero viene detto *dimensione di* L e si dice che L è un sottospazio lineare o vettoriale (di $\mathbb{R}^n$) di dimensione k:

$$\dim(L) = k.$$

***Definizione* 4.** Sia dato un sottospazio lineare X di $\mathbb{R}^n$. Si chiama *base* di X (più precisamente: base di Hamel) un insieme di vettori $x^1, x^2, ..., x^k$ di X tali che:

1. i vettori $x^1, x^2, ..., x^k$ sono l. i.;
2. sono un sostegno per X. ❑

Si osservi che un sostegno di un sottospazio lineare non è necessariamente una base per lo stesso, mentre è sempre vero il viceversa: da un sostegno può essere lecito eliminare alcuni vettori, senza che quelli che restano perdano la proprietà di generare lo spazio in questione; da una base ciò non è possibile. Analogia "pittorica" di base:

con i colori fondamentali rosso, blu, giallo, nero, riusciamo a generare ogni sfumatura di ogni colore.

Si può dimostrare inoltre il seguente

Teorema 5. Il sottospazio lineare X di $\mathbb{R}^n$ ammetta una base di k vettori. Allora:

1. ogni altra base di X è costituita da k vettori;
2. comunque si scelgano k vettori l. i. in X, essi costituiscono una base per X;
3. la rappresentazione dei vettori di X come c. l. degli elementi di una prefissata base, è *unica.*

Dalla proprietà 1 consegue che l'essere una base per X non è una proprietà inerente gli specifici vettori che la costituiscono, ma una caratteristica del sottospazio lineare in questione. Insomma ... una base vale l'altra! Possiamo di conseguenza dire anche che dim(X) è il numero dei vettori di una qualsiasi base di X. Quindi *dimensione* di un sottospazio lineare X di $\mathbb{R}^n$ è:

- il numero di vettori di una sua qualsiasi base;
- il numero minimo di vettori generatori di X;
- il numero massimo di vettori l. i. che è possibile trovare in X.

In particolare, la proprietà 3 del Teorema 5 si dimostra facilmente: vogliamo provare che esiste un'unica k-upla di coefficienti $\lambda_1, \lambda_2, ..., \lambda_k$ tale che, con $\bar{x}$ vettore qualsiasi di X, risulti

$$\bar{x} = \sum_{i=1}^{k} \lambda_i x^i .$$

Mostriamo che non possono sussistere due rappresentazioni diverse. Assumiamo, per assurdo, che

$$\bar{x} = \sum_{i=1}^{k} \lambda_i x^i = \sum_{i=1}^{k} \lambda'_i x^i ,$$

con almeno un $\lambda_i \neq \lambda'_i$. Non è restrittivo supporre che i coefficienti diversi riguardino l'ultimo vettore della base, ossia che $\lambda_k \neq \lambda'_k$. In tale caso da $\sum_{i=1}^{k} \lambda_i x^i = \sum_{i=1}^{k} \lambda'_i x^i$ si può trarre

$$x^k = \frac{1}{\lambda_k - \lambda'_k} \sum_{i=1}^{k-1} (\lambda_i - \lambda'_i) x^i = \sum_{i=1}^{k-1} \gamma_i x^i$$

ove

$$\gamma_i = \frac{(\lambda_i - \lambda'_i)}{(\lambda_k - \lambda'_k)}, \qquad \text{per } k = 1, 2, ..., k-1.$$

L'ultima relazione ci svela che in tal caso x^k è c. l. degli altri vettori $x^1, x^2, ..., x^{k-1}$, ma ciò è in contrasto con la indipendenza lineare dei vettori di una base.

***Definizione* 6.** Si chiama *base naturale o canonica o standard* di $\mathbb{R}^n$ la base formata dagli n vettori fondamentali (o di base) e^i, $i = 1, ..., n$. ❑

Infatti tali vettori sono l. i. (come già osservato) e generano, per c. l., ogni vettore di $\mathbb{R}^n$ (come già osservato). Dunque possiamo dire che $\mathbb{R}^n$ ha dimensione n:

$$\dim(\mathbb{R}^n) = n.$$

Va osservato che non sempre è possibile determinare una base, composta da un numero finito di elementi, per uno spazio lineare qualsiasi S. Un esempio particolarmente significativo di uno spazio lineare S in cui non esiste alcun sistema finito di elementi soddisfacenti la definizione di base, è fornito dall'insieme di *tutti* i polinomi a coefficienti reali, in una variabile x (insieme che costituisce uno spazio lineare). È assurdo pensare che possa esistere una base costituita da un numero finito di polinomi $P_1(x)$, $P_2(x)$, ..., $P_s(x)$, in grado di generare l'insieme di tutti i polinomi (di qualsiasi grado). Tali spazi lineari, molto importanti in Matematica, sono gli *spazi lineari a infinite dimensioni.* Da questi oggetti ci terremo rispettosamente alla larga.

