4. Richiami: sistemi lineari e matrici

Vettori

4.a Combinazioni lineari

Indichiamo con $\mathbb{R}^n$ l'insieme delle n-uple ordinate di elementi di $\mathbb{R}$,

$$\mathbb{R}^n := \Big\{x = (x^1, x^2, \dots, x^n) \,\Big|\, x^i \in \mathbb{R},\ i = 1, \dots, n\Big\}.$$

Si dice che $x = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$ è un *vettore* di $\mathbb{R}^n$ di componenti $x^1, x^2, \dots, x^n$.

4.1 Convenzioni. Indichiamo le componenti di un vettore con un indice *posto in alto*. Conviene distinguere tra i vettori o n-uple di $\mathbb{R}^n$ —che sono oggetti matematici— e la loro rappresentazione grafica sulla carta. In particolare, in connessione con il *prodotto righe per colonne* è d'uso disporre le componenti di un vettore *in colonna*. Ad esempio si dovrebbe scrivere: "sia

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \in \mathbb{R}^n \text{ ''},$$

una forma di scrittura assai poco pratica. Pertanto quando non assolutatmente necessario, si preferisce scrivere le componenti in riga come in "sia $x = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$ ".

Dati k scalari $\lambda^1, \lambda^2, \dots, \lambda^k \in \mathbb{R}$ e k vettori

$$a_1 := \begin{pmatrix} a_1^1 \\ a_1^2 \\ \vdots \\ a_1^n \end{pmatrix},\ a_2 := \begin{pmatrix} a_2^1 \\ a_2^2 \\ \vdots \\ a_2^n \end{pmatrix},\ \dots\ , a_k := \begin{pmatrix} a_k^1 \\ a_k^2 \\ \vdots \\ a_k^n \end{pmatrix},$$

si forma il vettore *combinazione lineare* di $a_1, \dots, a_k$ a coefficienti $\lambda^1, \lambda^2, \dots, \lambda^k$

$$\sum_{i=1}^{n} \lambda^i a_i := \begin{pmatrix} \lambda^1 a_1^1 + \lambda^2 a_2^1 + \dots \lambda^k a_k^1 \\ \lambda^1 a_1^2 + \lambda^2 a_2^2 + \dots \lambda^k a_k^2 \\ \dots \\ \lambda^1 a_1^n + \lambda^2 a_2^n + \dots \lambda^k a_k^n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^k \lambda^j a_j^1 \\ \sum_{j=1}^k \lambda^j a_j^2 \\ \dots \\ \sum_{j=1}^k \lambda^j a_j^n \end{pmatrix}. \tag{4.1}$$

4.2 Definizione. *Si dice che*

- *un insieme $W \subset \mathbb{R}^n$ è un* sottospazio *vettoriale se tutte le combinazioni lineari finite di vettori in W sono ancora in W,*
- *Dato un insieme di vettori $\{a_1, a_2, \dots, a_k\}$, lo spazio vettoriale di tutte le combinazioni lineari di $a_1, a_2, \dots, a_k$ si chiama* sottospazio vettoriale generato *da $\{a_1, a_2, \dots, a_k\}$ e si indica con*

$$\mathrm{Span}\Big\{a_1, a_2, \dots, a_k\Big\} := \Big\{w \in \mathbb{R}^n \,\Big|\, \exists\, \lambda^1, \dots, \lambda^k \text{ tali che } w = \sum_{i=1}^k \lambda^i a_i\Big\}$$

- *k vettori $a_1, a_2, \dots, a_k \in \mathbb{R}^n$ sono* linearmente indipendenti *se da*

$$\lambda^1 a_1 + \dots + \lambda^k a_k = 0, \qquad \lambda^1, \lambda^2, \dots, \lambda^k \in \mathbb{R}$$

  segue che $\lambda^1 = \dots = \lambda^k = 0$.

Sia W un sottospazio di $\mathbb{R}^n$. Un insieme di vettori $\mathcal{B} \subset W$ linearmente indipendenti che genera W si chiama una *base* di W. Se $(a_1, a_2, \dots, a_k)$ è una base *ordinata* di W, allora ogni $x \in W$ si scrive in modo *unico* come $x := \sum_{i=1}^k x^i a_i$. La n-upla $(x^1, x^2, \dots, x^k)$ sono le *coordinate* di x rispetto alla base $(a_1, a_2, \dots, a_k)$.

