APPLICAZIONI LINEARI

VINCENZO DI GENNARO

1. Le coordinate.

Sia V uno spazio vettoriale,[1] e fissiamo una base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ di V, quindi $\dim(V) = n$. Sappiamo che per ogni vettore $\mathbf{u}$ di V esiste un'unica $n-$pla di pesi $\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^n$ tali che

$$\mathbf{u} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n.$$

Chiameremo tale vettore numerico $\mathbf{x}$ il vettore delle coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, e scriveremo

$$\mathbf{x} =: [\mathbf{u}]_{\mathcal{B}}.$$

Nel seguito, per una questione formale che sara' chiara tra poco, si intendera' il vettore $\mathbf{x}$ messo in colonna, cioe' intenderemo

$$\mathbf{x} = (x_1, x_2, .., x_n)^T = \begin{bmatrix} x_1 \\ x_2 \\ : \\ x_n \end{bmatrix} \in \mathcal{M}(n, 1).$$

Naturalmente possiamo identificare $\mathbf{R}^n$ con $\mathcal{M}(n, 1)$ quando cio' non dia luogo a confusione.

Esempi.

1) Se $V = \mathbf{R}^n$, $\mathbf{u} = (u_1, u_2, \dots, u_n) \in \mathbf{R}^n$ e' un vettore numerico ed $\mathcal{E}$ e' la base canonica di $\mathbf{R}^n$, allora le coordinate di $\mathbf{u}$ rispetto ad $\mathcal{E}$ sono proprio le sue componenti messe in colonna: $[\mathbf{u}]_{\mathcal{E}} = (u_1, u_2, \dots, u_n)^T$.

2) Il seguente esempio mostra che le coordinate dipendono anche dall'ordine con cui sono disposti i vettori nella base $\mathcal{B}$. Nello spazio vettoriale V si considerino le seguenti basi: $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ e $\mathcal{C} = \{\mathbf{b}_3, \mathbf{b}_2, \mathbf{b}_1\}$. Allora le coordinate di $\mathbf{b}_1$ rispetto alla base $\mathcal{B}$ sono $(1, 0, 0)$ in quanto

$$\mathbf{b}_1 = 1 \cdot \mathbf{b}_1 + 0 \cdot \mathbf{b}_2 + 0 \cdot \mathbf{b}_3.$$

Invece le coordinate di $\mathbf{b}_1$ rispetto alla base $\mathcal{C}$ sono $(0, 0, 1)$ in quanto

$$\mathbf{b}_1 = 0 \cdot \mathbf{b}_3 + 0 \cdot \mathbf{b}_2 + 1 \cdot \mathbf{b}_1.$$

Riassumendo:

$$[\mathbf{b}_1]_{\mathcal{B}} = (1, 0, 0)^T, \quad [\mathbf{b}_1]_{\mathcal{C}} = (0, 0, 1)^T.$$

3) Se $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ e' una base dello spazio V, allora $[\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_i$, dove con $\mathbf{e}_i$ denotiamo il vettore canonico di $\mathbf{R}^n$ di posto i (messo in colonna).

[1] Ultimo aggiornamento: 17 novembre 2020

4) Nel secondo esempio abbiamo visto che le coordinate di un dato vettore possono cambiare al variare della base rispetto a cui si considerano le sue coordinate. Cioe' se $\mathbf{u} \in V$ e' un vettore, $\mathcal{B}$ e $\mathcal{C}$ sono basi di V, in generale $[\mathbf{u}]_{\mathcal{B}} \neq [\mathbf{u}]_{\mathcal{C}}$. Sorge naturale la domanda se c'e' comunque un nesso tra i vettori numerici $[\mathbf{u}]_{\mathcal{B}}$ e $[\mathbf{u}]_{\mathcal{C}}$. La risposta e' affermativa, in quanto c'e' una matrice P, che dipende solo dalle due basi, detta la matrice del cambiamento delle coordinate, che governa il passaggio da $[\mathbf{u}]_{\mathcal{B}}$ a $[\mathbf{u}]_{\mathcal{C}}$, cioe' e' tale che $[\mathbf{u}]_{\mathcal{C}} = P \cdot [\mathbf{u}]_{\mathcal{B}}$. Prima di dare la definizione di tale matrice, vediamo un altro esempio. Mettiamoci in $\mathbf{R}^2$, e consideriamo le basi $\mathcal{E} = \{(1,0),(0,1)\}$ e $\mathcal{B} = \{(1,1),(1,2)\}$. Sia $\mathbf{u} = (u_1, u_2) \in \mathbf{R}^2$ un qualunque vettore. Allora abbiamo:

$$[\mathbf{u}]_{\mathcal{E}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad e \quad [\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} 2u_1 - u_2 \\ -u_1 + u_2 \end{bmatrix}.$$

Quindi, come osservato, in generale $[\mathbf{u}]_{\mathcal{E}} \neq [\mathbf{u}]_{\mathcal{B}}$, pero' i due vettori delle coordinate si possono esprimere uno in funzione dell'altro tramite una matrice:

$$\begin{bmatrix} 2u_1 - u_2 \\ -u_1 + u_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Tale matrice

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

per la quale

$$[\mathbf{u}]_{\mathcal{B}} = P \cdot [\mathbf{u}]_{\mathcal{E}}$$

e' la matrice del cambiamento delle coordinate, dalla base $\mathcal{E}$ alla base $\mathcal{B}$, che ora andremo a studiare.

2. La matrice del cambiamento delle coordinate.

• *La definizione della matrice del cambiamento delle coordinate.*

Siano $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ e $\mathcal{B}' = \{\mathbf{b}'_1, \dots, \mathbf{b}'_n\}$ basi per un dato spazio vettoriale V. Si definisce matrice del cambiamento delle coordinate dalla base $\mathcal{B}'$ alla base $\mathcal{B}$ quella matrice P, $n \times n$, la cui j-esima colonna e' data dal vettore numerico $[\mathbf{b}'_j]_{\mathcal{B}}$ delle coordinate di $\mathbf{b}'_j$ rispetto alla base $\mathcal{B}$. Denoteremo tale matrice con il simbolo

$$M_{\mathcal{B}}^{\mathcal{B}'}(id_V) := P.$$

Esempi.

1) Si considerino le seguenti basi di $\mathbf{R}^2$:

$$\mathcal{E} = \{(1,0),(0,1)\} \quad e \quad \mathcal{B} = \{(1,1),(1,2)\}.$$

Poiche'

$$(1,0) = 2 \cdot (1,1) + (-1) \cdot (1,2) \quad e \quad (0,1) = (-1) \cdot (1,1) + 1 \cdot (1,2)$$

allora

$$M_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2}) = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Invece, poiche'

$$(1,1) = 1 \cdot (1,0) + 1 \cdot (0,1) \quad e \quad (1,2) = 1 \cdot (1,0) + 2 \cdot (0,1)$$

allora

$$M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Si osservi che $M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2})$ e' l'inversa di $M_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2})$.

2) Come si evince dall'esempio precedente, se $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ e' una base di $\mathbf{R}^n$ allora la matrice $M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^n})$ del cambiamento delle coordinate dalla base $\mathcal{B}$ alla base canonica $\mathcal{E}$ e' la matrice le cui colonne sono date dai vettori $\mathbf{b}_i$.

3) Sia $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ una base dello spazio V. Allora:

$$M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V) = I.$$

4) Si considerino le basi $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ e $\mathcal{B}' = \{\mathbf{b}_2, \mathbf{b}_1\}$ dello spazio V. Allora

$$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_{\mathbf{V}}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

• *Le proprieta' della matrice del cambiamento delle coordinate.*

1) *Per ogni vettore $\mathbf{u} \in V$ si ha:*

$$[\mathbf{u}]_{\mathcal{B}'} = M_{\mathcal{B}'}^{\mathcal{B}}(id_V) \cdot [\mathbf{u}]_{\mathcal{B}}.$$

2) *La matrice $M_{\mathcal{B}'}^{\mathcal{B}}(id_V)$ e' invertibile e*

$$M_{\mathcal{B}'}^{\mathcal{B}}(id_V)^{-1} = M_{\mathcal{B}}^{\mathcal{B}'}(id_V).$$

3) *Se $\mathcal{B}$, $\mathcal{B}'$, $\mathcal{B}''$ sono tre basi di V allora:*

$$M_{\mathcal{B}}^{\mathcal{B}''}(id_V) = M_{\mathcal{B}}^{\mathcal{B}'}(id_V) \cdot M_{\mathcal{B}'}^{\mathcal{B}''}(id_V).$$

Dimostrazione. 1) Questa proprieta' ci dice che la matrice del cambiamento delle coordinate $M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)$, che dipende solo dalle basi $\mathcal{B}$ e $\mathcal{B}'$, ci consente di calcolare le coordinate rispetto alla base $\mathcal{B}'$ a partire dalle coordinate rispetto alla base $\mathcal{B}$. Per dimostrare cio', innanzitutto semplifichiamo le notazioni, ponendo

$$\mathbf{x}' := [\mathbf{u}]_{\mathcal{B}'}, \quad \mathbf{x} := [\mathbf{u}]_{\mathcal{B}}, \quad P := M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V).$$

Quindi occorre provare che:

$$\mathbf{x}' = P \cdot \mathbf{x}.$$

Abbiamo

$$\mathbf{x}' = [\mathbf{u}]_{\mathcal{B}'} = [x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n]_{\mathcal{B}'}.$$

Per una proprieta' che dimostreremo nel prossimo paragrafo, quando studieremo la cosiddetta "applicazione delle coordinate", sappiamo che

$$[x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n]_{\mathcal{B}'} = x_1[\mathbf{b}_1]_{\mathcal{B}'} + \dots + x_n[\mathbf{b}_n]_{\mathcal{B}'}.$$

D'altra parte, per definizione di matrice del cambiamento delle coordinate, possiamo scrivere:

$$x_1[\mathbf{b}_1]_{\mathcal{B}'} + \dots + x_n[\mathbf{b}_n]_{\mathcal{B}'} = x_1P^1 + \dots + x_nP^n$$

dove con P^i denotiamo la colonna di posto i della matrice P. Tenuto conto che

$$P^i = P \cdot \mathbf{e}_i,$$

mettendo tutto insieme abbiamo:[2]

$$\mathbf{x}' = x_1(P \cdot \mathbf{e}_1) + \cdots + x_n(P \cdot \mathbf{e}_n).$$

Per la proprieta' distributiva del prodotto tra matrici sappiamo anche che

$$x_1(P \cdot \mathbf{e}_1) + \cdots + x_n(P \cdot \mathbf{e}_n) = P(x_1 \cdot \mathbf{e}_1 + \cdots + x_n \cdot \mathbf{e}_n) = P \cdot \mathbf{x}.$$

Quindi $\mathbf{x}' = P \cdot \mathbf{x}$, che e' quello che volevamo provare.

2) Ora denotiamo con Q la matrice $M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V)$. Si tratta di dimostrare che:

$$QP = I.$$

Per la proprieta' 1) appena dimostrata sappiamo che per ogni $\mathbf{u} \in V$ si ha:

$$\mathbf{x}' = P\mathbf{x}, \quad \mathbf{x} = Q\mathbf{x}'.$$

Quindi per ogni $\mathbf{u} \in V$ abbiamo

$$\mathbf{x} = Q\mathbf{x}' = Q(P\mathbf{x}) = (QP)\mathbf{x}.$$

Poiche' il vettore $\mathbf{u}$ varia tra tutti i vettori di V, allora $\mathbf{x}$ varia tra tutti i vettori di $\mathbf{R}^n$. Questo vuol dire che la matrice QP soddisfa la condizione $(QP)\mathbf{x} = \mathbf{x}$ per ogni $\mathbf{x} \in \mathbf{R}^n$. Applicando tale uguaglianza ai vettori canonici di $\mathbf{R}^n$, ne consegue che QP ha le stesse colonne di I, percio' $QP = I$.

3) Poniamo:

$$Q = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V), \quad R = M_{\mathcal{B}'}^{\mathcal{B}''}(\mathrm{id}_V), \quad S = M_{\mathcal{B}}^{\mathcal{B}''}(\mathrm{id}_V).$$

Occorre provare che

$$S = QR.$$

Ora per la prima proprieta' sappiamo che per ogni $\mathbf{u} \in V$ si ha:

$$\mathbf{x} = Q\mathbf{x}', \quad \mathbf{x}' = R\mathbf{x}'', \quad \mathbf{x} = S\mathbf{x}'',$$

dove naturalmente denotiamo con $\mathbf{x}''$ le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}''$. Quindi per ogni $\mathbf{u} \in V$, cioe' per ogni $\mathbf{x}'' \in \mathbf{R}^n$, abbiamo:

$$S\mathbf{x}'' = \mathbf{x} = Q\mathbf{x}' = Q(R\mathbf{x}'') = (QR)\mathbf{x}''.$$

Da qui, come in precedenza, deduciamo che $S = QR$. □

Esercizio. Nello spazio $\mathbf{R}^2$ *si considerino le basi* $\mathcal{E} = \{(1,0),(0,1)\}$, $\mathcal{B} = \{(1,3),(2,7)\}$, *ed il vettore* $\mathbf{u} = (1,7)$. *Calcolare* $M_{\mathcal{E}}^{\mathcal{B}}(id_{\mathbf{R}^2})$, $M_{\mathcal{B}}^{\mathcal{E}}(id_{\mathbf{R}^2})$, $[\mathbf{u}]_{\mathcal{E}}$, $[\mathbf{u}]_{\mathcal{B}}$.

Svolgimento. Abbiamo gia' osservato che la matrice $M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2})$ consiste dei vettori della base $\mathcal{B}$ messi in colonna, cioe':

$$M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}.$$

Ne consegue

$$M_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2}) = M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2})^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}.$$

Sappiamo che

$$[\mathbf{u}]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}.$$

[2] Come al solito, con $\mathbf{e}_i$ denotiamo il vettore canonico di $\mathbf{R}^n$ di posto i messo in colonna.

Quindi

$$[\mathbf{u}]_{\mathcal{B}} = M_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2}) \cdot [\mathbf{u}]_{\mathcal{E}} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 4 \end{bmatrix}.$$

Possiamo verificare questo risultato in quanto:

$$(1,7) = (-7)(1,3) + 4(2,7). \quad \square$$

Esercizio. Nello spazio V si considerino le basi $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, $\mathcal{B}' = \{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$, con $\mathbf{b}'_1 = \mathbf{b}_1$, $\mathbf{b}'_2 = \mathbf{b}_2 + 5\mathbf{b}_3$, $\mathbf{b}'_3 = \mathbf{b}_2 + 6\mathbf{b}_3$. Si consideri anche il vettore $\mathbf{u} = -2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3$. Calcolare $M_{\mathcal{B}'}^{\mathcal{B}}(id_V)$, $M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$, $[\mathbf{u}]_{\mathcal{B}}$, $[\mathbf{u}]_{\mathcal{B}'}$.