Ricordiamo che, con $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ dicesi *prodotto scalare o prodotto interno* tra x e y il numero

$$xy = x_1y_1 + x_2\,y_2 + ... + x_n\,y_n = \sum_{i=1}^{n} x_i\,y_i\,.$$

(È bene, per ragioni "pratiche", abituarci all'idea che x sia vettore riga e y sia vettore colonna).

Ricordiamo ancora che il prodotto scalare non gode della "legge dell'annullamento del prodotto": può cioè essere $xy = 0$ con $x \neq [0]$ e $y \neq [0]$ Allorché risulti, con $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ *non nulli*, $xy = 0$, i vettori x e y sono detti *ortogonali.* Ad esempio, sono ortogonali, i vettori fondamentali e^i, e^j, con $i \neq j$.

Teorema 7. Se $x^1, x^2, ..., x^m$ sono vettori non nulli (di $\mathbb{R}^n$) a due a due ortogonali, allora sono l. i.

Dimostrazione. Consideriamo la relazione

$$\lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_m x^m = [0],$$

ove $x^1, ..., x^m$ sono a due a due ortogonali. Eseguiamo il prodotto scalare di tale c. l. con il vettore x^i, $1 \leqq i \leqq m$. Si ha

$$x^i(\lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_m x^m) = 0.$$

Pertanto, poiché i vettori (non nulli) sono a due a due ortogonali, si ha $\lambda_1 x^i x^i = 0$ ed essendo $x^i x^i > 0$, segue $\lambda_i = 0$. Per l'arbitrarietà di i, segue $\lambda_1 = \lambda_2 = \lambda_m = 0$. ❑

Per ogni $x \in \mathbb{R}^n$ diciamo *norma (euclidea)*, denotata con $\|x\|$, la quantità non negativa

$$\|x\| = \sqrt{x^\top x} = \sqrt{\sum_{i=1}^{n} (x_i)^2}\,.$$

Nei casi $n = 1, 2, 3$, essa si riduce, rispettivamente, al modulo di un numero reale, alla lunghezza del vettore x in $\mathbb{R}^2$ alla lunghezza del vettore x in $\mathbb{R}^3$. La norma gode di alcune fondamentali proprietà:

1. $\|x\| \geqq 0, \forall x \in \mathbb{R}^n$;
2. $\|x\| = 0$ se e solo se $x = [0]$;
3. $\|\alpha x\| = |\alpha|\, \|x\|, \forall \alpha \in \mathbb{R}, \forall x \in \mathbb{R}^n$;
4. $\|x + y\| \leqq \|x\| + \|y\|, \forall x, y \in \mathbb{R}^n$.

La proprietà 4 è detta “disuguaglianza triangolare”. In due o tre dimensioni essa si riduce al ben noto teorema della geometria elementare che afferma che la lunghezza di un lato di un triangolo non supera la somma delle lunghezze degli altri due lati. Tale disuguaglianza è a sua volta conseguenza della proprietà

$$|xy| \leqq \|x\|\, \|y\|,$$

nota come “disuguaglianza di Cauchy-Schwarz”. Se x e y sono due vettori *non nulli* di $\mathbb{R}^n$ ha senso scrivere il quoziente

$$\frac{xy}{\|x\|\,\|y\|},$$

che risulta compreso tra – 1 e 1, a causa della disuguaglianza di Cauchy-Schwarz. Tale quoziente si chiama *coseno* dell’angolo tra x e y ed è denotato con $\cos(\widehat{xy})$. Risulta che due vettori non nulli sono ortogonali se e solo se il coseno dell’angolo da essi formato è zero.

Un vettore *non nullo* x di $\mathbb{R}^n$ dicesi *normalizzato* se la sua norma è unitaria: $\|x\| = 1$. È sempre possibile normalizzare un vettore $x \neq [0]$: basta porre $y = \dfrac{1}{\|x\|}x$. Il vettore y risulta normalizzato; infatti si ha

$$\|y\| = \left\|\frac{1}{\|x\|}x\right\| = \frac{1}{\|x\|}\|x\| = 1.$$

Notiamo che i vettori fondamentali e^i, $i = 1, ..., n$, sono normalizzati.