Ovviamente $\mathbb{R}^n$ è uno *spazio vettoriale* e le n-uple $e_1 := (1, 0, \dots, 0)$, $e_2 := (0, 1, \dots, 0)$,..., $e_n = (0, 0, \dots, 1)$ formano una base ordinata di $\mathbb{R}^n$ essendo

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x^1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x^2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x^n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^n x^i e_i. \tag{4.2}$$

$(e_1, e_2, \dots, e_n)$ è la cosiddetta base ordinata *canonica* o *standard* di $\mathbb{R}^n$.
Si dimostra che

4.3 Teorema. *Ogni sottospazio $W \subset \mathbb{R}^n$ ha una base e tutte le basi di W hanno lo stesso numero di elementi. Questo numero si chiama* la dimensione *di W e si indica con $\dim W$. Inoltre*

(i) *$\dim W \le n$,*
(ii) *se $k = \dim W$, allora k vettori di W linearmente indipendenti formano una base di W,*
(iii) *se $k = \dim W$, allora dati $h < k$ vettori $v_1, \dots, v_h$ linearmente indipendenti di W si possono trovare ulteriori $(k - h)$ vettori $v_{h+1}, \dots, v_k$ in modo che $(v_1, v_2, \dots, v_k)$ sia una base di W.*

$$\begin{pmatrix} (\mathbf{A}x)^1 \\ \vdots \\ \boxed{(\mathbf{A}x)^i} \\ \vdots \\ (\mathbf{A}x)^m \end{pmatrix} := \begin{pmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ & & \vdots & \\ \boxed{a^i_1 \quad a^i_2 \quad \dots \quad a^i_n} \\ & & \vdots & \\ a^m_1 & a^m_2 & \dots & a^m_n \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{matrix}} \end{pmatrix}$$

Figura 4.1. Il prodotto righe per colonne di $\mathbf{A} \in M_{m,n}$ con $x \in \mathbb{R}^n$.

Applicazioni lineari

4.b Prodotto righe per colonne di matrici

Sia $\mathbf{A} \in M_{m,n}(\mathbb{R})$ una matrice a m righe e n colonne, o come si dice una matrice $m \times n$. Indichiamo con a^i_j l'elemento di $\mathbf{A}$ nella posizione riga i e colonna j. Per ogni $x = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$, il *prodotto righe per colonne di* $\mathbf{A}$ *per* x è il vettore di $\mathbb{R}^m$ indicato con $\mathbf{A}x$,

$$\mathbf{A}x := ((\mathbf{A}x)^1, \dots (\mathbf{A}x)^m),$$

di componenti date da

$$(\mathbf{A}x)^i := \sum_{j=1}^n a^i_j x^j, \qquad i = 1, \dots, m, \tag{4.3}$$

o più esplicitamente

$$\begin{pmatrix} (\mathbf{A}x)^1 \\ \vdots \\ (\mathbf{A}x)^m \end{pmatrix} = \begin{pmatrix} a^1_1 x^1 + a^1_2 x^2 + \dots + a^1_n x^n \\ a^2_1 x^1 + a^2_2 x^2 + \dots + a^2_n x^n \\ \vdots \\ a^m_1 x^1 + a^m_2 x^2 + \dots + a^m_n x^n \end{pmatrix}.$$

Si osservi che, se $a_1, a_2, \dots, a_n$ sono le colonne di $\mathbf{A}$ e si scrive

$$\mathbf{A} = \Big[a_1 \Big| a_2 \Big| \dots \Big| a_n \Big],$$

allora

$$\mathbf{A}x = x^1 a_1 + x^2 a_2 + \dots x^n a_n. \tag{4.4}$$

4.c Applicazioni lineari e matrici

Una applicazione $A : \mathbb{R}^n \to \mathbb{R}^m$ si dice *lineare* se

$$A(\lambda x + \mu y) = \lambda A(x) + \mu A(y) \qquad \forall x, y \in \mathbb{R}^n, \ \forall \lambda, \mu \in \mathbb{R}.$$

In particolare $A(0) = 0$. Per induzione si verifica subito che $A : \mathbb{R}^n \to \mathbb{R}^m$ è lineare se e solo se

$$A\Big(\sum_{i=1}^k \lambda^i v_i\Big) = \sum_{i=1}^k \lambda^i A(v_i) \tag{4.5}$$

per ogni $k \in \mathbb{N}$, per ogni $\lambda^1, \lambda^2, \dots, \lambda^k \in \mathbb{R}$ e ogni $v_1, v_2, \dots, v_k \in \mathbb{R}^n$.

Applicazioni lineari da $\mathbb{R}^n$ su $\mathbb{R}^m$ e matrici $m \times n$ sono in corrispondenza biunivoca.