Svolgimento. Per la stessa definizione di matrice del cambiamento delle coordinate abbiamo:

$$M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 5 & 6 \end{bmatrix}.$$

Percio' $M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)$, che e' l'inversa della matrice precedente, e':

$$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & -1 \\ 0 & -5 & 1 \end{bmatrix}.$$

Per definizione di coordinate abbiamo:

$$[\mathbf{u}]_{\mathcal{B}} = (-2, 3, -1)^T.$$

Quindi

$$[\mathbf{u}]_{\mathcal{B}'} = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V) \cdot [\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & -1 \\ 0 & -5 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 19 \\ -16 \end{bmatrix}. \quad \square$$

Esercizio. Nello spazio $\mathbf{R}^2$ si considerino le basi $\mathcal{B} = \{(2,2),(2,3)\}$, $\mathcal{B}' = \{(1,1),(5,6)\}$, ed il vettore $\mathbf{u}$ che ha coordinate $(5,-4)^T$ rispetto alla base $\mathcal{B}$. Calcolare le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}'$.

Svolgimento. Applicando la proprieta' 3) abbiamo:

$$[\mathbf{u}]_{\mathcal{B}'} = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) \cdot [\mathbf{u}]_{\mathcal{B}} = M_{\mathcal{B}'}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2}) \cdot M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) \cdot [\mathbf{u}]_{\mathcal{B}}$$

dove $\mathcal{E}$ denota la base canonica. Poiche'

$$M_{\mathcal{B}'}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^2}) = M_{\mathcal{E}}^{\mathcal{B}'}(\mathrm{id}_{\mathbf{R}^2})^{-1} = \begin{bmatrix} 1 & 5 \\ 1 & 6 \end{bmatrix}^{-1} = \begin{bmatrix} 6 & -5 \\ -1 & 1 \end{bmatrix}$$

e

$$M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix},$$

otteniamo:

$$[\mathbf{u}]_{\mathcal{B}'} = \begin{bmatrix} 6 & -5 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} 22 \\ -4 \end{bmatrix}. \quad \square$$

Esercizio. Nello spazio $\mathbf{R}^3$ si consideri la base

$$\mathcal{B} = \{(0,1,1),(0,1,2),(1,1,1)\},$$

ed un'altra base $\mathcal{B}' = \{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ incognita. Per ogni vettore $\mathbf{u}$ di $\mathbf{R}^3$, siano $\mathbf{x}$ le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, ed $\mathbf{x}'$ le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}'$. Sapendo che

$$\begin{cases} x'_1 = x_1 \\ x'_2 = 2x_2 + 5x_3 \\ x'_3 = x_2 + 3x_3 \end{cases}$$

determinare i vettori della base $\mathcal{B}'$.

Svolgimento. Il vettore $\mathbf{b}'_1$ ha coordinate $\mathbf{x}' = (1, 0, 0)$ rispetto alla base $\mathcal{B}'$. Percio' dalle formule assegnate sappiamo che le sue coordinate $\mathbf{x}$ rispetto alla base $\mathcal{B}$ si ottengono dalle condizioni:

$$\begin{cases} 1 = x_1 \\ 0 = 2x_2 + 5x_3 \\ 0 = x_2 + 3x_3. \end{cases}$$

Risolvendo tale sistema otteniamo $\mathbf{x} = [\mathbf{b}'_1]_{\mathcal{B}} = (1, 0, 0)^T$, cioe' il vettore $\mathbf{b}'_1$ coincide con $(0, 1, 1)$:

$$\mathbf{b}'_1 = (0, 1, 1).$$

Poiche' il vettore $\mathbf{b}'_2$ ha coordinate $\mathbf{x}' = (0, 1, 0)$ rispetto alla base $\mathcal{B}'$, deduciamo:

$$\begin{cases} 0 = x_1 \\ 1 = 2x_2 + 5x_3 \\ 0 = x_2 + 3x_3. \end{cases}$$

Da cui otteniamo $\mathbf{x} = [\mathbf{b}'_2]_{\mathcal{B}} = (0, 3, -1)^T$, quindi

$$\mathbf{b}'_2 = 0 \cdot (0, 1, 1) + 3 \cdot (0, 1, 2) + (-1)(1, 1, 1) = (-1, 2, 5).$$

Similmente si vede che

$$\mathbf{b}'_3 = 0 \cdot (0, 1, 1) + (-5) \cdot (0, 1, 2) + 2(1, 1, 1) = (2, -3, -8). \quad \square$$

3. L'applicazione delle coordinate.

Fissata la base $\mathcal{B}$ di V, possiamo considerare la seguente applicazione:

$$[\ ]_{\mathcal{B}} : \mathbf{u} \in V \to [\mathbf{u}]_{\mathcal{B}} \in \mathcal{M}(n, 1).$$

Tale applicazione si dice *l'applicazione delle coordinate di V rispetto alla base $\mathcal{B}$*. L'interesse dell'applicazione delle coordinate consiste nelle seguenti proprieta' che, come sara' piu' chiaro in seguito, consentono di identificare lo spazio V con lo spazio numerico $\mathbf{R}^n$. Ma l'applicazione delle coordinate ha anche un difetto, cioe' che dipende dalla base fissata. La matrice del cambiamento delle coordinate pone un parziale rimedio.

• *Le proprieta' dell'applicazione delle coordinate.*

1) *L'applicazione delle coordinate e' biiettiva.*

2) *Per ogni $\mathbf{u}, \mathbf{v} \in V$ si ha*

$$[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}.$$

3) *Per ogni scalare $a \in \mathbf{R}$ ed $\mathbf{u} \in V$ si ha*

$$[a\mathbf{u}]_{\mathcal{B}} = a[\mathbf{u}]_{\mathcal{B}}.$$

4) *Per ogni* $a, b \in \mathbf{R}$ *ed* $\mathbf{u}, \mathbf{v} \in V$ *si ha*

$$[a\mathbf{u} + b\mathbf{v}]_{\mathcal{B}} = a[\mathbf{u}]_{\mathcal{B}} + b[\mathbf{v}]_{\mathcal{B}}.$$

Dimostrazione. Prima di procedere nella dimostrazione osserviamo che la proprieta' 1) e' ovvia, e che la 4) segue da 2) e 3). Osserviamo anche che nella proprieta' 2) appaiono due segni di addizione. Quello a sinistra indica l'addizione nello spazio vettoriale V, mentre quello a destra indica l'addizione tra vettori numerici di $\mathbf{R}^n$. Quindi tale proprieta' mostra una compatibilita' tra queste due addizioni. Un simile commento si puo' fare anche per la proprieta' 3), in cui appaiono due segni di moltiplicazione esterna. Infine osserviamo che abbiamo usato la proprieta' 4) nel corso della dimostrazione della proprieta' 1) della matrice del cambiamento delle coordinate.

Per provare la 2), denotiamo con $\mathbf{x}$ e con $\mathbf{y}$ le coordinate di $\mathbf{u}$ e $\mathbf{v}$. Allora

$$\mathbf{u} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \quad \text{e} \quad \mathbf{v} = y_1\mathbf{b}_1 + \cdots + y_n\mathbf{b}_n.$$

Percio', per la proprieta' associativa, commutativa, e distributiva, abbiamo:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) + (y_1\mathbf{b}_1 + \cdots + y_n\mathbf{b}_n) \\ &= (x_1 + y_1)\mathbf{b}_1 + \cdots + (x_n + y_n)\mathbf{b}_n.\end{aligned}$$

Cio' vuol dire che le coordinate di $\mathbf{u}+\mathbf{v}$ sono $\mathbf{x}+\mathbf{y}$, cioe' la somma in $\mathbf{R}^n$ delle coordinate di $\mathbf{u}$ con quelle di $\mathbf{v}$, che e' quello che volevamo dimostrare.

La dimostrazione della proprieta' 3) e' analoga. Infatti

$$a\mathbf{u} = a(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = (ax_1)\mathbf{b}_1 + \cdots + (ax_n)\mathbf{b}_n.$$

Cioe' le coordinate di $a\mathbf{u}$ sono date dal vettore numerico $a\mathbf{x} = a[\mathbf{u}]_{\mathcal{B}}$. $\square$

La seguente conseguenza ci fa capire l'interesse dell'applicazione delle coordinate.

Corollario. Sia $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ *una base di* V*, siano* $\mathbf{u}_1, \dots, \mathbf{u}_h$ *vettori di* V*, e per ciscuno di tali vettori* $\mathbf{u}_i$ *denotiamone con* $\mathbf{x}_i = [\mathbf{u}_i]_{\mathcal{B}} \in \mathbf{R}^n$ *le coordinate rispetto alla base* $\mathcal{B}$*. Allora i vettori* $\mathbf{u}_1, \dots, \mathbf{u}_h$ *sono linearmente indipendenti in* V *se e solo se lo sono i vettori numerici* $\mathbf{x}_1, \dots, \mathbf{x}_h$ *in* $\mathbf{R}^n$*.*

Dimostrazione. In altre parole, lo studio dei sistemi liberi di uno spazio vettoriale qualsiasi si riconduce, tramite l'applicazione delle coordinate, allo studio dei sistemi liberi di vettori numerici, che come sappiamo si puo' effettuare tramite algoritmi. Questo e' un principio generale, che vale anche per lo studio di altre proprieta'.

Supponiamo innanzitutto che i vettori $\mathbf{u}_1, \dots, \mathbf{u}_h$ siano linearmente indipendenti in V, e sia $(a_1, \dots, a_h) \in \mathbf{R}^h$ una relazione per il sistema $\mathbf{x}_1, \dots, \mathbf{x}_h$. Allora:

$$\mathbf{0}_{\mathbf{R}^n} = a_1\mathbf{x}_1 + \cdots + a_h\mathbf{x}_h = a_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + a_h[\mathbf{u}_h]_{\mathcal{B}}.$$

Per la proprieta' 4) sappiamo che:

$$a_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + a_h[\mathbf{u}_h]_{\mathcal{B}} = [a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h]_{\mathcal{B}}.$$

Cioe' il vettore $a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h$ di V ha coordinate nulle, quindi e' il vettore nullo di V:

$$a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h = \mathbf{0}_V.$$

Poiche' per ipotesi $\mathbf{u}_1, \dots, \mathbf{u}_h$ sono liberi, allora la relazione $(a_1, \dots, a_h)$ deve essere banale. Cio' prova che ogni relazione tra i vettori $\mathbf{x}_1, \dots, \mathbf{x}_h$ deve essere banale, cioe' che tale sistema e' libero.

Viceversa, supponiamo che i vettori $\mathbf{x}_1, \dots, \mathbf{x}_h$ siano linearmente indipendenti in $\mathbf{R}^n$, e sia $(a_1, \dots, a_h) \in \mathbf{R}^h$ una relazione per il sistema $\mathbf{u}_1, \dots, \mathbf{u}_h$. Allora $a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h = \mathbf{0}_V$. Percio' tale vettore ha coordinate nulle, cioe'

$$\mathbf{0}_{\mathbf{R}^n} = [a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h]_{\mathcal{B}}.$$

Ma

$$[a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h]_{\mathcal{B}} = a_1[\mathbf{u}_1]_{\mathcal{B}} + \dots + a_h[\mathbf{u}_h]_{\mathcal{B}} = a_1\mathbf{x}_1 + \dots + a_h\mathbf{x}_h.$$

Percio' $(a_1, \dots, a_h) \in \mathbf{R}^h$ e' anche una relazione per i vettori $\mathbf{x}_1, \dots, \mathbf{x}_h$. Poiche' tali vettori sono liberi, allora la relazione $(a_1, \dots, a_h) \in \mathbf{R}^h$ e' banale. $\square$

Esercizio. Si considerino i seguenti polinomi:

$$p(t) = 1 + 2t + 3t^2 + t^3, \quad q(t) = 1 - 4t + 7t^2 - 9t^3, \quad r(t) = 1 + 5t + t^2 + 6t^3.$$

Dire se il sistema $S = \{p(t), q(t), r(t)\}$ e' libero in $\mathbf{R}[t]$ oppure no.

Svolgimento. Sia $1, t, t^2, t^3$ la base canonica di $\mathbf{R}[t]_{h \leq 3}$. Rispetto a tale base i polinomi assegnati hanno coordinate date dai vettori numerici di $\mathbf{R}^4$:

$$(1, 2, 3, 1), (1, -4, 7, -9), (1, 5, 1, 6).$$

L'algoritmo di Gauss prova che tale sistema e' legato. Percio' anche S lo e', dunque la risposta e' no. $\square$

Esercizio. Sia W il sottoinsieme di $\mathcal{M}(2,2)$ formato dalle matrici della forma

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

al variare di $a, b, c \in \mathbf{R}$. Provare che W e' un sottospazio di $\mathcal{M}(2,2)$, e calcolarne una base e la dimensione.

Svolgimento. Poiche'

$$\begin{bmatrix} a & b \\ -b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

si deduce che

$$W = \mathrm{Span}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

Cio' prova che W e' un sottospazio di $\mathcal{M}(2,2)$, generato dal sistema di matrici

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Tramite l'applicazione delle coordinate di $\mathcal{M}(2,2)$ rispetto alla base canonica di $\mathcal{M}(2,2)$, i tre generatori corrispondono ai vettori di $\mathbf{R}^4$:

$$(1, 0, 0, 0), (0, 1, -1, 0), (0, 0, 0, 1).$$

Tali vettori sono liberi in $\mathbf{R}^4$. Percio' i vettori di S sono liberi. Dunque S e' una base per W, e W ha dimensione 3. $\square$

Esempio. Sia $\mathcal{S} : A\mathbf{x} = \mathbf{0}$ un sistema lineare omogeneo di rango p, in n variabili, e sia

$$\rho : \mathbf{R}^{n-p} \to Sol(\mathcal{S})$$

la rappresentazione parametrica delle soluzioni di $\mathcal{S}$. Sia $\mathcal{B}$ la base di $Sol(\mathcal{S})$ corrispondente alla base canonica di $\mathbf{R}^{n-p}$. Allora la mappa inversa di ρ:

$$\rho^{-1} : Sol(\mathcal{S}) \to \mathbf{R}^{n-p}$$

altro non e' che l'applicazione delle coordinate di $Sol(\mathcal{S})$ rispetto alla base $\mathcal{B}$:

$$[\ ]_{\mathcal{B}} = \rho^{-1}.$$

Osservazione. (i) Nel caso in cui $V = \mathcal{V}_O$ sia lo spazio dei vettori geometrici, o un sottospazio di $\mathcal{V}_O$, l'applicazione delle coordinate mette in corrispondenza biiettiva i vettori geometrici con i vettori numerici.