***Osservazione* 8.**

1. Altre notazioni usate in Matematica per denotare il prodotto interno tra due vettori x e y sono (x, y); $\langle x, y\rangle$.

2. I vettori normalizzati sono anche detti *versori*; in particolare, i vettori fondamentali e^i, $i = 1, ..., n$, sono anche detti *versori degli assi.*

3. I vettori normalizzati e ortogonali sono anche detti *ortonormali.* Se essi costituiscono una base per un certo sottospazio lineare $L \subseteq \mathbb{R}^n$ si parlerà di *base ortonormale.* I vettori e^i, $i = 1, ..., n$, costituiscono allora una base ortonormale per $\mathbb{R}^n$. Il cosiddetto "procedimento di Grahm-Schmidt" assicura che, a partire da una qualsiasi base di $\mathbb{R}^n$ è sempre possibile costruire una base ortonormale di $\mathbb{R}^n$.

Dati due vettori $x, y \in \mathbb{R}^n$ diciamo *distanza euclidea* tra x e y la quantità non negativa

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \ .$$

Risulta quindi che la norma $\|x\|$ altro non è che la distanza euclidea di x dal vettore nullo, ossia dall'origine di $\mathbb{R}^n$. Anche $d(x, y)$ gode di proprietà simili a quelle della norma:

a) $d(x, y) = 0$, $\forall x, y \in \mathbb{R}^n$, essendo $d(x,y) = 0$ se e solo se $x = y$;
b) $d(x, y) = d(y, x)$ (proprietà di simmetria);
c) $d(x, y) \leqq d(x, z) + d(z, y)$, $\forall x, y, z \in \mathbb{R}^n$ (disuguaglianza triangolare).

Attraverso la distanza possiamo introdurre la nozione di *intorno circolare* (o brevemente: intorno) di raggio $\varepsilon > 0$ di un punto x^* in $\mathbb{R}^n$: è l'insieme dei vettori $x \in \mathbb{R}^n$ la cui distanza euclidea da x^* è minore di ε, ossia

$$U(x^*, \varepsilon) = \{x : x \in \mathbb{R}^n, \ \|x - x^*\| < \varepsilon\}.$$

Quando $n = 1$, $U(x^*, \varepsilon)$ è l'intervallo $(x^* - \varepsilon, x^* + \varepsilon)$ quando $n = 2$ avremo il cerchio (privato della circonferenza, ossia del bordo) di centro x^* e raggio ε; quando $n = 3$ avremo la sfera (o bolla) di centro x^* e raggio ε (sfera privata della "superficie sferica").

Già che ci siamo, richiamiamo alcune nozioni fondamentali di topologia ordinaria per $\mathbb{R}^n$. Dato un insieme $A \subseteq \mathbb{R}^n$ e un punto $x^0 \in \mathbb{R}^n$ diciamo che x^0 è *interno* ad A se non solo $x^0 \in A$ ma anche esiste un intorno $U(x^0, \varepsilon)$ tutto contenuto in A. Diciamo che x^0 è *esterno* ad A se non solo $x^0 \notin A$ ma anche esiste un intorno $U(x^0, \varepsilon)$ che non ha punti in comune con $A : A \cap U(x^0, \varepsilon) = \varnothing$.

Per un insieme $A \subseteq \mathbb{R}^n$ i punti che non sono né interni né esterni sono detti *punti di frontiera.* Un punto x^0 è cioè di frontiera se ogni suo intorno contiene almeno un punto di A e un punto di A^c essendo A^c *l'insieme complementare* di A rispetto a $\mathbb{R}^n$: $A^c = \mathbb{R}^n \backslash A$. Si noti che i punti di frontiera possono appartenere o non appartenere ad A. L'insieme dei punti interni di A si denota con *int*(A) o con $\mathring{A}$ e si chiama *interno* di A. L'insieme dei punti di frontiera di A si denota con ∂A e si chiama *frontiera* di A. Il punto $x^0 \in \mathbb{R}^n$ è di *accumulazione* per A se ogni suo intorno contiene almeno un altro punto $x \in A$, con $x \neq x^0$ (si osservi che, in tale caso, ogni suddetto intorno contiene necessariamente infiniti punti $x \in A$). Il punto $x^0 \in A$ si dice *isolato* se esiste un suo intorno che non contiene nessun altro punto di A (oltre a x^0).