Se $\mathbf{A} \in M_{m,n}(\mathbb{R})$, si verifica subito che la mappa $A : \mathbb{R}^n \to \mathbb{R}^m$ definita da

$$A(x) := \mathbf{A}x \qquad \text{righe per colonne}$$

è lineare. L'applicazione $x \to \mathbf{A}x$ si chiama l'*applicazione lineare* associata alla matrice $\mathbf{A}$.

Viceversa, sia $L : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare, e sia $\mathbf{L}$ la matrice le cui colonne sono nell'ordine i vettori immagini $L(e_1), L(e_2), \dots, L(e_n)$ della base standard di $\mathbb{R}^n$,

$$\mathbf{L} := \Big[L(e_1)\Big|L(e_2)\Big|\dots\Big|L(e_n)\Big].$$

Allora per ogni $x = (x^1, x^2, \dots, x^n)$ segue dalle (4.2) e (4.5) che

$$L(x) := x^1 L(e_1) + x^2 L(e_2) + \dots x^n L(e_n) = \mathbf{L}x$$

righe per colonne. Si dice che $\mathbf{L}$ è la *matrice associata all'applicazione lineare L.*

4.d Nucleo e immagine

Sia $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{R})$ e sia $A : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare associata ad $\mathbf{A}$, $A(x) = \mathbf{A}x$, $x \in \mathbb{R}^n$.

L'immagine di A e il nucleo di A sono rispettivamente

$$\ker A := \Big\{x \in \mathbb{R}^n \,\Big|\, \mathbf{A}x = 0\Big\},$$
$$\operatorname{Im} A := \Big\{y \in \mathbb{R}^m \,\Big|\, \exists\, x \in \mathbb{R}^n \text{ tale che } \mathbf{A}x = y\Big\}.$$

$\ker A$ è un sottospazio vettoriale di $\mathbb{R}^n$ ed è facile provare che A è iniettiva se e solo se $\ker A = \{0\}$. In questo caso si dice che la matrice $\mathbf{A}$ è *non-singolare.*

$\operatorname{Im} A$ è un sottospazio lineare di $\mathbb{R}^m$ e in effetti è l'insieme delle combinazioni lineari delle colonne $(a_1, a_2, \dots, a_n)$ di $\mathbf{A}$,

$$\operatorname{Im} A = \operatorname{Span}\Big\{a_1, a_2, \dots, a_n\Big\}.$$

In altre parole *le colonne di $\mathbf{A}$ sono un insieme di generatori per* $\operatorname{Im} A$. La dimensione di $\operatorname{Im} A$ è dunque inferiore o uguale a $\min(n, m)$ e si chiama anche *rango* della matrice $\mathbf{A}$ e si indica con $\operatorname{Rank} \mathbf{A}$. Evidentemente A è surgettiva se e solo se $\operatorname{Im} A = \mathbb{R}^m$ o, equivalentemente, se e solo se $\dim \operatorname{Im} A = m$. Si prova poi l'importante *formula del rango*

$$\dim \ker A + \dim \operatorname{Im} A = n. \tag{4.6}$$

Segue in particolare che, *se $\mathbf{A}$ è una matrice quadrata, $\mathbf{A} \in M_{n,n}(\mathbb{R})$, allora la mappa $x \to \mathbf{A}x$, $x \in \mathbb{R}^m$, è surgettiva se e solo se è iniettiva.*

4.e Prodotto fra matrici

Se $\mathbf{A} \in \mathcal{M}_{p,n}(\mathbb{R})$ e $\mathbf{B} \in \mathcal{M}_{n,q}(\mathbb{R})$, è definito il *prodotto righe per colonne di* $\mathbf{A}$ *per* $\mathbf{B}$ come la matrice $\mathbf{C} \in \mathcal{M}_{p,q}(\mathbb{R})$ il cui elemento c^i_j di posto riga i-esima e colonna j-esima, $1 \leq i \leq p$, $1 \leq j \leq q$, è dato da

$$c^i_j := (\mathbf{AB})^i_j := \sum_{k=1}^{n} a^i_k b^k_j$$

essendo a^i_j, b^i_j rispettivamente gli elementi delle matrici $\mathbf{A}$, $\mathbf{B}$ di posto riga i-esima e colonna j-esima. È facile verificare che

(i) il prodotto di matrici è associativo, i.e., $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$,
(ii) il prodotto di matrici è distributivo, i.e., $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$,
(iii) il prodotto di matrici *non è in generale commutativo*, $\mathbf{AB} \neq \mathbf{BA}$; anzi, non è detto che si possa eseguire $\mathbf{BA}$, anche se si può eseguire $\mathbf{AB}$.
(iv) può accadere che $\mathbf{AB}$ sia la matrice nulla senza che né $\mathbf{A}$ né $\mathbf{B}$ lo siano.