(ii) Nel caso in cui $V = \mathbf{R}^n$ ed $\mathcal{E}$ e' la base canonica di $\mathbf{R}^n$, l'applicazione delle coordinate e' semplicemente quella applicazione che prende un vettore numerico $\mathbf{u} \in \mathbf{R}^n$ e lo mette in colonna.

4. Le applicazioni lineari.

In matematica il confronto tra oggetti diversi si effettua tramite lo studio delle applicazioni tra i due oggetti. Naturalmente, nel caso in cui gli oggetti siano due spazi vettoriali V e V', le applicazioni $V \to V'$ che sono "compatibili" rispetto alle operazioni di V e V' rivestono un grande interesse. Queste sono le applicazioni lineari, che generalizzano la nozione di applicazione delle coordinate.

• *La definizione di applicazione lineare.*

Sia $f : V \to V'$ un'applicazione tra due spazi vettoriali. La funzione f si dice lineare se soddisfa le seguenti condizioni. Per ogni scelta di vettori $\mathbf{u}$, $\mathbf{v}$ in V, ed ogni scalare $a \in \mathbf{R}$, si ha

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) \quad e \quad f(a \cdot \mathbf{u}) = a \cdot f(\mathbf{u}).$$

Osserviamo che nella condizione $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u})+f(\mathbf{v})$, il segno $+$ che appare a sinistra denota l'operazione di addizione in V, mentre quello a destra quello di V'. Cosi' come nella condizione $f(a\cdot\mathbf{u}) = a\cdot f(\mathbf{u})$, il segno $\cdot$ che appare a sinistra denota la moltiplicazione esterna di V, mentre a destra quella di V'. La letteratura non e' uniforme sul termine "applicazione lineare". In alcuni libri si usano anche termini del tipo "trasformazione lineare", "operatore lineare", "omomorfismo", "morfismo". Naturalmente l'applicazione delle coordinate e' un esempio di applicazione lineare. Prima di vedere altri esempi, conviene studiare le proprieta' di calcolo delle applicazioni lineari.

• *Proprieta' di calcolo per le applicazioni lineari.*

Sia $f : V \to V'$ un'applicazione lineare. Allora:

1) $f(\mathbf{0}_V) = \mathbf{0}_{V'}$.

2) *Per ogni* $\mathbf{u} \in V$ *si ha* $f(-\mathbf{u}) = -f(\mathbf{u})$.

3) *Per ogni* $\mathbf{u}, \mathbf{v} \in V$ *si ha* $f(\mathbf{u} - \mathbf{v}) = f(\mathbf{u}) - f(\mathbf{v})$.

4) *Per ogni* $\mathbf{u}, \mathbf{v} \in V$ *ed ogni* $a, b \in \mathbf{R}$ *si ha*

$$f(a\mathbf{u} + b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v}).$$

5) *Se f e' biiettiva anche f^{-1} e' lineare.*

6) *Se $f : V \to V'$ e $g : V' \to V''$ sono applicazioni lineari anche la funzione composta $g \circ f : V \to V''$ e' lineare.*

7) *Date due qualunque applicazioni lineari* $f : V \to V'$ *e* $g : V \to V'$, *e qualunque scalare* $c \in \mathbf{R}$, *le applicazioni*

$$f + g : V \to V' \quad ed \quad c \cdot f : V \to V'$$

definite ponendo

$$(f+g)(\mathbf{u}) := f(\mathbf{u}) + g(\mathbf{u}), \qquad (c \cdot f)(\mathbf{u}) := c \cdot f(\mathbf{u})$$

sono anch'esse lineari.

Dimostrazione. 1) Poiche' $f(\mathbf{0}_V) = f(\mathbf{0}_V + \mathbf{0}_V)$ ed f e' lineare abbiamo:

$$f(\mathbf{0}_V) = f(\mathbf{0}_V) + f(\mathbf{0}_V).$$

Quindi

$$\begin{aligned}\mathbf{0}_{V'} &= f(\mathbf{0}_V) - f(\mathbf{0}_V) = (f(\mathbf{0}_V) + f(\mathbf{0}_V)) - f(\mathbf{0}_V) \\ &= f(\mathbf{0}_V) + (f(\mathbf{0}_V) - f(\mathbf{0}_V)) = f(\mathbf{0}_V) + \mathbf{0}_{V'} = f(\mathbf{0}_V).\end{aligned}$$

2) Poiche' f e' lineare abbiamo:

$$\mathbf{0}_{V'} = f(\mathbf{0}_V) = f(\mathbf{u} - \mathbf{u}) = f(\mathbf{u} + (-\mathbf{u})) = f(\mathbf{u}) + f(-\mathbf{u}).$$

Cio' vuol dire proprio che $f(-\mathbf{u})$ e' l'opposto di $f(\mathbf{u})$.

3) Poiche' f e' lineare abbiamo:

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

4) Poiche' f e' lineare abbiamo:

$$f(a\mathbf{u} + b\mathbf{v}) = f(a\mathbf{u}) + f(b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v}).$$

5) Per provare che f^{-1} e' lineare dobbiamo provare che per ogni $\mathbf{u}', \mathbf{v}' \in V'$ ed ogni scalare a si ha:

$$f^{-1}(\mathbf{u}' + \mathbf{v}') = f^{-1}(\mathbf{u}') + f^{-1}(\mathbf{v}'), \quad \text{ed} \quad f^{-1}(a\mathbf{u}') = af^{-1}(\mathbf{u}').$$

A tale proposito, siano $\mathbf{u}, \mathbf{v} \in V$ tali che

$$\mathbf{u}' = f(\mathbf{u}), \quad \mathbf{v}' = f(\mathbf{v}).$$

Allora, tenuto conto che f e' lineare, che $f^{-1} \circ f = \mathrm{id}_V$, e che $\mathbf{u} = f^{-1}(\mathbf{u}')$ e $\mathbf{v} = f^{-1}(\mathbf{v}')$, abbiamo:

$$\begin{aligned}f^{-1}(\mathbf{u}' + \mathbf{v}') &= f^{-1}(f(\mathbf{u}) + f(\mathbf{v})) \\ &= f^{-1}(f(\mathbf{u} + \mathbf{v})) = \mathbf{u} + \mathbf{v} = f^{-1}(\mathbf{u}') + f^{-1}(\mathbf{v}').\end{aligned}$$

Similmente:

$$f^{-1}(a\mathbf{u}') = f^{-1}(af(\mathbf{u})) = f^{-1}(f(a\mathbf{u})) = a\mathbf{u} = af^{-1}(\mathbf{u}').$$

6) Occorre provare che per ogni $\mathbf{u}, \mathbf{v} \in V$ ed ogni scalare a si ha:

$$(g \circ f)(\mathbf{u} + \mathbf{v}) = (g \circ f)(\mathbf{u}) + (g \circ f)(\mathbf{v}), \quad \text{e} \quad (g \circ f)(a\mathbf{u}) = a(g \circ f)(\mathbf{u}).$$

Tenuto conto della definizione di funzione composta, e del fatto che f e g sono lineari, abbiamo:

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

Similmente

$$(g \circ f)(a\mathbf{u}) = g(f(a\mathbf{u})) = g(af(\mathbf{u})) = ag(f(\mathbf{u})) = a(g \circ f)(\mathbf{u}).$$

7) Per provare che $f + g$ e' lineare occorre provare che per ogni $\mathbf{u}, \mathbf{v} \in V$ ed ogni scalare a si ha:

$$(f+g)(\mathbf{u}+\mathbf{v}) = (f+g)(\mathbf{u}) + (f+g)(\mathbf{v}), \quad \text{e} \quad (f+g)(a\mathbf{u}) = a((f+g)(\mathbf{u})).$$

Tenuto conto della definizione di $f+g$, che f e g sono lineari, e delle proprieta' commutativa e distributiva, abbiamo:

$$\begin{aligned}(f+g)(\mathbf{u}+\mathbf{v}) &= f(\mathbf{u}+\mathbf{v}) + g(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) + g(\mathbf{u}) + g(\mathbf{v}) \\ &= f(\mathbf{u}) + g(\mathbf{u}) + f(\mathbf{v}) + g(\mathbf{v}) == (f+g)(\mathbf{u}) + (f+g)(\mathbf{v}),\end{aligned}$$

e

$$\begin{aligned}(f+g)(a\mathbf{u}) &= f(a\mathbf{u}) + g(a\mathbf{u}) = af(\mathbf{u}) + ag(\mathbf{u}) \\ &= a(f(\mathbf{u}) + g(\mathbf{u})) = a((f+g)(\mathbf{u})).\end{aligned}$$

Infine, per provare che $c \cdot f$ e' lineare occorre provare che per ogni $\mathbf{u}, \mathbf{v} \in V$ ed ogni scalare a si ha:

$$(c \cdot f)(\mathbf{u}+\mathbf{v}) = (c \cdot f)(\mathbf{u}) + (c \cdot f)(\mathbf{v}), \quad \text{e} \quad (c \cdot f)(a\mathbf{u}) = a((c \cdot f)(\mathbf{u})).$$

Ora tenuto conto della definizione di $c \cdot f$, che f e' lineare, della proprieta' distributiva e della proprieta' associativa della moltiplicazione esterna (e della proprieta' commutativa della moltiplicazione tra numeri reali), abbiamo:

$$\begin{aligned}(c \cdot f)(\mathbf{u}+\mathbf{v}) &= c \cdot (f(\mathbf{u}+\mathbf{v})) = c \cdot (f(\mathbf{u}) + f(\mathbf{v})) \\ &= c \cdot f(\mathbf{u}) + c \cdot f(\mathbf{v}) = (c \cdot f)(\mathbf{u}) + (c \cdot f)(\mathbf{v})\end{aligned}$$

e

$$\begin{aligned}(c \cdot f)(a\mathbf{u}) &= c \cdot f(a\mathbf{u}) = c \cdot (af(\mathbf{u})) = (ca)f(\mathbf{u}) = (ac)f(\mathbf{u}) \\ &= a(c(f(\mathbf{u}))) = a((c \cdot f)(\mathbf{u})). \quad \square\end{aligned}$$

Esempi.

1) L'applicazione delle coordinate e' un'applicazione lineare.

2) L'applicazione nulla $\mathbf{u} \in V \to \mathbf{0}_{V'} \in V'$ e' lineare.

3) L'applicazione identica $\mathrm{id}_V : \mathbf{u} \in V \to \mathbf{u} \in V$ e' lineare. Inoltre, se W e' un sottospazio di V, l'inclusione di W in V, cioe' l'applicazione $i : \mathbf{w} \in W \to \mathbf{w} \in V$ e' lineare.

4) La funzione derivata $p(t) \in \mathbf{R}[t] \to p'(t) \in \mathbf{R}[t]$ e' un'applicazione lineare.

5) Fissato un numero reale a, l'applicazione

$$f : x \in \mathbf{R} \to f(x) = ax \in \mathbf{R}$$

e' lineare. Cio' segue dalla proprieta' associativa e distributiva. Infatti $f(x+y) = a(x+y) = ax + ay = f(x) + f(y)$, e $f(bx) = a(bx) = (ab)x = (ba)x = b(ax) = bf(x)$ (e' intevenuta anche la proprieta' commutativa della moltiplicazione tra numeri reali). E' interessante notare che tutte le applicazioni lineari $f : \mathbf{R} \to \mathbf{R}$ sono di questo tipo. Infatti se f e' lineare, posto

$$a := f(1),$$

allora $f(x) = f(1 \cdot x) = f(x \cdot 1) = xf(1) = xa = ax$. Percio' le uniche applicazioni lineari $f : \mathbf{R} \to \mathbf{R}$ sono i polinomi di primo grado, con termine costante nullo. Questo e' un fatto particolare di una proprieta' generale che andremo a studiare tra poco. Come conseguenza osserviamo che allora le applicazioni di $\mathbf{R}$ in $\mathbf{R}$ definite ponendo $f(x) = x + 1$, $g(x) = x^2$, $h(x) = \cos x$ sono esempi di applicazioni non lineari.

6) L'esempio precedente si puo' generalizzare nel seguente modo. Fissiamo una matrice A, $m \times n$. Denotiamo con L_A la seguente applicazione:

$$L_A : \mathbf{x} \in \mathbf{R}^n \to L_A(\mathbf{x}) = A \cdot \mathbf{x} \in \mathbf{R}^m$$

(per poter eseguire la moltiplicazione righe per colonna $A \cdot \mathbf{x}$ dobbiamo pensare il vettore $\mathbf{x}$ messo in colonna; in altre parole, nella funzione precedente, stiamo identificando $\mathbf{R}^n$ con $\mathcal{M}(n, 1)$, ed $\mathbf{R}^m$ con $\mathcal{M}(m, 1)$). Si noti l'analogia con l'esempio precedente. Ebbene anche in questo caso l'applicazione L_A e' lineare. Cio' segue dalle proprieta' di calcolo delle matrici. Infatti:

$$L_A(\mathbf{x} + \mathbf{y}) = A \cdot (\mathbf{x} + \mathbf{y}) = A \cdot \mathbf{x} + A \cdot \mathbf{y} = L_A(\mathbf{x}) + L_A(\mathbf{y}),$$

e

$$L_A(a\mathbf{x}) = A \cdot (a\mathbf{x}) = a(A \cdot \mathbf{x}) = aL_A(\mathbf{x}).$$

Per fare un esempio esplicito, se poniamo

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

tenuto conto che

$$A \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y + z \\ x - y - 3z \end{bmatrix}$$

abbiamo:

$$L_A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y + z \\ x - y - 3z \end{bmatrix}.$$

7) L'esempio precedente ci consente di costruire applicazioni lineari $V \to V'$ tra spazi vettoriali qualsiasi. Infatti, fissiamo una base $\mathcal{B}$ di V ed una base $\mathcal{B}'$ di V'. Supponiamo che V abbia dimensione n, e che V' abbia dimensione m. Sia A una matrice $m \times n$. Possiamo allora definire un'applicazione lineare $f : V \to V'$ percorrendo il seguente diagramma:

(1)
$$\begin{tikzcd} V \arrow[r, dashed, "f"] \arrow[d, "{[\ ]_{\mathcal{B}}}"'] & V' \\ M(n,1) \arrow[r, "L_A"] & \mathcal{M}(m,1) \arrow[u, "{[\ ]_{\mathcal{B}'}^{-1}}"'] \end{tikzcd}$$

cioe' definiamo f come la funzione composta delle seguenti applicazioni lineari:

$$f := [\ ]_{\mathcal{B}'}^{-1} \circ L_A \circ [\ ]_{\mathcal{B}}.$$

Poiche' f e' costruita utilizzando le basi $\mathcal{B}$, $\mathcal{B}'$, e la matrice A, useremo la notazione:

$$f_{\mathcal{B}'}^{\mathcal{B}}(A) := f.$$

Questa funzione agisce al seguente modo: parte da un vettore $\mathbf{u} \in V$, poi calcola le coordinate $\mathbf{x} = [\mathbf{u}]_{\mathcal{B}}$ di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, moltiplica il vettore $\mathbf{x}$ per la matrice A, ottiene cosi' un nuovo vettore numerico $\mathbf{y} = A \cdot \mathbf{x}$ (per fare questo occorre considerare $\mathbf{x}$ come una colonna), e poi definisce $f(\mathbf{u})$ come quel vettore di V' che, rispetto alla base $\mathcal{B}'$, ha coordinate $\mathbf{y}$. E' notevole il fatto che questo modo di costruire una funzione lineare e' assolutamente generale, cioe' tutte le applicazioni lineari sono fatte in questo modo.