***Esempio* 9.** Per l'insieme

$$A = [3, 10) \cup \{15\}$$

risulta:

- 5 è interno ad A ed è anche di accumulazione per A;
- 3, 10, 15 sono punti di frontiera per A; 3 e 10 sono anche di accumulazione per A;
- 0 è punto esterno ad A;
- 15 è punto isolato di A.

L'insieme di tutti i punti di accumulazione di A si chiama *insieme derivato* di A e si denota con A'. Si chiama *chiusura* di A l'insieme $\bar{A}$ dato da $\mathring{A} \cup \partial A$, ovvero costituito da tutti i punti interni e da quelli di frontiera. Risulta anche $\bar{A} = A \cup A'$.

***Esempio* 10.** Con riferimento all'insieme dell'esempio 9, risulta:

- $\mathring{A} = (3, 10)$;
- $\partial A = \{3, 10, 15\}$;
- $A' = [3, 10]$;
- $\bar{A} = [3, 10] \cup \{15\}$.

Si chiamano *aperti* quegli insiemi che hanno solo punti interni (A è quindi aperto se $A = \mathring{A}$). Nessun punto di frontiera appartiene quindi ad un insieme aperto. A è invece *chiuso* se ogni suo punto di frontiera appartiene ad $A : \partial A \subseteq A$. Equivalentemente: se la sua chiusura coincide con $A : A = \partial A \cup \mathring{A}$. L'insieme vuoto $\varnothing$ e l'intero $\mathbb{R}^n$ sono gli unici insiemi contemporaneamente aperti e chiusi. Esistono anche insiemi che non sono né aperti né chiusi: ad esempio l'insieme A dell'Esempio 9. Inoltre si dimostra che A è chiuso se e solo se contiene tutti i suoi punti di accumulazione: $A' \subseteq A$. Si dimostra anche che A è chiuso (rispettivamente: aperto) se e solo se il suo complementare A^c è aperto (rispettivamente: chiuso).

Un insieme A che sia tutto dentro un intorno dell'origine, ossia tale per cui esiste $U([0], \varepsilon)$, con $A \subseteq U([0], \varepsilon)$ si dice *limitato*; se A (insieme di $\mathbb{R}^n$) è sia chiuso che limitato si chiama anche *compatto.* Passiamo infine alla *connessione.* Sia X un insieme chiuso (rispettivamente: aperto) di $\mathbb{R}^n$. Si dice che X è *connesso* se *non* esistono due sottoinsiemi X_1 e X_2 non vuoti, entrambi chiusi (rispettivamente: entrambi aperti), tali che $X_1 \cap X_2 = \varnothing$, $X_1 \cup X_2 = X$. Informalmente si può dire che X è connesso se è costituito da un solo "pezzo". È abbastanza intuitivo che in $\mathbb{R}$ gli unici insiemi connessi siano gli intervalli.

1.3. *Richiami su matrici, determinanti, sistemi lineari*

Definizione rigorosa di *matrice A di ordine* (m, n) è: A è un insieme di mn elementi dotato di un doppio ordinamento completo, ovvero in corrispondenza biunivoca con le mn coppie (i, j) del prodotto cartesiano $M \times N = \{1, 2, ..., m\} \times \{1, 2, ..., n\}$. Più prosaicamente diciamo che è un insieme di mn elementi disposti su una tabella di m righe e n colonne:

$$A = [a_{ij}], \qquad i = 1, ..., \quad m; j = 1, 2, ..., n.$$

Tale nozione, con ogni a_{ij} numero reale, è ben nota dal corso di Matematica Generale e quindi non staremo in questa sede a ripetere tutte le relative classificazioni e proprietà. Notiamo solo che anche l'insieme delle matrici (reali) di ordine (m, n) è uno spazio lineare (di dimensione mn). Similmente a quanto fatto per i vettori di $\mathbb{R}^n$ possiamo confrontare due matrici (ad elementi reali, non staremo più a ripeterlo) $A = [a_{ij}]$ e $B = [b_{ij}]$, dello stesso ordine (m, n):

• $A = B$ se (e solo se) $a_{ij} = b_{ij}$, $\forall i, j$. ($A \neq B$ in caso contrario);
• $A \geqq B$, se è $a_{ij} \geqq b_{ij}$, $\forall i, j$;
• $A \geq B$, se è $A \geqq B$, però è $A \neq B$. In altre parole: è $a_{ij} \geqq b_{ij}$, $\forall i, j$, però esiste almeno una coppia a_{ij}, b_{ij} per la quale è $a_{ij} > b_{ij}$;
• $A > B$, se è $a_{ij} > b_{ij}$, $\forall i, j$.

