Capitolo 2

MATRICI

Fra tutte le applicazioni su uno spazio vettoriale interessa esaminare quelle che mantengono la struttura di spazio vettoriale e che, per questo, vengono dette lineari. La loro importanza risiede soprattutto nel fatto che le applicazioni lineari possono essere espresse mediante matrici.

2.1 Definizione di applicazione lineare.

Siano V e W due spazi vettoriali su $\mathbf{R}$. Un'applicazione

$$f \ : \ V \ \to \ W$$

è detta *lineare* se soddisfa le seguenti condizioni

(a) per ogni $\mathbf{u},\ \mathbf{v} \in V$ è $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$,

(b) per ogni $\alpha \in \mathbf{R}$ e ogni $\mathbf{v} \in V$ è $f(\alpha\mathbf{v}) = \alpha f(\mathbf{v})$.

Ad esempio, l'applicazione $f \ : \ \mathbf{R}^n \ \to \ \mathbf{R}^2$ definita da

$$f(\mathbf{x}) = \mathbf{y}, \quad \text{dove} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} x_1 + x_2 + \cdots + x_n \\ 0 \end{bmatrix},$$

è lineare. Infatti soddisfa (a) e (b). Invece l'applicazione definita da

$$f(\mathbf{x}) = \mathbf{y}, \quad \text{dove} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} |x_1| + |x_2| + \cdots + |x_n| \\ 0 \end{bmatrix},$$

non è lineare, perché non soddisfa la (a).

2.2 Lo spazio vettoriale delle applicazioni lineari.

L'insieme $\mathcal{L}$ delle applicazioni lineari da V a W è uno spazio vettoriale se si definiscono le operazioni di addizione e di moltiplicazione per scalare

$$(f+g)\ :\ V\ \rightarrow\ W \quad \text{e} \quad (\alpha f)\ :\ V\ \rightarrow\ W$$

nel modo seguente

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

L'elemento O di $\mathcal{L}$ risulta l'applicazione f tale che $f(\mathbf{v}) = \mathbf{0}$ per ogni $\mathbf{v} \in V$. Se lo spazio W coincide con V, si può considerare anche l'applicazione *identica* i_V, cioè quella per cui $i_V(\mathbf{v}) = \mathbf{v}$ per ogni $\mathbf{v} \in V$. Inoltre date due applicazioni lineari $f\ :\ V\ \rightarrow\ W$ e $g\ :\ U\ \rightarrow\ V$, l'applicazione composta $f \circ g\ :\ U\ \rightarrow\ W$ è ancora un'applicazione lineare.

Per definire un'applicazione lineare da V a W è sufficiente assegnare i valori che l'applicazione assume sui vettori di una base di V. Vale infatti il seguente teorema.

2.3 Teorema.

Siano V e W due spazi vettoriali su $\mathbf{R}$, con $\dim V = n$, e sia $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base di V. Fissati n elementi $\mathbf{z}_1, \ldots, \mathbf{z}_n$ di W, esiste ed è unica l'applicazione lineare $f \in \mathcal{L}$ tale che

$$\mathbf{z}_i = f(\mathbf{v}_i) \quad \text{per} \quad i = 1, \ldots, n. \tag{1}$$

Dim. Per ogni $\mathbf{v} \in V$ si può scrivere in modo unico $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$. Si definisce

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

È facile verificare che l'applicazione così definita soddisfa la (1) e che è lineare. L'unicità dell'applicazione segue dall'unicità della rappresentazione di $\mathbf{v}$. ∎

Ad esempio, l'applicazione lineare $f\ :\ \mathbf{R}^2\ \rightarrow\ \mathbf{R}^2$ tale che

$$f(\mathbf{x}) = \mathbf{y}, \quad \text{dove} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix},$$

viene definita sulla base canonica $\{\mathbf{e}_1, \mathbf{e}_2\}$ di $\mathbf{R}^2$ nel modo seguente

$$f(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad f(\mathbf{e}_2) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Si ha infatti per ogni $\mathbf{x} \in \mathbf{R}^2$, $\quad \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$,

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ -x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}.$$

Si noti che non è richiesto che i vettori $\mathbf{z}_1, \ldots, \mathbf{z}_n$ siano linearmente indipendenti (cosa che non sarebbe possibile se fosse $\dim W < n$) e neppure che siano tutti distinti.

Ad esempio, l'applicazione $f \;:\; \mathbf{R}^3 \;\to\; \mathbf{R}^2$ tale che

$$f(\mathbf{x}) = \mathbf{y}, \quad \text{dove} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} x_1 + x_2 \\ x_3 \end{bmatrix},$$

viene definita sulla base canonica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ di $\mathbf{R}^3$ nel modo seguente

$$f(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad f(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad f(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Si ha infatti per ogni $\mathbf{x} \in \mathbf{R}^3$, $\quad \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$,

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + x_3 f(\mathbf{e}_3) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_3 \end{bmatrix}.$$

L'applicazione $f \;:\; \mathbf{R}^n \;\to\; \mathbf{R}^2$ tale che

$$f(\mathbf{x}) = \mathbf{y}, \quad \text{dove} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} x_1 + x_2 + \cdots + x_n \\ 0 \end{bmatrix},$$

viene definita sulla base canonica $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ di $\mathbf{R}^n$ nel modo seguente

$$f(\mathbf{e}_i) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{per} \quad i = 1, \ldots, n.$$

Si ha infatti per ogni $\mathbf{x} \in \mathbf{R}^n$, $\quad \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$,

$$f(\mathbf{x}) = \sum_{i=1}^{n} x_i f(\mathbf{e}_i) = \sum_{i=1}^{n} x_i \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + \cdots + x_n \\ 0 \end{bmatrix}.$$