Ne parleremo nel prossimo paragrafo, quando introdurremo la matrice rappresentativa di un'applicazione lineare. Vediamo un altro esempio.

8) Vediamo come agisce l'applicazione lineare $f = f_{\mathcal{E}_4}^{\mathcal{E}_3}(A)$ quando $V = \mathbf{R}^3$, $V' = \mathbf{R}^4$, $\mathcal{E}_3$ e' la base canonica di $\mathbf{R}^3$, $\mathcal{E}_4$ e' la base canonica di $\mathbf{R}^4$, ed A e' la matrice

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

Sia $\mathbf{u} = (x, y, z) \in \mathbf{R}^3$ un qualunque vettore. Tenuto conto che l'applicazione delle coordinate rispetto alla base canonica mette in colonna un vettore numerico riga, e che

$$A \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ x-z \\ -2x-y+z \\ 3x+y-2z \end{bmatrix},$$

allora abbiamo:

$$f(x, y, z) = (x + y, x - z, -2x - y + z, 3x + y - 2z).$$

Per esempio $f(0, 0, 0) = (0, 0, 0, 0)$, $f(1, 0, 0) = (1, 1, -2, 3)$, $f(0, 1, 0) = (1, 0, -1, 1)$, $f(0, 0, 1) = (0, -1, 1, -2)$, $f(1, 1, 1) = (2, 0, -2, 2)$. Dalle immagini dei vettori della base canonica di $\mathbf{R}^3$ ritroviamo le colonne di A. Inoltre osserviamo che, identificati i vettori con le rispettive coordinate, l'applicazione f trasforma un vettore numerico (x, y, z) in un altro vettore numerico $f(x, y, z)$ le cui componenti appaiono come polinomi omogenei di primo grado nelle variabili x, y, z. I coefficienti (ordinati) di tali polinomi rappresentano le righe della matrice A. Possiamo anche provare direttamente che la funzione $f(x, y, z) = (x + y, x - z, -2x - y + z, 3x + y - 2z)$ e' lineare. Infatti, per ogni $\mathbf{u} = (x, y, z), \mathbf{u'} = (x', y', z') \in \mathbf{R}^3$ ed ogni scalare $a \in \mathbf{R}$ abbiamo:

$$\begin{gathered} f(\mathbf{u}+\mathbf{v}) = f(x+x', y+y', z+z') = ((x+x')+(y+y'), (x+x')-(z+z'), \\ -2(x+x')-(y+y')+(z+z'), 3(x+x')+(y+y')-2(z+z')) \\ = ((x+y)+(x'+y'), (x-z)+(x'-z'), (-2x-y+z)+(-2x'-y'+z'), \\ (3x+y-2z)+(3x'+y'-2z')) = (x+y, x-z, -2x-y+z, 3x+y-2z) \\ +(x'+y', x'-z', -2x'-y'+z', 3x'+y'-2z') = f(\mathbf{u})+f(\mathbf{v}) \end{gathered}$$

e

$$\begin{gathered} f(a\mathbf{u}) = f((ax, ay, az)) \\ = (ax+ay, ax-az, -2ax-ay+az, 3ax+ay-2az) \\ = a(x+y, x-z, -2x-y+z, 3x+y-2z) = af(\mathbf{u}). \end{gathered}$$

Osservazione. Riprendiamo il diagramma (1) discusso in precedenza. Piu' precisamente, dati due spazi vettoriali V e V', di dimensione n ed m, fissiamo una base $\mathcal{B}$ di V ed una base $\mathcal{B}'$ di V'. Una volta fissate le basi, ne risulta *una corrispondenza biiettiva tra l'insieme* ${V'}^V$ *delle funzioni da* V *a* V', *e l'insieme* ${\mathbf{R}^m}^{\mathbf{R}^n}$ *delle funzioni da* $\mathbf{R}^n$ *ad* $\mathbf{R}^m$. Infatti, come in (1), se partiamo da una funzione qualsiasi $F : \mathbf{R}^n \to \mathbf{R}^m$, tramite il diagramma

$$\begin{array}{ccc} V & & V' \\ [\ ]_{\mathcal{B}} \downarrow & & \uparrow [\ ]_{\mathcal{B}'}^{-1} \\ \mathbf{R}^n & \xrightarrow{F} & \mathbf{R}^m \end{array}$$

possiamo costruire una funzione $f : V \to V'$ ponendo

$$f := [\ ]_{\mathcal{B}'}^{-1} \circ F \circ [\ ]_{\mathcal{B}}.$$

Cioe', definiamo la funzione f come quella funzione che ad ogni vettore $\mathbf{u}$ di V associa quel vettore di V' che ha come coordinate rispetto alla base $\mathcal{B}'$ l'immagine tramite F delle coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$:

$$[f(\mathbf{u})]_{\mathcal{B}'} := F([\mathbf{u}]_{\mathcal{B}}).$$

Viceversa, se partiamo da una funzione qualsiasi $f : V \to V'$, tramite il diagramma

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ [\ ]_{\mathcal{B}}^{-1} \uparrow & & \downarrow [\ ]_{\mathcal{B}'} \\ \mathbf{R}^n & & \mathbf{R}^m \end{array}$$

possiamo costruire una funzione $F : \mathbf{R}^n \to \mathbf{R}^m$ ponendo

$$F := [\ ]_{\mathcal{B}'} \circ f \circ [\ ]_{\mathcal{B}}^{-1}.$$

Cioe' definiamo la funzione F come quella funzione che ad ogni vettore $\mathbf{x} \in \mathbf{R}^n$ associa quel vettore di $\mathbf{R}^m$ che rappresenta le coordinate rispetto alla base $\mathcal{B}'$ dell'immagine tramite f del vettore $\mathbf{u}$ di V che ha coordinate $\mathbf{x}$ rispetto alla base $\mathcal{B}$:

$$F(\mathbf{x}) := [f(\mathbf{u})]_{\mathcal{B}'} \tag{2}$$

Le corrispondenze cosi' definite

$$f \in {V'}^{V} \to F \in {\mathbf{R}^m}^{\mathbf{R}^n}, \quad F \in {\mathbf{R}^m}^{\mathbf{R}^n} \to f \in {V'}^{V},$$

sono biiettive, una inversa dell'altra. Si osservi che *tali costruzioni non sono naturali, in quanto abbiamo dovuto fissare delle basi.* Percio', se cambiamo le basi, le corrispondenze precedenti cambiano. Tenuto conto che composte di applicazioni lineari sono lineari, dalle formule precedenti si evince che *f e' lineare se e solo se F e' lineare.* Nel prossimo paragrafo andremo a provare che se f e' lineare, allora esiste un'unica matrice A tale che $F = L_A$. Chiameremo tale matrice A la matrice rappresentativa di f rispetto alle basi fissate, e la denoteremo con il simbolo $M_{\mathcal{B}'}^{\mathcal{B}}(f)$. Cio' si puo' esprimere brevemente dicendo che *a patto di introdurre le coordinate, ogni applicazione lineare e' del tipo L_A, per un'opportuna matrice A.*

Siamo pronti per studiare la matrice rappresentativa.

5. La matrice rappresentativa.

- *La definizione di matrice rappresentativa.*

Sia $f : V \to V'$ un' applicazione lineare, $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ una base di V e $\mathcal{B}' = \{\mathbf{b}'_1, \dots, \mathbf{b}'_m\}$ una base di V'. Si definisce matrice rappresentativa di f rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$ quella matrice A, con m righe ed n colonne, la cui colonna di posto j e' data dal vettore $[f(\mathbf{b}_j)]_{\mathcal{B}'}$ delle coordinate di $f(\mathbf{b}_j)$ rispetto alla base $\mathcal{B}'$. Denoteremo tale matrice anche con il simbolo

$$M_{\mathcal{B}'}^{\mathcal{B}}(f) := A.$$

- *Le proprieta' della matrice rappresentativa.*

1) *Sia $f : V \to V'$ un' applicazione lineare, $\mathcal{B}$ una base di V e $\mathcal{B}'$ una base di V'. Allora per ogni $\mathbf{u} \in V$ si ha* [3]*:*

$$[f(\mathbf{u})]_{\mathcal{B}'} = M^{\mathcal{B}}_{\mathcal{B}'}(f) \cdot [\mathbf{u}]_{\mathcal{B}}.$$

2) *Se $f : V \to V'$ e $g : V' \to V''$ sono applicazioni lineari, $\mathcal{B}$ una base di V, $\mathcal{B}'$ una base di V', e $\mathcal{B}''$ una base di V'', allora $g \circ f$ e' un'applicazione lineare e*

$$M^{\mathcal{B}}_{\mathcal{B}''}(g \circ f) = M^{\mathcal{B}'}_{\mathcal{B}''}(g) \cdot M^{\mathcal{B}}_{\mathcal{B}'}(f).$$

3) *Se $f : V \to V'$ e' un' applicazione lineare biiettiva tra spazi della stessa dimensione, $\mathcal{B}$ una base di V e $\mathcal{B}'$ una base di V', allora anche f^{-1} e' lineare, la matrice $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ e' invertibile, e*

$$M^{\mathcal{B}'}_{\mathcal{B}}(f^{-1}) = M^{\mathcal{B}}_{\mathcal{B}'}(f)^{-1}.$$

Dimostrazione. 1) Denotiamo con $\mathbf{x}$ le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, con $\mathbf{y}$ le coordinate di $f(\mathbf{u})$ rispetto alla base $\mathcal{B}'$, e con A la matrice rappresentativa $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ di f. Dobbiamo provare che

$$\mathbf{y} = A \cdot \mathbf{x}.$$

Per fare cio', tenuto conto che f e' lineare e che tale e' l'applicazione delle coordinate, osserviamo che:

$$\begin{aligned} \mathbf{y} = [f(\mathbf{u})]_{\mathcal{B}'} &= [f(x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n)]_{\mathcal{B}'} \\ = [x_1 f(\mathbf{b}_1) + \dots + x_n f(\mathbf{b}_n)]_{\mathcal{B}'} &= x_1[f(\mathbf{b}_1)]_{\mathcal{B}'} + \dots + x_n[f(\mathbf{b}_n)]_{\mathcal{B}'} \\ &= x_1 A^1 + \dots + x_n A^n, \end{aligned}$$

dove con A^j denotiamo la colonna di posto j di A. Poiche' $A^j = A \cdot \mathbf{e}_j$ ($\mathbf{e}_j$ = vettore canonico colonna di posto j di $\mathbf{R}^n$), allora

$$x_1 A^1 + \dots + x_n A^n$$

$$= x_1(A \cdot \mathbf{e}_1) + \dots + x_n(A \cdot \mathbf{e}_n) = A \cdot (x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) = A \cdot \mathbf{x}.$$

2) Gia' sappiamo che $g \circ f$ e' lineare. Denotiamo con A, B e C le matrici rappresentative $M^{\mathcal{B}}_{\mathcal{B}'}(f)$, $M^{\mathcal{B}'}_{\mathcal{B}''}(g)$ e $M^{\mathcal{B}}_{\mathcal{B}''}(g \circ f)$. Occorre provare che

$$C = B \cdot A.$$

Per ogni vettore $\mathbf{u} \in V$ denotiamo con $\mathbf{x}$ le coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, con $\mathbf{y}$ le coordinate di $f(\mathbf{u})$ rispetto alla base $\mathcal{B}'$, e con $\mathbf{z}$ le coordinate di $(g \circ f)(\mathbf{u})$ rispetto alla base $\mathcal{B}''$. Per la proprieta' 1) appena dimostrata sappiamo che

$$\mathbf{y} = A \cdot \mathbf{x}, \quad \mathbf{z} = B \cdot \mathbf{y}, \quad \mathbf{z} = C \cdot \mathbf{x}.$$

Quindi sostituendo:

$$\mathbf{z} = B \cdot \mathbf{y} = B \cdot (A \cdot \mathbf{x}) = (B \cdot A) \cdot \mathbf{x}$$

e percio'

$$C \cdot \mathbf{x} = (B \cdot A) \cdot \mathbf{x}$$

per ogni $\mathbf{x} \in \mathbf{R}^n$ ($n = \dim(V)$). Sappiamo che cio' implica che $C = B \cdot A$.