Se denotiamo con [0] la *matrice nulla* (di ordine (m, n)), diremo:

• *A non negativa*, se è $A \geqq [0]$;
• *A semipositiva*, se è $A \geq [0]$;
• *A positiva*, se è $A > [0]$.

È ovvio il significato di: A non positiva, A seminegativa, A negativa.

Indicheremo con A_i, $i = 1, ..., m$, l'i-esimo vettore riga di A (brevemente: la i-esima riga di A) e con A^j, $j = 1, ..., n$, il j-esimo vettore colonna di A (brevemente: la j-esima colonna di A). Ciò permette di considerare A di ordine (m, n) come un insieme ordinato di m vettori riga (di $\mathbb{R}^n$), oppure di n vettori colonna (di $\mathbb{R}^m$). Di conseguenza, possiamo porre

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \\ A_m \end{bmatrix} \qquad \text{oppure} \qquad A = [A^1, A^2, ..., A^n].$$

Quando ci si vuole riferire indifferentemente alle righe o alle colonne di A, si parla anche di *linee* di A. Ricordiamo che se A è *quadrata* ($m = n$), A è *simmetrica* se $A = A^{\mathsf{T}}$, con A^{T} trasposta di A. Sempre con $m = n$, *matrice diagonale* è una matrice con elementi al di fuori della diagonale principale ("elementi extra-diagonali") tutti nulli:

$$D = \begin{bmatrix} \lambda_1 & 0 \ldots 0 \\ 0 & \lambda_2 \ldots 0 \\ \multicolumn{2}{c}{\ldots\ldots\ldots\ldots} \\ 0 & 0 \ldots \lambda_n \end{bmatrix}.$$

Introducendo il *simbolo di Kronecker*

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j, \end{cases}$$

risulta quindi $D = [\delta_{ij}, \lambda_i]$ $i, j = 1, ..., n$. Se nella matrice diagonale D è $\lambda_1 = \lambda_2 = ... = \lambda_n = \lambda$, parleremo di *matrice scalare* e se è $\lambda_1 = \lambda_2 = ... = \lambda_n = 1$ parleremo di *matrice identica o matrice unità*:

$$I = [\delta_{ij}], \qquad i, j = 1, ..., n.$$

Sempre con $m = n$ (matrici quadrate), si dice *matrice di permutazione o di scambio* una matrice, indicata usualmente con P o con Π, ottenuta scambiando tra loro due o più righe (o colonne; quindi due o più linee) della matrice identica I.

***Esempio* 1.** Da

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

si può ottenere

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \text{oppure} \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \text{oppure} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ ecc.}$$

Una matrice quadrata tale per cui è

$$AA^{\mathsf{T}} = A^{\mathsf{T}}A = I$$

si dice *ortogonale* (ad esempio le matrici di permutazione P). Quindi A è ortogonale se $A^{\mathsf{T}} = A^{-1}$, essendo ovviamente A^{-1} la *matrice inversa* di A, ossia tale che $AA^{-1} = A^{-1}A = I$. Data una matrice quadrata A, si può dimostrare che le seguenti condizioni sono equivalenti:

1. A è ortogonale;
2. i vettori colonna di A sono ortonormali;
3. i vettori riga di A sono ortonormali.

Si ripassino le operazioni sulle matrici e in particolare il concetto di *prodotto* tra matrici (conformabili).

Data A quadrata, chiameremo *potenza* n-esima di A (n intero positivo) il prodotto di n fattori uguali ad A:

$$A^{(n)} = \underbrace{AAA\ldots A}_{n \text{ fattori}}.$$

Ovviamente $I^{(n)} = I$, $A^{(1)} = A$; con A non nulla si pone poi $A^{(0)} = I$. Si noti che nell'algebra delle matrici non valgono, in genere, le regole che valgono nell'algebra degli scalari. Ad esempio, risulta in genere, $(A + B)^{(2)} \neq A^{(2)} + 2AB + B^{(2)}$, essendo in genere $AB \neq BA$.

Cosa sia il *determinante* $|A|$, oppure $\det(A)$, $\det A$, di una matrice *quadrata* A è noto dal corso di Matematica Generale e non staremo qui a ripetere le nozioni collegate a tale concetto. Ricordiamo però il seguente importante teorema che giustifica, per così dire, la nozione (un po' astrusa) di determinante.

Teorema 2. Le linee di una matrice quadrata A sono l. i. se e solo se $|A| \neq 0$. Di conseguenza le linee di A sono l. d. se e solo se $|A| = 0$.