2.4 Matrice associata ad una applicazione lineare.

Siano $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base di $\mathbf{R}^n$ e $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ una base di $\mathbf{R}^m$. Un'applicazione lineare $f : \mathbf{R}^n \to \mathbf{R}^m$ risulta definita, per il teorema 2.2, dai valori $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n) \in \mathbf{R}^m$, cioè in termini della base di $\mathbf{R}^m$ dagli scalari a_{ij}, per $i = 1, \ldots, m$ e $j = 1, \ldots, n$ tali che

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m, \\ f(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m, \\ &\vdots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m. \end{aligned} \tag{2}$$

Questi scalari si possono disporre in forma di tabella con m righe ed n colonne

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a cui si dà il nome di *matrice* associata ad f. Viceversa, data una matrice A di elementi reali a_{ij}, per $i = 1, \ldots, m$ e $j = 1, \ldots, n$, risulta definita l'applicazione lineare (2). Si stabilisce quindi una corrispondenza biunivoca tra l'insieme $\mathcal{L}$ delle applicazioni lineari da $\mathbf{R}^n$ a $\mathbf{R}^m$ e l'insieme $\mathbf{R}^{m\times n}$ delle matrici ad elementi reali con m righe ed n colonne.

Ad esempio, consideriamo l'applicazione $f : \mathbf{R}^3 \to \mathbf{R}^2$ definita sulla base canonica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ di $\mathbf{R}^3$ da

$$f(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad f(\mathbf{e}_2) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad f(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

Se anche in $\mathbf{R}^2$ si sceglie la base canonica, che chiameremo $\{\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2\}$ per distinguerla da quella di $\mathbf{R}^3$, risulta

$$f(\mathbf{e}_1) = \boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2, \quad f(\mathbf{e}_2) = 2\boldsymbol{\epsilon}_1 - \boldsymbol{\epsilon}_2, \quad f(\mathbf{e}_3) = 3\boldsymbol{\epsilon}_2,$$

e la matrice associata ad f è

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

2.5 Prodotto di matrice per vettore.

Dato un vettore $\mathbf{v}$ di $\mathbf{R}^n$, si esprimono $\mathbf{v}$ e $f(\mathbf{v})$ in termini delle basi di $\mathbf{R}^n$ e $\mathbf{R}^m$:

$$\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{v}_j \quad \text{e} \quad f(\mathbf{v}) = \sum_{i=1}^{m} y_i \mathbf{w}_i.$$

Vogliamo trovare la relazione fra i coefficienti x_j e i coefficienti y_i. Si ha, per linearità

$$f(\mathbf{v}) = \sum_{j=1}^{n} x_j f(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\mathbf{w}_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}x_j \Big) \mathbf{w}_i,$$

da cui

$$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad \text{per} \quad i = 1, \ldots, m. \tag{3}$$

Questa relazione si può esprimere in termini della matrice A e dei due vettori $\mathbf{x} \in \mathbf{R}^n$ e $\mathbf{y} \in \mathbf{R}^m$, le cui componenti sono rispettivamente gli x_j e gli y_i, mediante un'operazione definita su $\mathbf{R}^{m\times n} \times \mathbf{R}^n$, a valori in $\mathbf{R}^m$, che viene indicata con

$$\mathbf{y} = A\mathbf{x} \tag{4}$$

e detta *prodotto di matrice per vettore.* Graficamente si ha

$$\begin{bmatrix} \vdots \\ y_i \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

2.6 Lo spazio vettoriale $\mathbf{R}^{m\times n}$.

Dalle proprietà di linearità segue in modo naturale che se A e B sono due matrici associate alle applicazioni lineari f e g rispettivamente, all'applicazione $f + g$ è associata la matrice

$$C = A + B, \quad \text{i cui elementi sono} \quad c_{ij} = a_{ij} + b_{ij},$$

e all'applicazione αf, con α scalare, è associata la matrice

$$C = \alpha A, \quad \text{i cui elementi sono} \quad c_{ij} = \alpha a_{ij}.$$

Su $\mathbf{R}^{m\times n}$ si sono così definite le operazioni di addizione e di moltiplicazione per scalare, ed è facile verificare che $\mathbf{R}^{m\times n}$ è uno spazio vettoriale su $\mathbf{R}$, in cui l'elemento neutro è la matrice O con tutti gli elementi nulli.

2.7 Alcune definizioni.

Sia $A \in \mathbf{R}^{m\times n}$ (i numeri m e n indicano le *dimensioni* di A). Se $m = n$ la matrice è detta quadrata (in tal caso m è l'*ordine* di A). Gli elementi a_{ij} tali che $i = j$ vengono detti elementi *diagonali* o *principali* di A e formano la *diagonale principale* di A.

Una matrice quadrata è

diagonale se $a_{ij} = 0$ per $i \neq j$;
scalare se è diagonale e $a_{ii} = \alpha \in \mathbf{R}$;
triangolare superiore (inferiore) se $a_{ij} = 0$ per $i > j$ (per $i < j$);
triangolare superiore (inferiore) in senso stretto se $a_{ij} = 0$ per $i \geq j$ (per $i \leq j$);
tridiagonale se $a_{ij} = 0$ per $|i - j| > 1$.

Una particolare matrice diagonale è la matrice *identica* I

$$I = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

È facile verificare che la matrice I di ordine n è associata all'applicazione identica da $\mathbf{R}^n$ a $\mathbf{R}^n$, se si sceglie la stessa base sui due spazi.

Una matrice B ottenuta da A eliminando un certo numero di righe e colonne è detta *sottomatrice* di A.