3) Gia' sappiamo che l'inversa di f e' lineare. Denotiamo con A la matrice rappresentativa $M^{\mathcal{B}}_{\mathcal{B}'}(f)$, e con B la matrice $M^{\mathcal{B}'}_{\mathcal{B}}(f^{-1})$. Per ogni vettore $\mathbf{u}$ di V denotiamo con $\mathbf{x}$ le

[3] Con le notazioni dell'Osservazione precedente, tale proprieta' equivale a dire che $F = L_A$ (si confronti con (2)).

sue coordinate rispetto alla base $\mathcal{B}$, e con $\mathbf{y}$ le coordinate di $\mathbf{u}' := f(\mathbf{u})$ rispetto alla base $\mathcal{B}'$. Tenuto conto che $\mathbf{u} = f^{-1}(\mathbf{u}')$, per la proprieta' 1) appena dimostrata sappiamo che:

$$\mathbf{y} = A \cdot \mathbf{x}, \quad \mathbf{x} = B \cdot \mathbf{y}.$$

Quindi

$$\mathbf{x} = B \cdot \mathbf{y} = B \cdot (A \cdot \mathbf{x}) = (B \cdot A) \cdot \mathbf{x}.$$

Cioe' per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha

$$\mathbf{x} = I \cdot \mathbf{x} = (B \cdot A) \cdot \mathbf{x}.$$

Cio' implica che $I = B \cdot A$, percio' A e' invertibile e la sua inversa e' B. $\square$

Osservazioni. (i) Posto $A = M^{\mathcal{B}}_{\mathcal{B}'}(f)$, la proprieta' 1) ci dice proprio che $f = f^{\mathcal{B}}_{\mathcal{B}'}(A)$. Percio' i commenti che abbiamo fatto negli esempi precedenti si applicano a tutte le applicazioni lineari. In particolare, una volta identificati i vettori con le rispettive coordinate, un'applicazione lineare f trasforma un vettore numerico

$$(x_1, \dots, x_n) \in \mathbf{R}^n$$

in un altro vettore numerico

$$f(x_1, \dots, x_n) \in \mathbf{R}^m$$

dato da un'espressione del tipo:

$$f(x_1, \dots, x_n)$$
$$= (a_{11}x_1 + \dots + a_{1n}x_n,\ a_{21}x_1 + \dots + a_{2n}x_n,\ \dots,\ a_{m1}x_1 + \dots + a_{mn}x_n),$$

dove i numeri a_{ij} sono le componenti della matrice rappresentativa A. Tale espressione si chiama *l'espressione esplicita di f rispetto alle basi fissate.* Le componenti del vettore immagine $f(x_1, \dots, x_n)$ appaiono come polinomi omogenei di primo grado nelle variabili $x_1, \dots, x_n$. I coefficienti (ordinati) di tali polinomi rappresentano le righe della matrice rappresentativa A.

(ii) La proprieta' 2) ci dice che la matrice rappresentativa di una composizione di due applicazioni lineari e' uguale al prodotto delle due matrici rappresentative. Anche questa proprieta' e' una proprieta' di *giustificazione*, nel senso che giustifica la definizione di prodotto righe per colonna con cui abbiamo definito il prodotto tra matrici.

(iii) Cosi' come le coordinate di un vettore variano al variare della base a cui sono riferite, cosi' la matrice rappresentativa di una data applicazione lineare varia al variare delle basi.

(iv) E' chiaro dalla definizione che la matrice del cambiamento delle coordinate coincide con la matrice rappresentativa dell'applicazione identica.

(v) La matrice rappresentativa dell'applicazione delle coordinate $[\ ]_{\mathcal{B}} : V \to \mathbf{R}^n$, riferita alla base $\mathcal{B}$ in partenza ed alla base canonica in arrivo, e' la matrice unitaria I:

$$M^{\mathcal{B}}_{\mathcal{E}}([\ ]_{\mathcal{B}}) = I.$$

(vi) Si osservi che la matrice rappresentativa dell'applicazione $L_A : \mathbf{R}^n \to \mathbf{R}^m$ riferita alle basi canoniche, e' proprio A:

$$M^{\mathcal{E}_n}_{\mathcal{E}_m}(L_A) = A.$$

Esercizio. Sia $f : \mathbf{R}^2 \to \mathbf{R}^4$ l'applicazione definita ponendo

$$f(x, y) = (x - y, x - y, 2x - 2y, 3x - 3y).$$

Siano $\mathcal{E}_2$ ed $\mathcal{E}_4$ le basi canoniche di $\mathbf{R}^2$ ed $\mathbf{R}^4$, e $\mathcal{B}$ la base di $\mathbf{R}^2$ formata dai vettori $(2,1),(1,1)$. Sia $\mathbf{u}$ il vettore di $\mathbf{R}^2$ che ha coordinate $(-1,3)^T$ rispetto alla base $\mathcal{B}$. Provare che f e' lineare, e calcolare $M_{\mathcal{E}_4}^{\mathcal{E}_2}(f)$, $M_{\mathcal{E}_4}^{\mathcal{B}}(f)$, $f(\mathbf{u})$, l'espressione esplicita di f riferita alle basi canoniche, e l'espressione esplicita di f riferita alle basi $\mathcal{B}$ ed $\mathcal{E}_4$.

Svolgimento. A patto di mettere i vettori riga in colonna, l'applicazione f coincide con l'applicazione

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2 \to A \cdot \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^4$$

dove A e' la matrice

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

Cio' prova che $f = L_A$, dunque e' lineare, e la sua espressione esplicita riferita alle basi canoniche e':

$$f(x,y) = (x-y, x-y, 2x-2y, 3x-3y).$$

In particolare

$$A = M_{\mathcal{E}_4}^{\mathcal{E}_2}(f) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{bmatrix}.$$

Ora osserviamo che

$$f = f \circ \mathrm{id}_{\mathbf{R}^2}.$$

Percio'

$$M_{\mathcal{E}_4}^{\mathcal{B}}(f) = M_{\mathcal{E}_4}^{\mathcal{E}_2}(f) \cdot M_{\mathcal{E}_2}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^2}) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 0 \\ 3 & 0 \end{bmatrix}.$$

Quindi, denotate con $(x',y')^T$ le coordinate rispetto alla base $\mathcal{B}$, l'espressione esplicita di f diviene

$$f(x',y') = (x', x', 2x', 3x').$$

Infine,

$$[f(\mathbf{u})]_{\mathcal{E}_4} = M_{\mathcal{E}_4}^{\mathcal{B}}(f) \cdot [\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 0 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -2 \\ -3 \end{bmatrix}.$$

Cioe'

$$f(\mathbf{u}) = (-1,-1,-2,-3). \quad \square$$

Esercizio. Sia $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ una base dello spazio vettoriale V, e siano $\mathbf{u}_1, \dots, \mathbf{u}_n$ vettori qualsiasi di un altro spazio V'. Provare che esiste ed e' unica un'applicazione lineare $f : V \to V'$ tale che $f(\mathbf{b}_i) = \mathbf{u}_i$ per ogni $i = 1, \dots, n$.

Svolgimento. Esistenza. Fissiamo una base $\mathcal{B}' = \{\mathbf{b}'_1, \dots, \mathbf{b}'_m\}$ di V', e sia A quella matrice $m \times n$ che ha per colonne le coordinate dei vettori $\mathbf{u}_j$ rispetto alla base $\mathcal{B}'$. Allora la funzione $f_{\mathcal{B}'}^{\mathcal{B}}(A)$ soddisfa le condizioni richieste.

Unicita'. Sia $g : V \to V'$ un'altra applicazione lineare tale che $g(\mathbf{b}_i) = \mathbf{u}_i$ per ogni $i = 1, \dots, n$. Allora per ogni $\mathbf{u} \in V$, denotate con $\mathbf{x}$ le sue coordinate rispetto alla base $\mathcal{B}$, abbiamo:

$$g(\mathbf{u}) = g(x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n) = x_1 g(\mathbf{b}_1) + \dots + x_n g(\mathbf{b}_n)$$
$$= x_1 f(\mathbf{b}_1) + \dots + x_n f(\mathbf{b}_n) = f(x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n) = f(\mathbf{u}). \quad \square$$

Esercizio. Sia $\mathcal{B} = \{\mathbf{b}_1, , \mathbf{b}_2, \mathbf{b}_3\}$ *la base di* $\mathbf{R}^3$ *formata dai vettori* $\mathbf{b}_1 = (-1, 0, 1)$, $\mathbf{b}_2 = (1, 0, 0)$, $\mathbf{b}_3 = (-1, -1, 1)$. *Sia poi* $f : \mathbf{R}^3 \to \mathbf{R}^3$ *l'operatore lineare tale che* $f(\mathbf{b}_1) = 3\mathbf{b}_1 - 2\mathbf{b}_3$, $f(\mathbf{b}_2) = \mathbf{b}_1 - \mathbf{b}_2 - \mathbf{b}_3$, $f(\mathbf{b}_3) = f(\mathbf{b}_1)$. *Calcolare la matrice rappresentativa* $M_{\mathcal{E}}^{\mathcal{E}}(f)$ *di* f *rispetto alla base canonica* $\mathcal{E}$.

Svolgimento. Denotiamo i vettori canonici con $\mathbf{e}_1$, $\mathbf{e}_2$, ed $\mathbf{e}_3$. Osserviamo che

$$f(\mathbf{b}_1) = 3\mathbf{b}_1 - 2\mathbf{b}_3 = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3.$$

D'altra parte, poiche' f e' lineare, abbiamo anche

$$f(\mathbf{b}_1) = f(-\mathbf{e}_1 + \mathbf{e}_3) = -f(\mathbf{e}_1) + f(\mathbf{e}_3).$$

Quindi otteniamo:

$$-f(\mathbf{e}_1) + f(\mathbf{e}_3) = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3.$$

Svolgendo un calcolo analogo per $f(\mathbf{b}_2)$ ed $f(\mathbf{b}_3)$ otteniamo le seguenti relazioni:

$$\begin{cases} -f(\mathbf{e}_1) + f(\mathbf{e}_3) = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3 \\ f(\mathbf{e}_1) = -\mathbf{e}_1 + \mathbf{e}_2 \\ -f(\mathbf{e}_1) - f(\mathbf{e}_2) + f(\mathbf{e}_3) = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3. \end{cases}$$

Risolvendo otteniamo:

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

Da cui:

$$M_{\mathcal{E}}^{\mathcal{E}}(f) = \begin{bmatrix} -1 & 0 & -2 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}. \quad \square$$

Esercizio. Si consideri la base $\mathcal{B}$ *di* $\mathbf{R}^3$ *formata dai vettori* $(1, 0, 1)$, $(1, 0, 2)$, $(1, -1, 1)$. *Sia* $f : \mathbf{R}^3 \to \mathbf{R}^3$ *l'operatore lineare la cui rappresentazione esplicita rispetto alla base* $\mathcal{B}$, *in partenza ed in arrivo, e':* $f(x_1', x_2', x_3') = (x_1' + x_2' + x_3', x_1' + 2x_2' + 3x_3', 2x_1' + 3x_2' + 4x_3')$. *Calcolare la matrice rappresentativa* $M_{\mathcal{E}}^{\mathcal{E}}(f)$ *di* f *rispetto alla base canonica, e l'espressione esplicita corrispondente.*

Svolgimento. Sappiamo che

$$M_{\mathcal{E}}^{\mathcal{E}}(f) = M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^3}) M_{\mathcal{B}}^{\mathcal{B}}(f) M_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^3}) = M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^3}) M_{\mathcal{B}}^{\mathcal{B}}(f) M_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}_{\mathbf{R}^3})^{-1}$$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -4 & 2 \\ -1 & 2 & -1 \\ 2 & -6 & 3 \end{bmatrix}.$$

E l'espressione esplicita corrispondente e'

$$f(x_1, x_2, x_3) = (2x_1 - 4x_2 + 2x_3, -x_1 + 2x_2 - x_3, 2x_1 - 6x_2 + 3x_3). \quad \square$$

Esercizio. Calcolare la matrice rappresentativa $M_{\mathcal{E}}^{\mathcal{E}}(D)$ *dell'applicazione di derivata* $D : \mathbf{R}[t]_{\leq 3} \to \mathbf{R}[t]_{\leq 3}$, *dove* $\mathcal{E}$ *denota la base canonica* $\{1, t, t^2, t^3\}$.

Svolgimento. Poiche' $D(1) = 0$, $D(t) = 1$, $D(t^2) = 2t$, $D(t^3) = 3t^2$, allora

$$M_{\mathcal{E}}^{\mathcal{E}}(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Si osservi che l'espressione esplicita e'

$$f(x, y, z, t) = (y, 2z, 3t, 0). \quad \square$$

Adesso andiamo a studiare alcuni esempi geometrici di applicazioni lineari.

• *Rotazioni, riflessioni e proiezioni ortogonali nel piano.*

Le rotazioni.

Fissiamo un punto O nel piano ρ, fissiamo un angolo α, e sia $\overrightarrow{OX}$ un vettore geometrico applicato in O. Sia $\overrightarrow{OY}$ il vettore applicato in O che si ottiene ruotando in senso antiorario il vettore $\overrightarrow{OX}$, fino a percorrere un'ampiezza data dall'angolo fissato α. Abbiamo cosi' definito un'applicazione:

$$f : \overrightarrow{OX} \in \mathcal{V}_{O,\rho} \to \overrightarrow{OY} \in \mathcal{V}_{O,\rho},$$

detta *la rotazione di angolo α nel piano ρ.*

Ora introduciamo una base $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$ per lo spazio $\mathcal{V}_{O,\rho}$ dei vettori geometrici del piano, formata da due vettori di lunghezza 1 ed ortogonali fra loro, e siano $\mathbf{x} = (x_1, x_2)^T$ le coordinate di $\overrightarrow{OX}$. *Ci chiediamo come si esprimono le coordinate* $\mathbf{y} = (y_1, y_2)^T$ *di* $\overrightarrow{OY}$ *in funzione di* $\mathbf{x}$. A tale proposito, sia φ l'angolo formato dal vettore $\overrightarrow{OX}$ rispetto all'asse $\overrightarrow{OE_1}$. Si ha

$$x_1 = l \cos\varphi, \quad x_2 = l \sin\varphi,$$

dove l denota la lunghezza di $\overrightarrow{OX}$, che e' anche la lunghezza di $\overrightarrow{OY}$. Poiche' l'angolo che $\overrightarrow{OY}$ forma con l'asse $\overrightarrow{OE_1}$ e' $\varphi + \alpha$ segue anche che

$$y_1 = l \cos(\varphi + \alpha), \quad y_2 = l \sin(\varphi + \alpha),$$

da cui si deduce, tramite le formule di addizione e sottrazione, che

$$\mathbf{y} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \mathbf{x}.$$

Quindi possiamo riguardare la rotazione di un vettore come un'applicazione lineare, con matrice rappresentativa (riferita ad una base formata da vettori ortogonali di lunghezza 1 *del tipo* $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$*) data dalla matrice*

$$R_\alpha := \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix},$$

dove α e' l'angolo della rotazione.

Esercizio. Scrivere l'espressione esplicita di una rotazione di angolo $\alpha = \frac{\pi}{2}$.

Svolgimento. Poiche'

$$R_{\frac{\pi}{2}} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

l'espressione esplicita e':

$$f(x_1, x_2) = (-x_2, x_1). \quad \square$$

Esercizio. Scrivere l'espressione esplicita di una rotazione di angolo $\alpha = \frac{\pi}{4}$.