Ricordiamo poi che se $|A| \neq 0$, allora A viene detta *regolare, invertibile, non singolare*; di contro, se $|A| = 0$, la matrice A viene detta *non regolare, non invertibile, singolare.*

Ricordiamo che dicesi *minore complementare* di a_{ij}, indicato con A_{ij} il determinante della matrice (quadrata) $\bar{A}$, ottenuta da A sopprimendo la i-esima riga e la j-esima colonna. Dicesi poi *complemento algebrico* di a_{ij} indicato con M_{ij} il numero

$$(-1)^{i+j} A_{ij}.$$

Richiamiamo il cosiddetto *primo teorema di Laplace.*

Teorema 3 (Primo teorema di Laplace). Il determinante di A è dato dalla somma degli elementi di una linea qualsiasi di A per i rispettivi complementi algebrici.

Ricordiamo che, dato n $\in \mathbb{N}_+$ (naturali positivi), si chiama *fattoriale di n* il numero

$$n! = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1.$$

Si pone, per definizione, $0! = 1$. Si chiama poi *coefficiente binomiale* il numero, che si indica con $\binom{n}{k}$ (leggere: "n su k"):

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Ciò premesso, diamo le seguenti definizioni.

***Definizione* 4.** Data A di ordine (m, n) dicesi *minore di ordine t* di A ogni determinante della sottomatrice quadrata di ordine t, composta da t righe e t colonne (non scelte necessariamente come le righe) di A. ❑

Ovviamente risulta $1 \leqq t \leqq \min(m, n)$. Si dimostra che da A, di ordine (m, n), è possibile estrarre

$$\binom{m}{t}\binom{n}{t}$$

minori di ordine t. Ad esempio, se A è di ordine (3, 4), possiamo estrarre da A:

• 4 minori di ordine 3, ossia

$$[A^1, A^2, A^3]; \quad [A^1, A^2, A^4]; \quad [A^1, A^3, A^4]; \quad [A^2, A^3, A^4];$$

- 18 minori di ordine 2;
- 12 minori di ordine 1.

Con l'occasione, ricordiamo che il *rango o caratteristica* di A, di ordine (m, n), è l'ordine massimo dei suoi minori *non nulli.* Tale numero, denotato con $r(A)$ oppure semplicemente con r, coincide con il massimo numero di righe (oppure, indifferentemente, di colonne) linearmente indipendenti che è possibile trovare in A. Si ha perciò:

$$A \text{ di ordine}(m, n) \ \Rightarrow 0 \leqq r(A) \leqq \min(m, n),$$

essendo $r(A) = 0$ se e solo se $A = [0]$. È poi $r(A) = r(A^\top)$. Se risulta $r(A) = \min(m, n)$, si dice che A ha rango *pieno o massimo.*

***Definizione* 5.** Data A *quadrata* di ordine n, dicesi *minore principale di ordine t* di A ogni determinante della sottomatrice quadrata di ordine t composta da t righe e dalle *corrispondenti t* colonne di A. ❑

Si dimostra che da A, di ordine n, è possibile estrarre $\binom{n}{t}$ minori principali di ordine t. Ad esempio, da

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

è possibile estrarre:

- 1 minore principale di ordine 3, ossia $|A|$;
- $\binom{3}{2} = 3$ minori principali di ordine 2, ossia:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$

- $\binom{3}{1} = 3$ minori principali di ordine 1: $|a_{11}|, |a_{22}|, |a_{33}|$.

La somma t_k di tutti gli $\binom{n}{k}$ minori principali di ordine k è detta *traccia di A di ordine k.* La traccia di ordine 1, ossia

$$t_1 = a_{11} + a_{22} + \ldots + a_{nn}$$

viene detta semplicemente *traccia* di A e viene spesso indicata con $tr(A)$. Ovviamente $t_n = |A|$.