Per esempio dalla matrice

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

si può ottenere la sottomatrice di ordine 2

$$\begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}.$$

2.8 Prodotto di matrici.

Oltre all'applicazione $f : \mathbf{R}^n \to \mathbf{R}^m$ consideriamo anche un'applicazione $g : \mathbf{R}^p \to \mathbf{R}^n$, dove in $\mathbf{R}^p$ si fissa la base $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. All'applicazione f è associata la matrice $A \in \mathbf{R}^{m\times n}$, all'applicazione g è associata la matrice $B \in \mathbf{R}^{n\times p}$. Vogliamo determinare la matrice $C \in \mathbf{R}^{m\times p}$ associata alla applicazione composta $f \circ g : \mathbf{R}^p \to \mathbf{R}^m$.

Esprimendo $g(\mathbf{u}_j)$, per $j = 1, \ldots, p$ in termini della base di $\mathbf{R}^n$, si ha

$$g(\mathbf{u}_j) = \sum_{k=1}^{n} b_{kj}\mathbf{v}_k,$$

quindi per la linearità della f e per la (2)

$$f(g(\mathbf{u}_j)) = \sum_{k=1}^{n} b_{kj} f(\mathbf{v}_k) = \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{m} a_{ik}\mathbf{w}_i = \sum_{i=1}^{m} \Big(\sum_{k=1}^{n} a_{ik} b_{kj} \Big) \mathbf{w}_i.$$

Gli elementi c_{ij} della matrice C risultano quindi

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad \text{per} \quad i = 1, \ldots, m, \quad j = 1, \ldots, p. \tag{5}$$

La (5) definisce un'operazione da $\mathbf{R}^{m\times n} \times \mathbf{R}^{n\times p}$ in $\mathbf{R}^{m\times p}$, indicata con

$$C = A\, B \tag{6}$$

e detta *prodotto di matrici.* Il prodotto (6) si chiama anche prodotto *riga per colonna* perché il generico elemento di C risulta per la (5) uguale al prodotto scalare tra la i-esima riga di A e la j-esima colonna di B. Graficamente si ha

$$\begin{bmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{bmatrix} = \begin{bmatrix} \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \end{bmatrix} \begin{bmatrix} \cdots & b_{1j} & \cdots \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{bmatrix}.$$

È evidente che è possibile moltiplicare due matrici A e B solo se il numero delle colonne di A è uguale al numero delle righe di B. A parte questo, A potrebbe anche avere una sola riga e quindi coincidere con un vettore $\mathbf{x}^T$ o B una sola colonna e quindi coincidere con un vettore $\mathbf{x}$, come nel caso del prodotto (4) di matrice per vettore.

Per esempio, dati i due vettori

$$\mathbf{x} = [1, 2, -1]^T, \quad \text{e} \quad \mathbf{y} = [2, 1, -3]^T,$$

risulta

$$\mathbf{x}\,\mathbf{y}^T = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -3 \\ 4 & 2 & -6 \\ -2 & -1 & 3 \end{bmatrix}.$$

Attenzione a non confondere questo prodotto con il prodotto scalare fra vettori, che dà come risultato lo scalare

$$\mathbf{x}^T\mathbf{y} = 7.$$

L'elemento neutro rispetto alla moltiplicazione è la matrice identica I, associata all'applicazione identica $i_{\mathbf{R}^n}$.

2.9 Il prodotto di matrici non è commutativo.

Nel caso $m = p$, è definita anche l'applicazione $g \circ f$ a cui è associata la matrice BA. In generale è $f \circ g \neq g \circ f$ e quindi $AB \neq BA$, cioè il prodotto di matrici non è commutativo.

Ad esempio, fissiamo in $\mathbf{R}^3$ la base canonica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ e consideriamo le due applicazioni $f, g \ : \ \mathbf{R}^3 \ \to \ \mathbf{R}^3$ definite da

$$\begin{aligned} &f(\mathbf{e}_1) = \mathbf{e}_2, \quad f(\mathbf{e}_2) = \mathbf{e}_3, \quad f(\mathbf{e}_3) = \mathbf{e}_1, \\ &g(\mathbf{e}_1) = \mathbf{e}_1 + \mathbf{e}_2, \quad g(\mathbf{e}_2) = \mathbf{0}, \quad g(\mathbf{e}_3) = \mathbf{e}_3. \end{aligned}$$

Le due matrici associate a f e g sono

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Alla applicazione $f \circ g$ è associata la matrice

$$AB = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

mentre alla applicazione $g \circ f$ è associata la matrice

$$BA = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

2.10 Matrice inversa.

Un'applicazione lineare $f : \mathbf{R}^n \to \mathbf{R}^n$ si dice *invertibile* se esiste un'altra applicazione $g : \mathbf{R}^n \to \mathbf{R}^n$ tale che le due applicazioni composte $g \circ f$ e $f \circ g$ sono uguali all'applicazione identica,

$$g \circ f = f \circ g = i_{\mathbf{R}^n}.$$

g è detta *applicazione inversa* di f e si indica con f^{-1}.

Sia $A \in \mathbf{R}^{n\times n}$ la matrice associata ad f. Se f è invertibile, si dice che A è *invertibile* e la matrice associata ad f^{-1} è detta matrice *inversa* di A e si indica con A^{-1}. Dalla definizione risulta

$$AA^{-1} = A^{-1}A = I.$$

Ad esempio, consideriamo l'applicazione $f : \mathbf{R}^2 \to \mathbf{R}^2$ definita sulla base canonica da

$$f(\mathbf{e}_1) = 2\mathbf{e}_1 - \mathbf{e}_2, \quad f(\mathbf{e}_2) = 2\mathbf{e}_2.$$

La matrice associata è

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}. \tag{7}$$

È facile verificare che $A^2 = 4A - 4I$, da cui si ricava che

$$A - \frac{1}{4}\, A^2 = I,$$

cioè

$$A \left(I - \frac{1}{4}\, A\right) = \left(I - \frac{1}{4}\, A\right) A = I.$$

Ne segue che

$$A^{-1} = I - \frac{1}{4}\, A = \begin{bmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{bmatrix}. \tag{8}$$

Quindi l'applicazione f è invertibile e l'applicazione inversa f^{-1} è definita da

$$f^{-1}(\mathbf{e}_1) = \frac{1}{2}\,\mathbf{e}_1 - \frac{1}{4}\,\mathbf{e}_2, \quad f^{-1}(\mathbf{e}_2) = \frac{1}{2}\,\mathbf{e}_2.$$

Nel seguito si vedrà come il calcolo dell'inversa di una matrice, quando esiste, possa essere affrontato con procedimenti di carattere generale.