Svolgimento. Poiche'

$$R_{\frac{\pi}{4}} := \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix},$$

l'espressione esplicita e':

$$f(x_1, x_2) = \left(\frac{\sqrt{2}}{2}x_1 - \frac{\sqrt{2}}{2}x_2, \frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}x_2 \right). \quad \square$$

Le riflessioni e le proiezioni ortogonali.

Fissiamo un punto O nel piano ρ, e sia r una retta per O. Sia $\overrightarrow{OX}$ un vettore geometrico applicato in O. Sia h la retta del piano passante per X ortogonale ad r. Tale retta interseca r in un punto H. Sia Y il punto di h speculare di X rispetto alla retta r. Allora $\overrightarrow{OY}$ *si chiama la riflessione di* $\overrightarrow{OX}$ *rispetto alla retta* r*, mentre* $\overrightarrow{OH}$ *si chiama la proiezione ortogonale di* $\overrightarrow{OX}$ *su* r. Abbiamo cosi' definito due applicazioni:

$$\overrightarrow{OX} \in \mathcal{V}_{O,\rho} \to \overrightarrow{OY} \in \mathcal{V}_{O,\rho}, \quad \overrightarrow{OX} \in \mathcal{V}_{O,\rho} \to \overrightarrow{OH} \in \mathcal{V}_{O,\rho},$$

la prima detta *la riflessione (o anche la simmetria ortogonale) di asse* r *nel piano* ρ, e la seconda detta *la proiezione ortogonale sulla retta* r *nel piano* ρ.

Come nel caso delle rotazioni, introduciamo una base $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$ per lo spazio $\mathcal{V}_{O,\rho}$ dei vettori geometrici del piano, formata da due vettori di lunghezza 1 ed ortogonali fra loro, e siano $\mathbf{x} = (x_1, x_2)^T$ le coordinate di $\overrightarrow{OX}$. *Ci chiediamo come si esprimono le coordinate* $\mathbf{y} = (y_1, y_2)^T$ *di* $\overrightarrow{OY}$ *in funzione di* $\mathbf{x}$. A tale proposito, sia α l'angolo formato dalla retta r rispetto all'asse $\overrightarrow{OE_1}$, e sia φ l'angolo formato dal vettore $\overrightarrow{OX}$ rispetto all'asse $\overrightarrow{OE_1}$. Si ha

$$x_1 = l\cos\varphi, \quad x_2 = l\sin\varphi,$$

dove l denota la lunghezza di $\overrightarrow{OX}$, che e' anche la lunghezza di $\overrightarrow{OY}$. Poiche' l'angolo che $\overrightarrow{OY}$ forma con l'asse $\overrightarrow{OE_1}$ e' $2\alpha - \varphi$ segue anche che

$$y_1 = l\cos(2\alpha - \varphi), \quad y_2 = l\sin(2\alpha - \varphi),$$

da cui si deduce, tramite le formule di addizione e sottrazione, che

$$\mathbf{y} = \begin{bmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{bmatrix} \mathbf{x}.$$

Quindi possiamo riguardare la riflessione di un vettore rispetto ad una retta come una applicazione lineare, con matrice rappresentativa (riferita ad un sistema monometrico ortogonale del tipo $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$*) data dalla matrice*

$$S_\alpha := \begin{bmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{bmatrix},$$

dove α *e' l'angolo che l'asse della riflessione forma rispetto al vettore* $\overrightarrow{OE_1}$.

Una descrizione analoga sussiste per la proiezione ortogonale. Denotate con $\mathbf{z}$ le coordinate di $\overrightarrow{OH}$, si ha

$$\mathbf{z} = \begin{bmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha \end{bmatrix} \mathbf{x}.$$

Infatti sia $\mathbf{u}$ un vettore di lunghezza 1 su r, per esempio il vettore di coordinate

$$(\cos\alpha, \sin\alpha).$$

Allora il vettore $\overrightarrow{OH}$ sara' del tipo $\overrightarrow{OH} = c\mathbf{u}$. Poiche' l'angolo che $\overrightarrow{OX}$ forma con r e' $\varphi - \alpha$ allora $c = l\cos(\varphi - \alpha)$. Quindi $\mathbf{z} = l\cos(\varphi - \alpha)(\cos\alpha, \sin\alpha)$, da cui la formula precedente.

Esercizio. Scrivere l'espressione esplicita della riflessione rispetto alla bisettrice del primo quadrante formato dalla base $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$.

Svolgimento. In questo caso $\alpha = \frac{\pi}{4}$, quindi

$$S_{\frac{\pi}{4}} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

e l'espressione esplicita e':

$$f(x_1, x_2) = (x_2, x_1). \quad \square$$

Esercizio. Scrivere l'espressione esplicita della proiezione ortogonale sulla bisettrice del primo quadrante formato dalla base $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$.

Svolgimento. Come prima $\alpha = \frac{\pi}{4}$, quindi la matrice rappresentativa della proiezione ortogonale e' la matrice

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

e l'espressione esplicita e':

$$f(x_1, x_2) = \frac{1}{2}(x_1 + x_2, x_1 + x_2). \quad \square$$

6. La struttura di un'applicazione lineare: il nucleo e l'immagine.

Ad ogni applicazione lineare $f : V \to V'$ si possono associare due spazi vettoriali, detti il nucleo e l'immagine di f, che ne rappresentano in qualche modo la struttura. Le definizioni sono le seguenti.

• *La definizione del nucleo e dell'immagine.*

Sia $f : V \to V'$ un'applicazione lineare. Si definisce nucleo di f, e si denota $\ker f$, *l'insieme formato da tutti i vettori di V che hanno immagine nulla tramite f. Cioe':*

$$\ker f := \{\mathbf{u} \in V : \ f(\mathbf{u}) = \mathbf{0}_{V'}\}.$$

Si definisce immagine di f, e si denota imf o anche $\Im f$, l'insieme formato da tutti i vettori di V' della forma $f(\mathbf{u})$, al variare di $\mathbf{u}$ in V. Cioe':

$$\Im f := \{\mathbf{u}' \in V' : \ \textit{esiste } \mathbf{u} \in V \ \textit{tale che } \mathbf{u}' = f(\mathbf{u})\}.$$

Si osservi che

$$\ker f \subseteq V, \quad \Im f \subseteq V'.$$

• *Le proprieta' del nucleo.*

1) *Il nucleo e' un sottospazio di V.*

2) *f e' iniettiva se e solo se* $\ker f = \{\mathbf{0}_V\}$.

3) *Se* $\ker f = \{\mathbf{0}_V\}$, *allora f trasforma sistemi liberi in sistemi liberi.*

4) *Se $f = L_A : \mathbf{R}^n \to \mathbf{R}^m$, allora* $\ker f$ *coincide con lo spazio delle soluzioni del sistema lineare omogeneo $A \cdot \mathbf{x} = \mathbf{0}_{\mathbf{R}^m}$.*

5) *Sia* $\mathcal{B}$ *una base di* V *e* $\mathcal{B}'$ *una base di* V'. *Sia* $A = M^{\mathcal{B}}_{\mathcal{B}'}(f)$ *la matrice rappresentativa di* f, *ed* $L_A : \mathbf{x} \in \mathbf{R}^n \to A \cdot \mathbf{x} \in \mathbf{R}^m$ *l'applicazione lineare corrispondente* ($\dim V = n$, $\dim V' = m$). *Siano* $\mathbf{x}_1, \dots, \mathbf{x}_h$ *vettori di una base di* $\ker L_A$, *e siano* $\mathbf{u}_1, \dots, \mathbf{u}_h$ *i vettori di* V *ad essi corrispondenti tramite l'applicazione delle coordinate* $[\ ]_{\mathcal{B}} : V \to \mathbf{R}^n$. *Allora i vettori* $\mathbf{u}_1, \dots, \mathbf{u}_h$ *formano una base di* $\ker f$. *In particolare,* $\ker f$ *e* $\ker L_A$ *hanno la stessa dimensione.*

6) *Se* A *e' una qualunque matrice rappresentativa di* $f : V \to V'$, *allora*

$$\dim \ker f = \dim V - rkA.$$

Dimostrazione. 1) Poiche' f e' lineare sappiamo che $f(\mathbf{0}_V) = \mathbf{0}_{V'}$. Questo vuol dire che il vettore nullo sta in $\ker f$. Ora siano $\mathbf{u}, \mathbf{v}$ vettori qualsiasi del nucleo, e sia a uno scalare qualsiasi. Poiche' f e' lineare abbiamo $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) = \mathbf{0}_{V'} + \mathbf{0}_{V'} = \mathbf{0}_{V'}$, e $f(a\mathbf{u}) = af(\mathbf{u}) = a\mathbf{0}_{V'} = \mathbf{0}_{V'}$. Cio' vuol dire che anche $\mathbf{u}+\mathbf{v}$, ed $a\mathbf{u}$ stanno in $\ker f$. Percio' $\ker f$ e' un sottospazio di V.

2) Supponiamo che f sia iniettiva, e sia $\mathbf{u}$ un vettore del nucleo. Allora

$$f(\mathbf{u}) = \mathbf{0}_{V'} = f(\mathbf{0}_V).$$

Poiche' f e' iniettiva allora $\mathbf{u} = \mathbf{0}_V$. Cio' vuol dire che in $\ker f$ c'e' solo il vettore nullo.

Viceversa, supponiamo che $\ker f = \{\mathbf{0}_V\}$, e che $f(\mathbf{u}) = f(\mathbf{v})$. Allora poiche' f e' lineare abbiamo:

$$\mathbf{0}_{V'} = f(\mathbf{u}) - f(\mathbf{v}) = f(\mathbf{u} - \mathbf{v}).$$

Quindi $\mathbf{u} - \mathbf{v}$ sta nel nucleo, che per ipotesi e' costituito solo dal vettore nullo. Percio' $\mathbf{u} - \mathbf{v} = \mathbf{0}_V$, cioe' $\mathbf{u} = \mathbf{v}$. Quindi f e' iniettiva.

3) Sia $\mathbf{u}_1, \dots, \mathbf{u}_h$ un sistema libero di vettori di V, e sia

$$a_1 f(\mathbf{u}_1) + \dots + a_h f(\mathbf{u}_h) = \mathbf{0}_{V'}$$

una qualunque relazione tra i vettori immagine $f(\mathbf{u}_1), \dots, f(\mathbf{u}_h)$. Poiche' f e' lineare abbiamo:

$$a_1 f(\mathbf{u}_1) + \dots + a_h f(\mathbf{u}_h) = f(a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h).$$

Percio' il vettore $a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h$ sta nel nucleo di f, che per ipotesi e' nullo. Percio'

$$a_1\mathbf{u}_1 + \dots + a_h\mathbf{u}_h = \mathbf{0}_V.$$

Poiche' il sistema $\mathbf{u}_1, \dots, \mathbf{u}_h$ e' libero, allora la relazione $(a_1, \dots, a_h)$ e' banale. Cio' prova che $f(\mathbf{u}_1), \dots, f(\mathbf{u}_h)$ e' libero.

4) E' ovvia, perche' la funzione L_A porta il vettore $\mathbf{x} \in \mathbf{R}^n$ nel vettore $L_A(\mathbf{x}) = A \cdot \mathbf{x} \in \mathbf{R}^m$.

5) Fissiamo una base $\mathcal{B}$ di V ed una base $\mathcal{B}'$ di V'. Sia $A = M^{\mathcal{B}}_{\mathcal{B}'}(f)$ la matrice rappresentativa. Posto $n = \dim V$ ed $m = \dim V'$, consideriamo l'applicazione lineare $L_A : \mathbf{x} \in \mathbf{R}^n \to A \cdot \mathbf{x} \in \mathbf{R}^m$. Siano $\mathbf{x}_1, \dots, \mathbf{x}_h$ vettori di una base di $\ker L_A$, e siano $\mathbf{u}_1, \dots, \mathbf{u}_h$ i vettori di V ad essi corrispondenti tramite l'applicazione delle coordinate $[\ ]_{\mathcal{B}} : V \to \mathbf{R}^n$. Andiamo a provare che tali vettori sono una base per il nucleo di f. Sappiamo che i vettori $\mathbf{u}_1, \dots, \mathbf{u}_h$ sono liberi, perche' corrispondono, tramite l'applicazione delle coordinate, ad un sistema libero. Inoltre, per le proprieta' della matrice rappresentativa abbiamo:

$$[f(\mathbf{u}_i)]_{\mathcal{B}'} = A \cdot [\mathbf{u}_i]_{\mathcal{B}} = A \cdot \mathbf{x}_i = \mathbf{0}_{\mathbf{R}^m}.$$

Questo vuol dire che i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_h$ stanno nel nucleo di f. Per concludere, occorre provare che tali vettori generano il nucleo di f. Sia $\mathbf{u}$ un vettore del nucleo di f. Allora, come prima, $[\mathbf{u}]_{\mathcal{B}} \in \ker L_A$. Percio' esistono opportuni pesi $a_1, \ldots, a_h$ tali che:

$$[\mathbf{u}]_{\mathcal{B}} = a_1\mathbf{x}_1 + \cdots + a_h\mathbf{x}_h.$$

Ma

$$a_1\mathbf{x}_1 + \cdots + a_h\mathbf{x}_h = a_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + a_h[\mathbf{u}_h]_{\mathcal{B}} = [a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h]_{\mathcal{B}}.$$

Poiche' l'applicazione delle coordinate e' biiettiva, ne consegue:

$$\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h.$$

Percio' i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_h$ formano una base per il nucleo di f. In particolare $\dim\ker f = \dim\ker L_A$.