Il prodotto righe per colonne fra matrici corrisponde alla composizione delle applicazioni lineari associate. Infatti, siano $\mathbf{A} \in M_{p,n}(\mathbb{R})$, $\mathbf{B} \in \mathcal{M}_{n,q}(\mathbb{R})$ e siano $A : \mathbb{R}^n \to \mathbb{R}^p$ e $B : \mathbb{R}^q \to \mathbb{R}^n$ le applicazioni lineari associate rispettivamente ad $\mathbf{A}$ e $\mathbf{B}$. Allora $\forall x \in \mathbb{R}^q$

$$A \circ B(x) = A(B(x)) = A(\mathbf{B}x) = \mathbf{A}(\mathbf{B}x) = (\mathbf{AB})x$$

righe per colonne.

Segue dalla formula del rango

4.4 Proposizione. *Sia* $\mathbf{A} \in M_{n,n}(\mathbb{R})$ *e* $A(x) := \mathbf{A}x$ *righe per colonne. Sono fatti equivalenti*

(i) A *è bigettiva,*
(ii) A *è non singolare,* $\ker A = \{0\}$,
(iii) *esiste* $\mathbf{B} \in M_{n,n}(\mathbb{R})$ *tale che* $\mathbf{AB} = \mathrm{Id}_n$,
(iv) *esiste* $\mathbf{B} \in M_{n,n}(\mathbb{R})$ *tale che* $\mathbf{BA} = \mathrm{Id}_n$,

4.f Matrici e sistemi lineari

Le proprietà delle applicazioni lineari sono strettamente legate alla risolubilità dei sistemi lineari. Sia dato il sistema lineare

$$\begin{cases} a^1_1 x^1 + a^1_2 x^2 + \cdots + a^1_n x^n = b^1, \\ a^2_1 x^1 + a^2_2 x^2 + \cdots + a^2_n x^n = b^2, \\ \dots\dots \\ a^m_1 x^1 + a^m_2 x^2 + \cdots + a^m_n x^n = b^m. \end{cases} \tag{4.7}$$

Se si ordinano in una matrice $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{R})$,

$$\mathbf{A} = \begin{pmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^m_1 & a^m_2 & \dots & a^m_n \end{pmatrix},$$

i coefficienti del *sistema lineare* in (4.7), in un vettore $x = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$ le n incognite $x^1, x^2, \dots, x^n \in \mathbb{R}$ e se si indica con $b := (b^1, \dots, b^m) \in \mathbb{R}^m$ il vettore dei termini noti, il sistema (4.7) si riscrive in modo più compatto come

$$\mathbf{A}x = b. \tag{4.8}$$

Si prova, ad esempio usando la formula del rango e il metodo di eliminazione di Gauss per la risoluzione dei sistemi lineari

4.5 Proposizione. *Sia* $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{R})$*. Si ha*

(i) *se* $n > m$*, il sistema* $\mathbf{A}x = 0$ *ha sempre soluzioni non nulle,*
(ii) (Rouché-Capelli) $\mathbf{A}x = b$ *è risolubile se e solo se, dette* $a_1, a_2, \dots, a_n$ *le* n *colonne di* $\mathbf{A}$*, si ha* $\operatorname{Rank}[a_1 \,|\, \dots \,|\, a_n] = \operatorname{Rank}[a_1 \,|\, \dots \,|\, a_n \,|\, b]$.
(iii) *se* $m = n$*, il sistema* $\mathbf{A}x = b$ *è risolubile per ogni* $b \in \mathbb{R}^n$ *se e solo se* $\mathbf{A}x = 0$ *ha la sola soluzione nulla,*

Inoltre, se $\mathbf{A}^T \in \mathcal{M}_{n,m}(\mathbb{R})$ indica la *matrice trasposta* di $\mathbf{A}$, i.e., $(\mathbf{A}^T)^i_j := \mathbf{A}^j_i$, allora

$$\operatorname{Rank} \mathbf{A}^T = \operatorname{Rank} \mathbf{A}. \tag{4.9}$$

4.g Determinante

Sia $A \in M_{n,n}(\mathbb{R})$. Se $a_1, a_2, \dots, a_n$ sono le n colonne di $\mathbf{A}$, scriviamo

$$\mathbf{A} = \Big[a_1 \Big| a_2 \Big| \dots \Big| a_n \Big].$$

4.6 Definizione. *Il* determinante *è l'unica funzione* $\det : M_{n,n}(\mathbb{R}) \to \mathbb{R}$ *con le seguenti proprietà.*