Consideriamo le due applicazioni $f, g : \mathbf{R}^n \to \mathbf{R}^n$ e le corrispondenti matrici associate A e $B \in \mathbf{R}^{n\times n}$. Se f e g sono invertibili, anche l'applicazione $h = f \circ g$ è invertibile e risulta $h^{-1} = g^{-1} \circ f^{-1}$, come è facile verificare. Da questo segue che

$$(AB)^{-1} = B^{-1}A^{-1}.$$

2.11 Matrice trasposta.

È spesso utile, come vedremo più avanti, considerare una matrice come un insieme di vettori colonna affiancati; vi sono però molte proprietà che valgono anche quando alle colonne si sostituiscono le righe della matrice. In questi casi è utile sfruttare la notazione di trasposizione già introdotta per i vettori.

Se $A \in \mathbf{R}^{m\times n}$, la matrice *trasposta* è la matrice A^T di $\mathbf{R}^{n\times m}$ che si ottiene scrivendo le righe come colonne. Ad esempio

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

È facile dimostrare che valgono le proprietà

$$(A^T)^T = A, \;\; (A+B)^T = A^T + B^T, \;\; (A\,B)^T = B^T A^T, \;\; (A^T)^{-1} = (A^{-1})^T.$$

2.12 Immagine, nucleo e rango di una matrice.

Sia $A \in \mathbf{R}^{m\times n}$. Si definisce *immagine* di A l'insieme

$$S(A) = \{\mathbf{y} \in \mathbf{R}^m \;:\; \mathbf{y} = A\mathbf{x} \quad \text{per qualche} \quad \mathbf{x} \in \mathbf{R}^n\}$$

e *nucleo* di A l'insieme

$$N(A) = \{\mathbf{x} \in \mathbf{R}^n \;:\; A\mathbf{x} = \mathbf{0}\}.$$

È facile dimostrare che $S(A)$ è un sottospazio di $\mathbf{R}^m$ e $N(A)$ è un sottospazio di $\mathbf{R}^n$. Si definisce *rango* di A il numero

$$\mathrm{rk}(A) = \dim S(A).$$

2.13 Teorema.

Sia $A \in \mathbf{R}^{m\times n}$. rk($A$) è uguale al numero delle colonne linearmente indipendenti di A.

Dim. Sia $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canonica di $\mathbf{R}^n$ e sia c il numero delle colonne linearmente indipendenti di A. Per $i = 1, \ldots, n$ sia $\mathbf{a}_i = A\mathbf{e}_i$ la i-esima colonna di A. Poiché $\mathbf{a}_i \in S(A)$, è $c \le \dim S(A)$. Consideriamo poi un qualsiasi $\mathbf{y} \in S(A)$, quindi esiste $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{y} = A\mathbf{x}$ e si ha

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i, \quad \mathbf{y} = A \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n x_i A\mathbf{e}_i = \sum_{i=1}^n x_i \mathbf{a}_i.$$

Quindi $\mathbf{y}$ può essere espresso come combinazione delle colonne di A, delle quali c sono linearmente indipendenti, cioè ogni $\mathbf{y} \in S(A)$ risulta combinazione lineare delle c colonne linearmente indipendenti di A. ■

Ad esempio, consideriamo la matrice (detta *diade*) di $\mathbf{R}^{n\times n}$

$$A = \mathbf{x}\,\mathbf{y}^T, \quad \text{dove} \quad \mathbf{x},\, \mathbf{y} \in \mathbf{R}^n,\ \mathbf{x},\, \mathbf{y} \neq \mathbf{0}.$$

A ha rango 1. Infatti le colonne di A sono i vettori

$$y_1\,\mathbf{x},\ y_2\,\mathbf{x},\ \ldots,\ y_n\,\mathbf{x},$$

che sono a due a due linearmente dipendenti.

2.14 Teorema.

Sia $A \in \mathbf{R}^{m\times n}$*. Allora*

$$\dim S(A) + \dim N(A) = n.$$

Dim. Se fosse $\dim S(A) = 0$, sarebbe $S(A) = \{\mathbf{0}\}$ e quindi $N(A) = \mathbf{R}^n$ e $\dim N(A) = n$. Consideriamo allora il caso che $\dim S(A) = r > 0$ e sia $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ una base di $S(A)$. Siano $\mathbf{v}_1, \ldots, \mathbf{v}_r \in \mathbf{R}^n$ tali che $A\mathbf{v}_i = \mathbf{y}_i$ per $i = 1, \ldots, r$. Distinguiamo due casi:

a) Se $N(A) \neq \{\mathbf{0}\}$, esiste una base $\{\mathbf{u}_1, \ldots, \mathbf{u}_s\}$ di $N(A)$ con $s \geq 1$. I vettori $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{u}_1, \ldots, \mathbf{u}_s$ di $\mathbf{R}^n$ sono linearmente indipendenti. Sia infatti

$$\mathbf{z} = \sum_{i=1}^{r} \alpha_i \mathbf{v}_i + \sum_{i=1}^{s} \beta_i \mathbf{u}_i = \mathbf{0}$$

una loro combinazione nulla, risulta

$$\mathbf{0} = A\Big(\sum_{i=1}^{r} \alpha_i \mathbf{v}_i + \sum_{i=1}^{s} \beta_i \mathbf{u}_i\Big) = \sum_{i=1}^{r} \alpha_i A\mathbf{v}_i + \sum_{i=1}^{s} \beta_i A\mathbf{u}_i = \sum_{i=1}^{r} \alpha_i \mathbf{y}_i,$$

da cui segue che $\alpha_i = 0$ per $i = 1, \ldots, r$ perché gli $\mathbf{y}_i$ per $i = 1, \ldots, r$ formano una base. Resta che

$$\mathbf{z} = \sum_{i=1}^{s} \beta_i \mathbf{u}_i = \mathbf{0},$$

ma anche qui risulta $\beta_i = 0$ per $i = 1, \ldots, s$ perché gli $\mathbf{u}_i$ per $i = 1, \ldots, s$ formano una base. Ne segue che $r + s \leq n$. Vediamo adesso che i vettori