6) Il nucleo di $L_A : \mathbf{x} \in \mathbf{R}^n \to A \cdot \mathbf{x} \in \mathbf{R}^m$ coincide con lo spazio delle soluzioni del sistema lineare omogeneo $A \cdot \mathbf{x} = \mathbf{0}_{\mathbf{R}^m}$. Dalla teoria dei sistemi lineari sappiamo che tale spazio, cioe' il nucleo di L_A, ha dimensione

$$\dim\ker L_A = n - rkA = \dim V - rkA.$$

La dimostrazione finisce qui, perche' abbiamo appena provato che il nucleo di f ed il nucleo di L_A hanno la stessa dimensione. $\Box$

- *Le proprieta' dell'immagine.*

1) *$\Im f$ e' un sottospazio di V'.*

2) *f e' suriettiva se e solo se $\Im f = V'$.*

3) *Se i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_h$ generano V allora $f(\mathbf{u}_1), \ldots, f(\mathbf{u}_h)$ generano $\Im f$.*

4) *Se $f = L_A : \mathbf{R}^n \to \mathbf{R}^m$, allora $\Im L_A$ coincide con lo spazio delle colonne di A.*

5) *Sia $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base di V e $\mathcal{B}'$ una base di V'. Sia $A = M^{\mathcal{B}}_{\mathcal{B}'}(f)$ la matrice rappresentativa di f. Se le colonne $A^{i_1}, \ldots, A^{i_k}$ formano una base per lo spazio delle colonne di A, allora $\{f(\mathbf{b}_{i_1}), \ldots, f(\mathbf{b}_{i_k})\}$ e' una base per $\Im f$.*

6) *Se A e' una qualunque matrice rappresentativa di f, allora $\dim\Im f = rkA$.*

7) $\dim V = \dim\ker f + \dim\Im f$ *(Teorema della dimensione).*

Dimostrazione. 1) Poiche' $f(\mathbf{0}_V) = \mathbf{0}_{V'}$ allora $\mathbf{0}_{V'} \in \Im f$. Siano ora $\mathbf{u}', \mathbf{v}'$ due qualunque vettori di $\Im f$, ed a uno scalare qualsiasi. Poiche' $\mathbf{u}', \mathbf{v}' \in \Im f$ allora esistono vettori $\mathbf{u}, \mathbf{v} \in V$ tali che $\mathbf{u}' = f(\mathbf{u})$ e $\mathbf{v}' = f(\mathbf{v})$. Quindi:

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

Cio' implica che $\mathbf{u}' + \mathbf{v}' \in \Im f$. Similmente

$$a\mathbf{u}' = af(\mathbf{u}) = f(a\mathbf{u}) \in \Im f.$$

Quindi $\Im f$ e' stabile, dunque e' un sottospazio di V'.

2) E' proprio la definizione di applicazione suriettiva.

3) Sia $\mathbf{u}'$ un vettore di $\Im f$. Allora esiste un vettore $\mathbf{u} \in V$ tale che $\mathbf{u}' = f(\mathbf{u})$. Poiche' V e' generato dai vettori $\mathbf{u}_1, \ldots, \mathbf{u}_h$ potremo scrivere $\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h$ secondo opportuni scalari $a_1, \ldots, a_h$. Quindi

$$\mathbf{u}' = f(\mathbf{u}) = f(a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h) = a_1 f(\mathbf{u}_1) + \cdots + a_h f(\mathbf{u}_h).$$

Cio' prova che $\Im f$ e' generabile con i vettori $f(\mathbf{u}_1), \ldots, f(\mathbf{u}_h)$.

4) Per la proprieta' precedente, $\Im L_A$ e' generato dalle immagini dei vettori canonici di $\mathbf{R}^n$. L'asserto segue dal fatto che $L_A(\mathbf{e}_i) = A \cdot \mathbf{e}_i$ coincide con la colonna di posto i di A.

5) Ricordiamo che l'applicazione delle coordinate fa corrispondere sistemi liberi a sistemi liberi. Percio', poiche' i vettori numerici $A^{i_1}, \dots, A^{i_k}$ rappresentano i vettori delle coordinate dei vettori $f(\mathbf{b}_{i_1}), \dots, f(\mathbf{b}_{i_k})$, allora il sistema $\{f(\mathbf{b}_{i_1}), \dots, f(\mathbf{b}_{i_k})\}$ e' libero, in quanto tale e' $A^{i_1}, \dots, A^{i_k}$. Quindi rimane da provare che il sistema $\{f(\mathbf{b}_{i_1}), \dots, f(\mathbf{b}_{i_k})\}$ genera $\Im f$. Sia allora $\mathbf{u}' = f(\mathbf{u})$ un vettore dell'immagine. Sappiamo che

$$[\mathbf{u}']_{\mathcal{B}'} = [f(\mathbf{u})]_{\mathcal{B}'} = A \cdot [\mathbf{u}]_{\mathcal{B}}.$$

Quindi, se denotiamo con $\mathbf{x}$ il vettore delle coordinate di $\mathbf{u}$, abbiamo:

$$\begin{aligned}[\mathbf{u}']_{\mathcal{B}'} &= A \cdot (x_1, \dots, x_n)^T = A \cdot (x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) \\ &= x_1(A \cdot \mathbf{e}_1) + \dots + x_n(A \cdot \mathbf{e}_n) = x_1 A^1 + \dots + x_n A^n,\end{aligned}$$

dove con A^j abbiamo denotato le colonne di A. Poiche' $A^{i_1}, \dots, A^{i_k}$ generano le colonne di A, allora esistono opportuni scalari $z_1, \dots, z_k$ tali che

$$[\mathbf{u}']_{\mathcal{B}'} = x_1 A^1 + \dots + x_n A^n = z_1 A^{i_1} + \dots + z_k A^{i_k}.$$

D'altra parte

$$z_1 A^{i_1} + \dots + z_k A^{i_k} = z_1[f(\mathbf{b}_{i_1})]_{\mathcal{B}'} + \dots + z_k[f(\mathbf{b}_{i_k})]_{\mathcal{B}'} = [z_1 f(\mathbf{b}_{i_1}) + \dots + z_k f(\mathbf{b}_{i_k})]_{\mathcal{B}'}.$$

E poiche' l'applicazione delle coordinate e' biiettiva segue che

$$\mathbf{u}' = z_1 f(\mathbf{b}_{i_1}) + \dots + z_k f(\mathbf{b}_{i_k}).$$

La proprieta' 6) e' una diretta conseguenza della proprieta' precedente, e la proprieta' 7) segue combinando la 6) con le proprieta' del nucleo. $\square$

Osservazione. Il Teorema della dimensione si puo' dimostrare anche senza introdurre le coordinate, cioe' senza utilizzare la matrice rappresentativa. Una dimostrazione *intrinseca* e' la seguente. Sia $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h\}$ una base del nucleo di f, e sia $\{f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_k)\}$ una base dell'immagine. Il teorema della dimensione sara' provato se faremo vedere che il sistema

$$\mathcal{S} := \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$$

e' una base di V.

Cominciamo a provare che tale sistema e' libero.

A tale proposito, sia

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \dots + a_h\mathbf{u}_h + b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \dots + b_k\mathbf{v}_k = \mathbf{0}_V \quad (*)$$

una relazione tra i vettori di $\mathcal{S}$. Applicando l'operatore f ad entrambi i membri, e tenuto conto che f e' lineare, e che i vettori $\mathbf{u}_i$ sono vettori del nucleo, deduciamo:

$$b_1 f(\mathbf{v}_1) + b_2 f(\mathbf{v}_2) + \dots + b_k f(\mathbf{v}_k) = \mathbf{0}_{V'}.$$

Cio' implica che $b_1 = b_2 = \dots = b_k = 0$, in quanto i vettori $f(\mathbf{v}_j)$ sono indipendenti. Ma allora la relazione $(*)$ si riduce ad una relazione tra i vettori $\mathbf{u}_i$, che sono liberi. Percio' anche i pesi a_i sono tutti nulli. Cio' prova che il sistema $\mathcal{S}$ e' linearmente indipendente.

Infine, andiamo a provare che $\mathcal{S}$ genera V.

Sia $\mathbf{w}$ un qualunque vettore di V. Allora esistono pesi $b_1, b_2, \dots, b_k$ tali che

$$f(\mathbf{w}) = b_1 f(\mathbf{v}_1) + b_2 f(\mathbf{v}_2) + \dots + b_k f(\mathbf{v}_k).$$

Poiche' f e' lineare abbiamo:

$$f(\mathbf{w}) = f(b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k).$$

Portando tutto a sinistra, e applicando ancora la linearita' di f, deduciamo:

$$f(\mathbf{w} - (b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k)) = \mathbf{0}_{V'}.$$

Quindi il vettore $\mathbf{w} - (b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k)$ e' un vettore del nucleo. Percio' esistono pesi $a_1, a_2, \dots, a_h$ tali che

$$\mathbf{w} - (b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k) = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_h\mathbf{u}_h.$$

Da cui:

$$\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_h\mathbf{u}_h + b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k.$$

Esercizio. Sia $f : \mathbf{R}^4 \to \mathbf{R}^3$ *l'applicazione lineare definita ponendo* $f(x, y, z, t) = (x + t, x + y - z, x + t)$. *Calcolare una base per* $\ker f$ *e per* $\Im f$.

Svolgimento. L'espressione esplicita di f e' riferita alle basi canoniche, quindi f si identifica con l'applicazione L_A con

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

Il nucleo di f e' lo spazio delle soluzioni del sistema lineare omogeneo $A \cdot \mathbf{x} = \mathbf{0}_{\mathbf{R}^3}$. Cioe' del sistema lineare

$$\begin{cases} x + t = 0 \\ x + y - z = 0 \\ x + t = 0. \end{cases}$$

Risolvendo tale sistema vediamo che una base per $\ker f$ e' costituita dai vettori $(0, 1, 1, 0)$, $(-1, 1, 0, 1)$. Poiche' $\Im f$ e' lo spazio delle colonne di A, ne consegue che una base per $\Im f$ e' data dai vettori $(1, 1, 1)$, $(0, 1, 0)$. $\square$

Esercizio. Sia $f : \mathbf{R}^3 \to \mathbf{R}^3$ *l'applicazione lineare la cui espressione esplicita rispetto alla base* $\mathcal{B} = \{(1, 1, 1), (1, 2, 1), (2, 5, 3)\}$, *in partenza ed in arrivo, e' data da* $f(x', y', z') = (x' + z', y', x' + z')$. *Calcolare una base per* $\ker f$ *e per* $\Im f$.

Svolgimento. Ricordando che le coordinate sono riferite alla base $\mathcal{B}$, possiamo identificare la funzione f con L_A, dove $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$, cioe'

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

Il sistema lineare $A \cdot \mathbf{x} = \mathbf{0}$ ammette come soluzioni i vettori numerici $(-z', 0, z')$, al variare di $z' \in \mathbf{R}$. Percio' il nucleo di f ha dimensione 1 ed una base per $\ker f$ e' data dal vettore che ha coordinate $(-1, 0, 1)^T$ rispetto alla base $\mathcal{B}$, cioe' dal vettore:

$$(1, 4, 2) = (-1) \cdot (1, 1, 1) + 0 \cdot (1, 2, 1) + 1 \cdot (2, 5, 3).$$

Lo spazio delle colonne di A ammette come base i vettori colonna $(1, 0, 1)^T$, $(0, 1, 0)^T$. Quindi una base per $\Im f$ e' data dai vettori $f(1, 1, 1)$ ed $f(1, 2, 1)$, cioe' dai vettori $(3, 6, 4)$, $(1, 2, 1)$.

In conclusione, una base per $\ker f$ e' data dal vettore $(1, 4, 2)$, ed una base per $\Im f$ e' formata dai vettori $(3, 6, 4)$, $(1, 2, 1)$. $\square$

Osservazione. Le rotazioni e le riflessioni hanno nucleo banale, ed immagine tutto il piano. Invece la proiezione ortogonale sulla retta r ammette come nucleo i vettori giacenti sulla retta ortogonale ad r, e come immagine i vettori della retta r.

Esercizio. Sia f un'applicazione lineare del tipo $f : V \to V$ [4]*. Per ogni intero $i \geq 0$, si consideri l'applicazione lineare $f^i : V \to V$ ottenuta componendo f con se stessa i volte. Provare che esistono le seguenti sequenze:*

$$\{\mathbf{0}\} = \ker f^0 \subseteq \ker f \subseteq \ker f^2 \subseteq \dots \subseteq \ker f^i \subseteq \dots$$

$$V = \Im f^0 \supseteq \Im f \supseteq \Im f^2 \supseteq \dots \supseteq \Im f^i \supseteq \dots .$$

Svolgimento. Se $\mathbf{u} \in \ker f^i$ allora $f^i(\mathbf{u}) = \mathbf{0}$. Quindi $f(f^i(\mathbf{u})) = \mathbf{0}$, cioe' $f^{i+1}(\mathbf{u}) = \mathbf{0}$. Il che implica $\mathbf{u} \in \ker f^{i+1}$.

Se $\mathbf{u} \in \Im f^{i+1}$ allora esiste un vettore $\mathbf{v}$ tale che $f^{i+1}(\mathbf{v}) = \mathbf{u}$. Quindi $f(f^i(\mathbf{v})) = \mathbf{u}$. Il che implica $\mathbf{u} \in \Im f^i$. $\square$

Osservazione. Sia $f : V \to V$ un'applicazione lineare. Sia $\mathbf{u} \in \Im f^i$. Allora esiste un vettore $\mathbf{v}$ tale che $f^i(\mathbf{v}) = \mathbf{u}$. Quindi $f(\mathbf{u}) = f^{i+1}(\mathbf{v}) = f^i(f(\mathbf{v}))$. Percio' $f(\mathbf{u}) \in \Im f^i$. In altre parole, l'applicazione f invia i vettori di $\Im f^i$ in $\Im f^i$. Percio', per restrizione, l'applicazione f induce un'applicazione $\varphi_i : \Im f^i \to \Im f^i$, dove

$$\varphi_i(\mathbf{u}) := f(\mathbf{u}).$$

Naturalmente, tale applicazione e' ancora lineare. Si osservi che $\ker \varphi_i = \Im f^i \cap \ker f$, e che $\Im \varphi_i = \Im f^{i+1}$.

Esercizio. Sia $f : V \to V$ un'applicazione lineare. Provare che, per ogni intero $i \geq 0$, si ha:

$$\dim(\Im f^i \cap \ker f) = \dim \Im f^i - \dim \Im f^{i+1}.$$

Svolgimento. Segue dal Teorema della dimensione, applicato all'operatore $\varphi_i : \Im f^i \to \Im f^i$. $\square$

7. Gli isomorfismi.

Un'applicazione lineare *biiettiva* $f : V \to V'$ si dice *isomorfismo tra lo spazio V e lo spazio V'.*

• *Le proprieta' degli isomorfismi.*

Sia $f : V \to V'$ un isomorfismo tra spazi vettoriali finitamente generabili. Allora valgono le seguenti proprieta':

1) *l'applicazione inversa $f^{-1} : V' \to V$ e' anch'essa lineare e quindi e' un isomorfismo tra V' e V;*

2) *un sistema di vettori $\mathbf{b}_1, \dots, \mathbf{b}_n$ e' una base di V se e solo se il sistema di vettori $f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)$ e' una base di V' (cioe' un isomorfismo trasforma basi in basi);*

3) $\dim(V) = \dim(V')$*;*

[4]Una applicazione lineare del tipo $f : V \to V$, in cui lo spazio di partenza coincide con quello di arrivo, dicesi *endomorfismo di V*

4) *se* $\mathcal{B}$ *e' una base di* V *e* $\mathcal{B}'$ *e' una base di* V' *allora la matrice rappresentativa* $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ *e' invertibile e*

$$M^{\mathcal{B}}_{\mathcal{B}'}(f)^{-1} = M^{\mathcal{B}'}_{\mathcal{B}}(f^{-1}).$$

Dimostrazione. Le proprieta' 1) e 4) gia' ci sono note. La 3) segue da 2). Percio' occorre solo provare la proprieta' 2).