***Definizione* 6.** Data A quadrata di ordine n, i *minori principali di Nord-Ovest o minori principali di guida* di A (in inglese: *N.W. principal minors, leading principal minors*) sono tutti quei determinanti formati dalle *prime t* righe e dalle *prime t* colonne di A ($t = 1, ..., n$). I minori principali di Nord-Ovest di A sono perciò dati dalla seguente successione di determinanti:

$$\left|a_{11}\right|, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, ..., \left|A\right|. \qquad \square$$

Ripassiamo ora sinteticamente le principali questioni relative ai *sistemi lineari.* Come è noto dal corso di Matematica Generale, per *sistema lineare* si intende un "sistema di equazioni algebriche lineari", ossia un sistema di m equazioni nelle incognite $x_1, x_2, ..., x_n$ del tipo:

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

che riscriviamo nella *forma matriciale*

$$Ax = b,$$

essendo *A matrice dei coefficienti*, di ordine (m, n), *x vettore (colonna) delle incognite* (di $\mathbb{R}^n$) e *b vettore (colonna) dei termini noti* (di $\mathbb{R}^n$). Le domande sensate relative ad un sistema siffatto sono:

- stabilire sotto quali condizioni esso ammette soluzioni;
- nel caso in cui esistano soluzioni, quante siano;
- quali siano le soluzioni (nel caso che il sistema sia risolubile).

Osserviamo subito che possiamo scrivere il sistema nella forma

$$x_1A^1 + x_2A^2 + ... + x_nA^n = b,$$

che evidenzia che il vettore dei termini noti b è dato dalla c. l. delle colonne di A: le incognite non sono altro che i coefficienti di tale combinazione lineare. Di conseguenza, $Ax = b$ ammette soluzioni se e solo se b può esprimersi come c. l. delle colonne di A. Ciò è equivalente al ben noto

Teorema di Rouché-Capelli. Condizione necessaria e sufficiente affinché b sia combinazione lineare delle colonne di A, ossia che $Ax = b$ ammetta soluzioni, è che $r(A) = r(A \mid b)$, ove $r(A)$ esprime il rango di A e $r(A \mid b)$ è il rango della *matrice completa (o matrice aumentata)*, matrice ottenuta da A accostando alle colonne di A il vettore dei termini noti b.

Se poi risulta $r(A) = r(A \mid b) = n$, la soluzione è *unica*, ossia il sistema è *determinato*; se è $r(A) = r(A \mid b) = r < n$, il sistema è *indeterminato* e si dice allora che vi sono

∞^{n-r} soluzioni, che il sistema ammette $n - r$ "gradi di libertà". Facciamo notare che se un sistema lineare ammette due soluzioni distinte, ne ammette allora infinite, ossia un sistema lineare che ammette soluzioni o ne ammette una sola o ne ammette infinite. Siano infatti x^1 e x^2 due soluzioni *diverse*: quindi $Ax^1 = b$ e $Ax^2 = b$. Sia t un numero reale qualsiasi e si considerino gli infiniti vettori $x(t) = tx^1 + (1 - t)x^2$. Tali infiniti vettori sono pure soluzioni del sistema; basta calcolare $Ax(t) = A[tx^1 + (1 - t)x^2] = tAx^1 + Ax^2 - tAx^2 = tb + b - tb = b$.

Riassumendo: dato il sistema lineare

$$Ax = b$$

con A di ordine (m, n), $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ possono verificarsi i seguenti casi.

a) Sistema impossibile, ossia

$$r(A \mid b) \neq r(A).$$

Più precisamente: $r(A) = k$, $r(A|b) = k + 1$.

b) Sistema possibile (risolubile, compatibile), ossia

$$r(A \mid b) = r(A) = r.$$

Notiamo che il teorema di Rouché-Capelli è *automaticamente* verificato (e quindi i relativi sistemi sono sempre solubili) nel caso di $r(A) = m$ ("sistemi normali") e nel caso $b = [0]$ ("sistemi omogenei").

b_1) Se (e solo se) è $r(A) = r(A \mid b) = n$, la soluzione è *unica* (ad esempio, i sistemi *crameriani*, ossia i sistemi quadrati con $|A| \neq 0$).

b_2) Se è $r(A) = r(A \mid b) = r < n$, il sistema ammette ∞^{n-r} soluzioni (ad esempio, i sistemi normali, con $r(A) = m < n$).

Spendiamo ancora qualche parola sui *sistemi lineari omogenei*, ossia

$$Ax = [0].$$

Come già notato, il teorema di Rouché-Capelli vale qui "by default": difatti tali sistemi ammettono sempre almeno la "soluzione banale" (o "soluzione nulla") $\bar{x} = [0]$. In base a quanto precedentemente osservato, possiamo asserire che:

- se (e solo se) è $r(A) = n$, la soluzione $\bar{x} = [0]$ è *l'unica* che il sistema ammette;
- se (e solo se) è $r(A) < n$, ci sono anche soluzioni $\bar{x} \neq [0]$. dette "autosoluzioni" o "soluzioni proprie" o "soluzioni non banali";
- quindi se $Ax = [0]$ è un sistema *quadrato*, esso ammette autosoluzioni se e solo se $|A| = 0$.