$\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{u}_1, \ldots, \mathbf{u}_s$ formano una base di $\mathbf{R}^n$. Per ogni $\mathbf{v} \in \mathbf{R}^n$ è $A\mathbf{v} \in S(A)$, per cui esistono $\gamma_1, \ldots, \gamma_r$ tali che

$$A\mathbf{v} = \sum_{i=1}^{r} \gamma_i \mathbf{y}_i = \sum_{i=1}^{r} \gamma_i A\mathbf{v}_i = A \sum_{i=1}^{r} \gamma_i \mathbf{v}_i,$$

da cui

$$A\Big(\mathbf{v} - \sum_{i=1}^{r} \gamma_i \mathbf{v}_i\Big) = \mathbf{0}.$$

Quindi $\mathbf{v} - \sum\limits_{i=1}^{r} \gamma_i \mathbf{v}_i \in N(A)$ ed esistono $\delta_1, \ldots, \delta_s$ tali che

$$\mathbf{v} - \sum_{i=1}^{r} \gamma_i \mathbf{v}_i = \sum_{i=1}^{s} \delta_i \mathbf{u}_i, \quad \text{cioè} \quad \mathbf{v} = \sum_{i=1}^{r} \gamma_i \mathbf{v}_i + \sum_{i=1}^{s} \delta_i \mathbf{u}_i.$$

Perciò $n = r + s$.

b) Se $N(A) = \{\mathbf{0}\}$, cioè $\dim(N(A)) = 0$, si segue lo stesso procedimento del caso a) tralasciando i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_s$. Si dimostra quindi che ogni $\mathbf{v} \in \mathbf{R}^n$ è combinazione lineare di $\mathbf{v}_1, \ldots, \mathbf{v}_r$ e quindi che $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ è una base di $\mathbf{R}^n$ e che $n = r$. ■

Se si applica il teorema 2.14 alla matrice trasposta $A^T \in \mathbf{R}^{n\times m}$ si ricava la relazione

$$\dim S(A^T) + \dim N(A^T) = m. \tag{9}$$

Consideriamo ad esempio

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

Dal teorema 2.13, per determinare $\dim S(A) = \mathrm{rk}(A)$ basta calcolare quante sono le colonne linearmente indipendenti di A applicando al sistema lineare $A\mathbf{x} = \mathbf{0}$ il procedimento di riduzione a forma a scalini anticipato nell'esercizio 5 del Capitolo 1, e che verrà descritto più adeguatamente nel seguito. Si ottengono dal sistema iniziale

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

i sistemi

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

e

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

da cui risulta $\text{rk}(A) = 2$, poiché 2 sono le equazioni del sistema finale con coefficienti non tutti nulli.

Dal teorema 2.14 si ha immediatamente $\dim N(A) = 4 - \dim S(A) = 2$.

In modo analogo si verifica che $\dim S(A^T) = 2$ e, dalla (9), $\dim N(A^T) = 1$.

2.15 Teorema.

$$S(A)^\perp = N(A^T).$$

Dim. Sia $\mathbf{x} \in N(A^T)$, quindi $A^T\mathbf{x} = \mathbf{0}$. Per ogni $\mathbf{y} \in S(A)$ è $\mathbf{y} = A\mathbf{z}$ per qualche $\mathbf{z} \in \mathbf{R}^n$, e quindi

$$\mathbf{x}^T\mathbf{y} = \mathbf{x}^T(A\mathbf{z}) = (\mathbf{x}^T A)\mathbf{z} = (A^T\mathbf{x})^T\mathbf{z} = \mathbf{z}^T(A^T\mathbf{x}) = \mathbf{0}.$$

Quindi $\mathbf{x}$ risulta ortogonale a tutti gli $\mathbf{y} \in S(A)$, per cui $\mathbf{x} \in S(A)^\perp$. Ne segue che

$$N(A^T) \subset S(A)^\perp. \tag{10}$$

Viceversa, sia $\mathbf{x} \in S(A)^\perp$. Allora $\mathbf{x}^T(A\mathbf{z}) = \mathbf{0}$ per ogni $\mathbf{z} \in \mathbf{R}^n$. Ma $\mathbf{x}^T(A\mathbf{z}) = (A^T\mathbf{x})^T\mathbf{z} = \mathbf{0}$ per ogni $\mathbf{z} \in \mathbf{R}^n$, cioè $A^T\mathbf{x} \in (\mathbf{R}^n)^\perp = \{\mathbf{0}\}$, quindi $\mathbf{x} \in N(A^T)$. Ne segue che

$$S(A)^\perp \subset N(A^T). \tag{11}$$

Da (10) e (11) segue la tesi. ■

Poiché $S(A)$ e $N(A^T)$ sono entrambi sottospazi di $\mathbf{R}^m$, dal teorema 1.16 segue che

$$S(A) \oplus N(A^T) = \mathbf{R}^m,$$

da cui

$$\dim S(A) + \dim N(A^T) = m. \tag{12}$$

Dal confronto fra la (9) e la (12) si ricava che

$$\dim S(A) = \dim S(A^T). \tag{13}$$

Per il teorema 2.13 $\dim S(A)$ dà il numero delle colonne linearmente indipendenti di A e $\dim S(A^T)$ dà il numero delle colonne linearmente indipendenti di A^T, cioè delle righe linearmente indipendenti di A. Dalla (13) segue che in una matrice il numero delle colonne linearmente indipendenti, e quindi il rango, coincide con il numero delle righe linearmente indipendenti.