(i) (Linearità su ciascuna colonna) *per ogni* $a'_i, a''_i \in \mathbb{R}^n$, $i = 1, \dots, n$, *e* $\lambda \in \mathbb{R}$

$$\det \Big[\dots, a'_i + a''_i, \dots \Big] = \det \Big[\dots, a'_i, \dots \Big] + \det \Big[\dots, a''_i, \dots \Big],$$
$$\det \Big[\dots, \lambda a_i, \dots \Big] = \lambda \det \Big[\dots, a_i, \dots \Big],$$

(ii) (Alternanza) *scambiando tra loro due colonne adiacenti, il determinante cambia segno,*

$$\det \Big[\dots, a_i, a_{i+1}, \dots \Big] = -\det \Big[\dots, a_{i+1}, a_i, \dots \Big].$$

(iii) (Normalizzazione) $\det \mathrm{Id}_n = 1$, *essendo* Id_n *la matrice* $n \times n$ *data da*

$$\mathrm{Id}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

La proprietà di alternanza e di linearità implicano in particolare che $\det \mathbf{A} = 0$ *se* $\mathbf{A}$ *ha due colonne coincidenti*,

$$\det \Big[\dots, a_i, \dots, a_j, \dots \Big] = 0$$

se $a_i = a_j$ per qualche $i \neq j$.

Si calcola il determinante mediante le *formule* induttive *di Laplace* di sviluppo per colonna e per riga

$$\delta^{kh} \det \mathbf{A} = \sum_{j=1}^{n} (-1)^{h+j} A_j^h \det \mathbf{M}_j^k(\mathbf{A}),$$

$$\delta_{kh} \det \mathbf{A} = \sum_{i=1}^{n} (-1)^{j+h} A_k^i \det \mathbf{M}_h^i(\mathbf{A}),$$

essendo $\mathbf{M}_k^j(\mathbf{A})$ la matrice ottenuta da $\mathbf{A}$ cancellando la *riga j-esima* e la *colonna k-esima* e

$$\delta_{kh} = \begin{cases} 1 & \text{se } h = k, \\ 0 & \text{se } h \neq k \end{cases}$$

il *simbolo di Kronecker*. Un'altra formula importante è

$$\det \mathbf{A} := \sum_{\sigma \in \mathcal{P}_n} (-1)^{\sigma} a^1_{\sigma(1)} a^2_{\sigma(2)} \dots a^n_{\sigma(n)},$$

dove la sommatoria è estesa all'insieme $\mathcal{P}_n$ delle permutazioni di n elementi e naturalmente $\mathbf{A} = (a^i_j)$.

4.7 Teorema. *Si ha*

(i) *una matrice* $\mathbf{A}$ $n \times n$ *è non singolare se e solo se* $\det \mathbf{A} \neq 0$,
(ii) (Determinante della trasposta) $\det \mathbf{A}^T = \det \mathbf{A}$,
(iii) (Formula di Binet) $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$.

La formula induttiva di Laplace perpette di calcolare i coefficienti dell'inversa di una matrice invertibile. Per ogni $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{R})$ indichiamo con $\mathbf{cof}(\mathbf{A}) \in \mathcal{M}_{n,n}(\mathbb{R})$ la *matrice dei cofattori di* $\mathbf{A}$,

$$\mathbf{cof}(\mathbf{A})^i_j := (-1)^{i+j} \det \mathbf{M}^j_i(\mathbf{A}).$$

Attenzione, qui la posizione degli indici è scambiata. La formula di Laplace di sviluppo del determinante per colonna si riscrive allora come

$$\mathbf{A}\mathbf{cof}(\mathbf{A}) = \det \mathbf{A} \; \mathrm{Id}_n, \tag{4.10}$$

e si ha

4.8 Proposizione. *Sia* $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{R})$ *non singolare,* $\det \mathbf{A} \neq 0$*. Allora*

(i) (Formula per l'inversa)

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathbf{cof}(\mathbf{A}).$$

(ii) (Regola di Cramer) *Il sistema* $\mathbf{A}x = b$, $b \in \mathbb{R}^n$, *ha una unica soluzione data da*

$$x = (x^1, x^2, \dots, x^n), \qquad x^i = \frac{\det \mathbf{B}_i}{\det \mathbf{A}},$$

dove

$$\mathbf{B}_i := \Big[a_1 \Big| \dots \Big| a_{i-1} \Big| b \Big| a_{i+1} \Big| \dots \Big| a_n \Big]$$

è la matrice $n \times n$ *ottenuta sostituendo la colonna* i*-esima di* $\mathbf{A}$ *con la colonna* b*.*