Si rifletta su quanto appena evidenziato. Risulta

$$Ax = [0] \Leftrightarrow x_1A^1 + x_2A^2 + ... + x_nA^n = [0].$$

ed è $x_1 = x_2 = ... = x_n = 0$ soluzione *unica* se e solo se le colonne $A^1, A^2, ..., A^n$, sono l. i., ossia se e solo se $r(A) = n$.

Importante è poi il seguente teorema, chiamato talvolta (ad esempio da Simon e Blume: si veda la bibliografia posta alla fine del volume) *Teorema fondamentale dell'Algebra Lineare.*

Teorema 7. Sia A di ordine (m, n) e rango r. L'insieme delle soluzioni X^0 di $Ax = [0]$ costituisce un sottospazio lineare di $\mathbb{R}^n$, di dimensione $(n - r)$. Come vedremo più avanti, X^0 si chiama anche *nucleo* di A.

Dunque, se $r = n$, la dimensione di tale sottospazio è zero, ossia tale sottospazio contiene solo il vettore nullo. Se è $r = 0$, la dimensione è n: infatti in tale caso è $Ax = [0]$ ed è ovviamente $[0]x = [0]$, $\forall x \in \mathbb{R}^n$.

Dimostrazione del Teorema 7. Dimostrare che l'insieme delle soluzioni di $Ax = [0]$ è un sottospazio lineare di $\mathbb{R}^n$ è facile. Sia $X^0 = \{x : Ax = [0]\}$ da $x^1 \in X^0$, $x^2 \in X^0$, $\lambda \in R$ segue:

a) $A(x^1 + x^2) = Ax^1 + Ax^2 = [0] + [0] = [0]$ cioè $(x^1 + x^2) \in X^0$;
b) $A(\lambda x^1) = \lambda Ax^1 = \lambda[0] = [0]$, $\lambda x^1 \in X^0$.

Circa la dimensione di X^0 abbiamo già osservato che nelle ipotesi $r = 0$ oppure $r = n$, essa vale $(n - r)$. Resta allora da considerare solo l'ipotesi $0 < r < n$. A tale scopo occorre: *a*) individuare in X^0 un insieme di $(n - r)$ vettori l. i.; *b*) mostrare che ogni vettore $x \in X^0$ può essere espresso come c. l. di quelli.

Siano dati k vettori $u^1, ..., u^k$ che costituiscano una base per il nucleo X^0. Dobbiamo dimostrare che $k = n - r$. Estendiamo $u^1, ..., u^k$ a formare una base per $\mathbb{R}^n$ (ciò è sempre possibile), aggiungendo gli $(n - k)$ vettori $u^{k+1}, ..., u^n$. Dimostreremo il teorema mostrando che i vettori

$$Au^{k+1}, ..., Au^n$$

costituiscono una base per lo spazio generato dalle colonne di A, spazio denotato con $C(A)$. Osserviamo che $Au^{k+1}, ..., Au^n$ sono combinazioni lineari delle colonne A^j, $j = 1, ..., n$. Si ricordi infatti che $Ax = b$ può essere scritto come

$$A^1x_1 + A^2x_2 + ... + A^nx_n = b.$$

Quindi $Au^{k+1}, ..., Au^n$ appartengono a $C(A)$. Inoltre, tali vettori sono linearmente indipendenti: supponiamo che esistano scalari $c_{k+1}, ..., c_n$, tali che

$$c_{k+1}\, Au^{k+1} + ... + c_n\, Au^n = [0]. \tag{1}$$

Risulta quindi

$$A(c_{k+1}\, u^{k+1} + ... + c_n\, u^n) = [0].$$

Di conseguenza il vettore $c_{k+1}\, u^{k+1} + ... + c_n\, u^n$ appartiene al nucleo X^0. Ma allora risulta essere combinazione lineare dei vettori $u^1, ..., u^k$ che per ipotesi formano una base per X^0. Esistono quindi scalari $c_1, ..., c_k$ tali che

$$c_1\, u^1 + ... + c_k\, u^k = c_{k+1}\, u^{k+1} + ... + c_n\, u^n,$$

ossia

$$c_1\, u^1 + ... + c_k\, u^k - c_{k+1}\, u^{k+1} - ... - c_n\, u^n = [0].$$

I vettori $u^1, ..., u^n$ sono linearmente indipendenti, in quanto formano, per ipotesi,