2.16 Teorema.

Una matrice $A \in \mathbf{R}^{n\times n}$ è invertibile se e solo se $N(A) = \{\mathbf{0}\}$.

Dim. Sia $N(A) = \{\mathbf{0}\}$. Dal teorema 2.14 segue che $\dim S(A) = n$, quindi $S(A) = \mathbf{R}^n$ e per ogni $\mathbf{y} \in \mathbf{R}^n$ esiste $\mathbf{x} \in \mathbf{R}^n$ tale che $A\mathbf{x} = \mathbf{y}$. Verifichiamo che $\mathbf{x}$ è unico. Se esistesse $\overline{\mathbf{x}} \neq \mathbf{x}$ tale che $\mathbf{y} = A\overline{\mathbf{x}}$, sarebbe $A(\mathbf{x} - \overline{\mathbf{x}}) = \mathbf{0}$, quindi $\mathbf{x} - \overline{\mathbf{x}} \in N(A)$, cioè $\mathbf{x} - \overline{\mathbf{x}} = \mathbf{0}$. Si può allora definire un'applicazione $g : \mathbf{R}^n \to \mathbf{R}^n$ tale che $g(\mathbf{y}) = \mathbf{x}$; è facile verificare che g è lineare. Sia $B \in \mathbf{R}^{n\times n}$ la matrice associata a g. Risulta $B\mathbf{y} = \mathbf{x}$, quindi

$$\mathbf{y} = A\mathbf{x} = AB\mathbf{y} \quad \text{e} \quad \mathbf{x} = B\mathbf{y} = BA\mathbf{x} \quad \text{per ogni} \quad \mathbf{x}, \mathbf{y} \in \mathbf{R}^n.$$

Perciò $AB = BA = I$ e $B = A^{-1}$. Viceversa, supponiamo che A^{-1} esista. Per ogni $\mathbf{x} \in N(A)$ è $A\mathbf{x} = \mathbf{0}$, da cui segue che $\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$ e quindi la tesi. ■

Dai teoremi 2.14 e 2.16 segue che una matrice A è invertibile se e solo se $\mathrm{rk}(A) = n$, cioè se e solo se A ha n righe (e colonne) linearmente indipendenti.

2.17 Cambiamento di base.

Date due basi diverse $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ di $\mathbf{R}^n$, si chiama matrice del *cambiamento di base* la matrice P associata all'applicazione identica $i_{\mathbf{R}^n} : \mathbf{R}^n \to \mathbf{R}^n$, dove su $\mathbf{R}^n$ come spazio di definizione si considera la base $\{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ e su $\mathbf{R}^n$ come spazio di arrivo si considera la base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Gli elementi p_{ij} di P sono tali che

$$\mathbf{v}'_i = i_{\mathbf{R}^n}(\mathbf{v}'_i) = \sum_{j=1}^{n} p_{ji}\mathbf{v}_j, \quad i = 1, \dots, n.$$

Se le due basi coincidono, cioè $\mathbf{v}'_i = \mathbf{v}_i$ per $i = 1, \dots, n$, risulta $P = I$.

Ad esempio, consideriamo in $\mathbf{R}^3$ le due basi $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ e $\{\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}'_3\}$, dove

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \tag{14}$$

$$\mathbf{v}_1' = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3' = \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}. \tag{15}$$

Poiché, come è facile verificare, è

$$\mathbf{v}_1' = \mathbf{v}_1 - \mathbf{v}_2, \quad \mathbf{v}_2' = \mathbf{v}_2 - \mathbf{v}_3, \quad \mathbf{v}_3' = 2\mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3,$$

la matrice del cambiamento di base è

$$P = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 0 & -1 & -2 \end{bmatrix}. \tag{16}$$

La matrice associata ad una applicazione $f : \mathbf{R}^n \to \mathbf{R}^m$ dipende dalle basi scelte in $\mathbf{R}^n$ e $\mathbf{R}^m$. Sia A la matrice associata all'applicazione f quando in $\mathbf{R}^n$ si considera la base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e in $\mathbf{R}^m$ si considera la base $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ Vogliamo adesso vedere come cambia la matrice associata all'applicazione quando si cambiano la base di $\mathbf{R}^n$ in $\{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$ e la base di $\mathbf{R}^m$ in $\{\mathbf{w}_1', \ldots, \mathbf{w}_m'\}$. Siano P la matrice del cambiamento di base in $\mathbf{R}^n$, associata a $i_{\mathbf{R}^n}$, e Q la matrice del cambiamento di base in $\mathbf{R}^m$, associata a $i_{\mathbf{R}^m}$. Poiché

$$f(\mathbf{v}) = f(i_{\mathbf{R}^n}(\mathbf{v})) = {i_{\mathbf{R}^m}}^{-1}(f(i_{\mathbf{R}^n}(\mathbf{v}))) = ({i_{\mathbf{R}^m}}^{-1} \circ f \circ i_{\mathbf{R}^n})(\mathbf{v}),$$

per la (6) la matrice B associata all'applicazione f quando in $\mathbf{R}^n$ si considera la base $\{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$ e in $\mathbf{R}^m$ si considera la base $\{\mathbf{w}_1', \ldots, \mathbf{w}_m'\}$ è così legata alla matrice A

$$B = Q^{-1}AP, \quad \text{ovvero} \quad A = QBP^{-1}. \tag{17}$$

Se $n = m$ e $\mathbf{v}_i = \mathbf{w}_i$ e $\mathbf{v}_i' = \mathbf{w}_i'$ per $i = 1, \ldots, n$, allora $P = Q$ e

$$B = P^{-1}AP, \quad \text{ovvero} \quad A = PBP^{-1}. \tag{18}$$

Ad esempio, consideriamo l'applicazione $f : \mathbf{R}^3 \to \mathbf{R}^2$ definita sulla base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ di $\mathbf{R}^3$ data in (14) nel modo seguente

$$f(\mathbf{v}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad f(\mathbf{v}_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad f(\mathbf{v}_3) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Se in $\mathbf{R}^2$ fissiamo la base canonica

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

è

$$f(\mathbf{v}_1) = \mathbf{w}_1, \quad f(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2, \quad f(\mathbf{v}_3) = \mathbf{w}_2,$$

e la matrice associata ad f risulta

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

Cambiamo adesso le basi. In $\mathbf{R}^3$ scegliamo la base $\{\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}'_3\}$ data in (10) e $\mathbf{R}^2$ la base $\{\mathbf{w}'_1, \mathbf{w}'_2\}$, dove

$$\mathbf{w}'_1 = 2\mathbf{w}_1 - \mathbf{w}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{w}'_2 = 2\mathbf{w}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Le matrici dei cambiamenti di base sono per $\mathbf{R}^3$ la P data in (11) e per $\mathbf{R}^2$

$$Q = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}$$

(Q è la stessa matrice chiamata A in (7)). Rispetto alle nuove basi all'applicazione f risulta associata la matrice $B = Q^{-1}AP$ e per la (8) è

$$B = \begin{bmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 0 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 3/2 \\ -1/2 & 1/4 & 1/4 \end{bmatrix}.$$

Quindi sulle nuove basi l'applicazione f è definita da

$$f(\mathbf{v}'_1) = -\frac{1}{2}\,\mathbf{w}'_2, \quad f(\mathbf{v}'_2) = \frac{1}{2}\,\mathbf{w}'_1 + \frac{1}{4}\,\mathbf{w}'_2, \quad f(\mathbf{v}'_3) = \frac{3}{2}\,\mathbf{w}'_1 + \frac{1}{4}\,\mathbf{w}'_2.$$

2.18 Esercizi.

- **1.** Siano V e W due spazi vettoriali. Verificare che l'insieme $\mathcal{L}$ delle applicazioni lineari da V a W è uno spazio vettoriale rispetto alle usuali operazioni di addizione di applicazioni e di moltiplicazione per scalare.

- **2.** Dire se le seguenti applicazioni da $\mathbf{R}^n$ a $\mathbf{R}^m$ sono lineari e, in caso affermativo, definire l'applicazione sui vettori della base canonica di $\mathbf{R}^n$ e scrivere la matrice associata:

  a) per $n = 2$ e $m = 2$ $\quad f\left(\begin{bmatrix} x_1, \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 1 \\ x_2 + 1 \end{bmatrix}$,

  b) per $n = 2$ e $m = 3$ $\quad f\left(\begin{bmatrix} x_1, \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \\ x_2 \end{bmatrix}$,

  c) per $n = 2$ e $m = 3$ $\quad f\left(\begin{bmatrix} x_1, \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 1 \\ x_2 \end{bmatrix}$,

  d) per $n = 3$ e $m = 2$ $\quad f\left(\begin{bmatrix} x_1, \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_3 \\ 2x_2 \end{bmatrix}$.

- **3.** Dare agli elementi della matrice

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

  dei valori numerici reali e scrivere l'applicazione lineare a cui è associata la matrice A rispetto alla base canonica o rispetto ad altre basi.

- **4.** Dare una interpretazione geometrica dell'applicazione lineare a cui è associata la matrice

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

  per $\theta \in \mathbf{R}$.

- **5.** Verificare che l'insieme di $\mathbf{R}^{m\times n}$ delle matrici aventi m righe e n colonne è uno spazio vettoriale rispetto alle usuali operazioni di addizione di matrici e di moltiplicazione per scalare.

6. Un sottoinsieme di $\mathbf{R}^{n\times n}$ si dice *chiuso rispetto all'operazione di moltiplicazione*, se date due matrici A e B appartenenti al sottoinsieme, anche il prodotto AB appartiene al sottoinsieme. Dire quali dei seguenti sottoinsiemi di $\mathbf{R}^{n\times n}$ sono chiusi rispetto all'operazione di moltiplicazione:

• a) matrici scalari,

•• b) matrici triangolari superiori (inferiori),

•• c) matrici triangolari superiori (inferiori) in senso stretto

•• d) matrici tridiagonali.

• **7.** Calcolare i prodotti AB e BA nei seguenti casi:

a) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 1 \\ -4 & 3 \end{bmatrix},$

b) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 \\ 3 & 2 \end{bmatrix},$

c) $A = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ -4 & -2 \\ 0 & 2 \end{bmatrix}.$

•• **8.** Costruire due matrici A e $B \in \mathbf{R}^{2\times 2}$ tali che

$$AB \neq BA \neq A^TB \neq AB^T \neq A^TB^T \neq B^TA^T.$$

•• **9.** Data

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

costruire A^n per $n \geq 1$.

10. Si dice che due matrici A e B *commutano* se $AB = BA$.

• a) Costruire due matrici A e B di ordine 2 che commutano;

• b) dimostrare che le matrici diagonali commutano;

•• c) dimostrare che la relazione

$$(A+B)\,(A-B) = A^2 - B^2$$

vale se e solo se A e B commutano;

• d) dimostrare che per ogni A le due matrici $B = A^i$ e $C = A^j$, i, j interi positivi, commutano;

•• e) verificare che le sole matrici di $\mathbf{R}^{n\times n}$ che commutano con ogni matrice di $\mathbf{R}^{n\times n}$ sono le matrici scalari.

•• **11.** Verificare che se A e B commutano, anche A^T e B^T commutano e A^{-1} e B^{-1}, se esistono, commutano.

•• **12.** Dire se i seguenti sottoinsiemi di $\mathbf{R}^{n\times n}$ sono sottospazi vettoriali:

a) matrici tali che $A^2 = A$,

b) matrici che commutano con una matrice quadrata T,

c) matrici tridiagonali.

•• **13.** Sia $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Calcolare A^2 e determinare due scalari α e β tali che $\alpha A^2 + \beta A = I$. Da questa relazione ricavare A^{-1}.

•• **14.** Verificare che valgono le proprietà

$$(A^T)^T = A, \quad (A+B)^T = A^T + B^T,$$

$$(AB)^T = B^T A^T, \qquad (A^T)^{-1} = (A^{-1})^T.$$

• **15.** Sia $A \in \mathbf{R}^{m\times n}$. Verificare che $S(A)$ è un sottospazio di $\mathbf{R}^n$ e che $N(A)$ è un sottospazio di $\mathbf{R}^m$.

•• **16.** Siano $A \in \mathbf{R}^{n\times n}$ e $\mathbf{x}_1 \in \mathbf{R}^n$. Considerare la successione di vettori

$$\mathbf{x}_i = A\mathbf{x}_{i-1}, \quad i = 2, 3, \ldots$$

Se per un k è $\mathbf{x}_k \neq \mathbf{0}$ e $\mathbf{x}_{k+1} = \mathbf{0}$, allora i vettori $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ sono linearmente indipendenti.

(Traccia: considerare $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0}$ e moltiplicare per $A^{k-1}, A^{k-2}, \ldots, A$. Risulta che $\alpha_i = 0$ per $i = 1, \ldots, k$.)

••• **17.** Siano A e $B \in \mathbf{R}^{n\times n}$. Dimostrare le seguenti relazioni:

a) $\mathrm{rk}(A+B) \ \leq \mathrm{rk}A + \mathrm{rk}B$;

b) $\mathrm{rk}A + \mathrm{rk}B - n \ \leq \mathrm{rk}(AB) \leq \min\,\{\mathrm{rk}A\,,\, \mathrm{rk}B\}$.

(Traccia: b) per la prima disuguaglianza, notare che $\dim N(AB) \leq \dim N(A) + \dim N(B)$ e sfruttare il teorema 2.14, per la seconda, notare che le colonne (risp. le righe) di AB sono combinazioni lineari delle colonne di A (risp. delle righe di B).)

••• **18.** Sia $A \in \mathbf{R}^{n\times n}$. Verificare che

$$N(A^T A) = N(A), \quad S(A^T A) = S(A^T).$$

(Traccia: se $A\mathbf{x} = \mathbf{0}$, allora $A^T A\mathbf{x} = \mathbf{0}$ e quindi $N(A) \subset N(A^T A)$; viceversa se $A^T A\mathbf{x} = \mathbf{0}$ e per assurdo $A\mathbf{x} \neq \mathbf{0}$, il vettore $\mathbf{y} = A\mathbf{x}$ è tale che $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y} \in S(A)$ e $\mathbf{y} \in N(A^T)$, ma ciò non è possibile perché $N(A^T) = S(A)^\perp$. La seconda relazione deriva dalla prima tenendo conto che $N(A) = S(A^T)^\perp$.)

••• **19.** Siano $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in \mathbf{R}^n$. Determinare nucleo e rango delle due matrici

$$A = \mathbf{x}\mathbf{y}^T, \quad B = \mathbf{x}\mathbf{y}^T + \mathbf{u}\mathbf{v}^T.$$

•• **20.** In $\mathbf{R}^3$ si considerano le basi $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ e $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, dove

$$\mathbf{u}_1 = [1,0,1]^T, \quad \mathbf{u}_2 = [0,1,1]^T, \quad \mathbf{u}_3 = [1,1,0]^T.$$

a) Scrivere la matrice P che fa passare dalla prima alla seconda base e la matrice Q che fa passare dalla seconda alla prima.

b) Verificare che $P = Q^{-1}$.

c) Sia

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

la matrice associata all'applicazione $f : \mathbf{R}^3 \to \mathbf{R}^3$ quando la base considerata in $\mathbf{R}^3$ è quella canonica. Scrivere la matrice B associata alla stessa applicazione quando la base di $\mathbf{R}^3$ viene cambiata nella $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

•• **21.** Data una matrice A reale quadrata di dimensione 2 e un vettore $\mathbf{x} \in \mathbf{R}^2$ si consideri lo scalare $\alpha = \mathbf{x}^T A \mathbf{x}$. Verificare che esistono matrici non nulle A tali che $\alpha = 0$ per ogni scelta di $\mathbf{x}$.

• **22.** Assegnato un vettore $\mathbf{v} = [v_1, v_2]^T \in \mathbf{R}^2$, si consideri l'applicazione $f : \mathbf{R}^2 \to \mathbf{R}^2$ così definita:

$$f(\mathbf{x}) = \mathbf{v}\mathbf{x}^T\mathbf{v}.$$

Dire se f è un'applicazione lineare, e in tal caso determinare la matrice ad essa associata se su $\mathbf{R}^2$ si sceglie la base canonica.

•• **23.** Si consideri l'applicazione $f : \mathbf{R}^4 \to \mathbf{R}^4$ così definita:

$$f(\mathbf{x}) = \begin{bmatrix} x_1 + x_2 + 2x_3 + 2x_4 \\ x_1 + 2x_3 + x_4 \\ x_1 + x_2 + 2x_3 + 2x_4 \\ x_1 + x_3 \end{bmatrix},$$

per ogni $\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4]^T \in \mathbf{R}^4$. Si dica se f è lineare e, in tal caso, si determini la matrice ad essa associata quando in $\mathbf{R}^4$ si considera la base canonica. Infine si determini la dimensione dell'immagine di f.