Poiche' f e' iniettiva allora sappiamo che $f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)$ e' libero. Inoltre sappiamo che $f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)$ genera $\Im f$, che e' tutto V' perche' f e' suriettiva. Cio' prova che se $\mathbf{b}_1, \dots, \mathbf{b}_n$ e' una base di V allora $f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)$ e' una base di V'. Il viceversa segue dallo stesso ragionamento applicato ad f^{-1}. $\square$

Due spazi vettoriali V e V' si dicono *isomorfi* se esiste un isomorfismo $f : V \to V'$ tra lo spazio V e lo spazio V'. Possiamo riguardare la nozione di isomorfismo come una relazione nell'insieme di tutti gli spazi vettoriali. Tale relazione e' una relazione di equivalenza. Infatti ogni spazio vettoriale V e' isomorfo a se stesso in virtu' dell'applicazione identica $\mathrm{id}_V : V \to V$, che e' un isomorfismo. Poi la proprieta' 1) precedente ci dice che tale relazione e' anche simmetrica. Infine se $f : V \to V'$ e $g : V' \to V$" sono isomorfismi allora tale e' anche l'applicazione composta $g \circ f : V \to V$". Quindi la relazione di isomorfismo e' anche una relazione transitiva.

Un esempio importante di isomorfismo e' l'applicazione delle coordinate $[\ \ ]_{\mathcal{B}}$. Assegnata una base $\mathcal{B}$ in uno spazio vettoriale V di dimensione n, tale applicazione e' quella che associa al vettore $\mathbf{u}$ di V il vettore numerico $\mathbf{x} \in \mathbf{R}^n$ delle coordinate di $\mathbf{u}$ rispetto alla base $\mathcal{B}$:

$$[\ \ ]_{\mathcal{B}} : \mathbf{u} \in V \to \mathbf{x} = [\mathbf{u}]_{\mathcal{B}} \in \mathbf{R}^n.$$

L'esistenza di tale isomorfismo consente di dedurre il seguente

Teorema. Ogni spazio vettoriale di dimensione n *e' isomorfo ad* $\mathbf{R}^n$.

Per transitivita' otteniamo il corollario

Corollario. Due spazi vettoriali sono isomorfi se e solo se hanno la stessa dimensione.

Sia $\mathcal{V}$ l'insieme di tutti gli spazi vettoriali di dimensione finita, e sia $\widetilde{\mathcal{V}}$ l'insieme di tutte le classi di equivalenza rispetto alla relazione di isomorfismo in $\mathcal{V}$. Per ogni spazio $V \in \mathcal{V}$ denotiamo con $[V] \in \widetilde{\mathcal{V}}$ la sua classe di equivalenza. In base al corollario precedente la seguente applicazione

$$[V] \in \widetilde{\mathcal{V}} \to \dim(V) \in \mathbf{N}_0$$

e' ben definita ed e' biiettiva. In altre parole, *a meno di isomorfismi, ci sono tanti spazi vettoriali di dimensione finita quanti sono i numeri naturali.*

8. Lo spazio delle applicazioni lineari.

• *Lo spazio Hom*(V, V').

Siano V e V' spazi vettoriali, e denotiamo con

$$\mathrm{Hom}(V, V')$$

l'insieme di tutte le applicazioni lineari da V a V'. Se f e g sono due applicazioni lineari tra V e V', e $c \in \mathbf{R}$ e' uno scalare, possiamo definire

$$f + g : V \to V'$$

come l'applicazione che trasforma $\mathbf{u}$ in $f(\mathbf{u}) + g(\mathbf{u})$, cioe'

$$(f + g)(\mathbf{u}) := f(\mathbf{u}) + g(\mathbf{u}),$$

e

$$c \cdot f : V \to V'$$

come l'applicazione che trasforma $\mathbf{u}$ in $c \cdot f(\mathbf{u})$, cioe'

$$(c \cdot f)(\mathbf{u}) := c \cdot f(\mathbf{u}).$$

In precedenza abbiamo provato che, poiche' f e g sono lineari, anche $f + g$ e $c \cdot f$ lo sono. Pertanto queste due operazioni definiscono una struttura algebrica

$$(\mathrm{Hom}(V, V'), +, \cdot).$$

Si puo' provare che tale struttura e' uno spazio vettoriale. In tale spazio vettoriale il vettore nullo e' l'applicazione nulla $\mathbf{u} \in V \to \mathbf{0}_{V'} \in V'$. E la funzione opposta di f e' l'applicazione $-f : \mathbf{u} \in V \to -f(\mathbf{u}) \in V'$. Andiamo a provare che:

• *Gli spazi Hom*(V, V') *ed* $\mathcal{M}(m, n)$ *sono isomorfi.*

Se $\mathcal{B}$ e $\mathcal{B}'$ sono basi di V e V', allora

$$M^{\mathcal{B}}_{\mathcal{B}'}(f + g) = M^{\mathcal{B}}_{\mathcal{B}'}(f) + M^{\mathcal{B}}_{\mathcal{B}'}(g)$$

e

$$M^{\mathcal{B}}_{\mathcal{B}'}(c \cdot f) = c \cdot M^{\mathcal{B}}_{\mathcal{B}'}(f).$$

In altre parole l'applicazione

$$M^{\mathcal{B}}_{\mathcal{B}'} : f \in \mathrm{Hom}(V, V') \to M^{\mathcal{B}}_{\mathcal{B}'}(f) \in \mathcal{M}(m, n)$$

e' un'applicazione lineare (qui naturalmente n e' la dimensione di V, ed m quella di V'). Questa applicazione e' in realta' un isomorfismo, in quanto e' biiettiva. Per provare che e' biiettiva possiamo costruire direttamente la sua inversa

$$f^{\mathcal{B}}_{\mathcal{B}'} : \mathcal{M}(m, n) \to \mathrm{Hom}(V, V')$$

ponendo

$$f^{\mathcal{B}}_{\mathcal{B}'}(A) := [\quad]^{-1}_{\mathcal{B}'} \circ L_A \circ [\quad]_{\mathcal{B}},$$

dove con L_A denotiamo l'applicazione

$$L_A : \mathbf{x} \in \mathbf{R}^n \to A \cdot \mathbf{x} \in \mathbf{R}^m.$$

Poiche'

$$f^{\mathcal{B}}_{\mathcal{B}'}(M^{\mathcal{B}}_{\mathcal{B}'}(f)) = f \quad \text{e} \quad M^{\mathcal{B}}_{\mathcal{B}'}(f^{\mathcal{B}}_{\mathcal{B}'}(A)) = A$$

allora

$$f^{\mathcal{B}}_{\mathcal{B}'} \circ M^{\mathcal{B}}_{\mathcal{B}'} = \mathrm{id}_{\mathrm{Hom}(V,V')} \quad \text{e} \quad M^{\mathcal{B}}_{\mathcal{B}'} \circ f^{\mathcal{B}}_{\mathcal{B}'} = \mathrm{id}_{\mathcal{M}(m,n)}.$$

Ne consegue che $M^{\mathcal{B}}_{\mathcal{B}'}$ e' un isomorfismo.

L'esistenza dell'isomorfismo $M^{\mathcal{B}}_{\mathcal{B}'}$ significa che lo studio delle applicazioni lineari e' equivalente allo studio delle matrici. Tuttavia occorre tener presente che l'isomorfismo $M^{\mathcal{B}}_{\mathcal{B}'}$ non e' naturale, nel senso che se si cambiano le basi allora cambia l'isomorfismo. Quando $V' = \mathbf{R}$, lo spazio $\mathrm{Hom}(V, \mathbf{R})$ si chiama *il duale di* V e riveste un grande interesse nell'algebra lineare, nella ricerca di argomenti dimostrativi che non utilizzino l'applicazione delle coordinate.

- *L'algebra degli endomorfismi End*(V).

Quando $V = V'$ allora su $\mathrm{Hom}(V, V)$, oltre all'addizione interna ed alla moltiplicazione esterna, abbiamo anche una moltiplicazione interna data dalla composizione. Cioe', cosi' come $\mathcal{M}(n, n)$, $\mathrm{Hom}(V, V)$ e' una algebra:

$$(\mathrm{Hom}(V, V), +, \cdot, \circ),$$

detta anche *l'algebra degli endomorfismi di* V, e denotata anche $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ (un'applicazione lineare del tipo $f : V \to V$ si dice anche *endomorfismo di* V). Poiche' la matrice rappresentativa di $g \circ f$ e' il prodotto della matrice rappresentativa di g per quella di f, allora l'isomorfismo di spazi vettoriali (adesso assumiamo $\mathcal{B} = \mathcal{B}'$, ed $n = \dim V$)

$$M^{\mathcal{B}}_{\mathcal{B}} : \mathrm{End}(V) \to \mathcal{M}(n, n),$$

oltre ad essere compatibile rispetto all'addizione ed alla moltiplicazione esterna, e' compatibile anche rispetto alla moltiplicazione interna. Cioe'

$$M^{\mathcal{B}}_{\mathcal{B}}(g \circ f) = M^{\mathcal{B}}_{\mathcal{B}}(g) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f).$$

Cio' si esprime dicendo che $M^{\mathcal{B}}_{\mathcal{B}}$ e' un isomorfismo di algebre.

- *La matrice rappresentativa di un endomorfismo e la relazione di similitudine.*

Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale V, che assumiamo di dimensione finita n. In questo caso particolare possiamo fissare una stessa base $\mathcal{B}$ in partenza ed in arrivo, e quindi considerare la matrice rappresentativa $M^{\mathcal{B}}_{\mathcal{B}}(f)$, che e' una matrice quadrata $n \times n$. Ci si puo' chiedere come varia tale matrice al variare della base $\mathcal{B}$. La risposta e' fornita dalla seguente formula:

$$M^{\mathcal{B}'}_{\mathcal{B}'}(f) = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V),$$

a cui si perviene facilmente osservando che

$$f = \mathrm{id}_V \circ f \circ \mathrm{id}_V$$

e ricordando che la matrice rappresentativa di una composizione di applicazioni lineari e' il prodotto delle matrici rappresentative delle singole applicazioni. Possiamo scrivere la formula precedente in modo piu' semplice se poniamo $P = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V)$, $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$, e $A' = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$:

$$A' = P^{-1} \cdot A \cdot P.$$

Questa formula e' alla base della seguente definizione. Date due matrici quadrate A ed A' di ordine n, diremo che *A e' simile (o coniugata) ad A' se esiste una matrice invertibile P tale che $A' = P^{-1} \cdot A \cdot P$.* Per quanto detto in precedenza, tutte le matrici rappresentative di un dato endomorfismo f, della forma $M^{\mathcal{B}}_{\mathcal{B}}(f)$, al variare della base $\mathcal{B}$, sono simili tra loro.

Osservazione. (i) La relazione di similitudine e' una relazione di equivalenza nell'insieme delle matrici quadrate di ordine n. Infatti, data una matrice quadrata A, tenuto conto che $I = I^{-1}$, possiamo scrivere $A = I \cdot A \cdot I$, percio' la relazione di similitudine e' riflessiva. E' anche simmetrica, infatti se $A' = P^{-1} \cdot A \cdot P$ allora $A = P \cdot A' \cdot P^{-1}$, cioe' se A e' simile ad A' allora A' e' simile ad A. La relazione e' anche transitiva. Infatti, se A e' simile ad A' ed A' e' simile ad A'', allora, secondo opportune matrici invertibili P e Q potremo scrivere $A' = P^{-1} \cdot A \cdot P$ e $A'' = Q^{-1} \cdot A' \cdot Q$. Sostituendo otterremo $A'' = (PQ)^{-1} \cdot A \cdot (PQ)$, cioe' A e' simile ad A''.

(ii) Siano A ed A' due matrici quadrate di ordine n, sia V uno spazio vettoriale di dimensione n. Allora *A e' simile ad A' se e solo se esiste un endomorfismo $f : V \to V$ ed esistono basi $\mathcal{B}$ e $\mathcal{B}'$ di V tali che $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ ed $A' = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$.* Per provare cio', assumiamo che A sia simile ad A', cioe' che esista una matrice invertibile P tale che $A' = P^{-1} \cdot A \cdot P$. Fissata una base qualsiasi $\mathcal{B}$ di V, sia $f = f^{\mathcal{B}}_{\mathcal{B}}(A)$. Gia' sappiamo che $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Adesso sia $\mathcal{B}'$ quella base di V formata da quei vettori che hanno per coordinate rispetto alla base $\mathcal{B}$ le colonne della matrice P ($\mathcal{B}'$ e' una base perche' P e' invertibile, percio' le sue colonne sono una base di $\mathbf{R}^n$, e l'applicazione delle coordinate, in quanto isomorfismo, porta basi in basi). Percio' abbiamo:

$$M^{\mathcal{B}'}_{\mathcal{B}'}(f) = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) = P^{-1} \cdot A \cdot P = A'.$$

Cio' conclude la dimostrazione (il viceversa ci e' gia' noto).

(iii) Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale di dimensione n, e sia $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ una matrice rappresentativa di f. Denotiamo con $[A]$ la classe di equivalenza di A rispetto alla relazione di similitudine. Cioe' $[A]$ sia l'insieme di tutte le matrici simili ad A. Con lo stesso argomento dell'osservazione precedente (ii), possiamo provare che *$[A]$ e' uguale all'insieme di tutte le matrici rappresentative di f, del tipo $M^{\mathcal{C}}_{\mathcal{C}}(f)$, al variare di $\mathcal{C}$ nell'insieme delle basi di V*. Cioe' che:

$$[A] = \left\{ M^{\mathcal{C}}_{\mathcal{C}}(f) \,:\, \mathcal{C} \text{ base di } V \right\}.$$

• *La sostituzione di una matrice e di un endomorfismo in un polinomio.*

Sia

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_h t^h$$

un polinomio. Dato uno scalare c, ha senso calcolare il numero $p(c)$, in quanto, per definizione, il polinomio $p(t)$ e' una funzione $p(t) : \mathbf{R} \to \mathbf{R}$. Tale numero si dice anche *la sostituzione di c in $p(t)$*, ed abbiamo:

$$p(c) = a_0 + a_1 c + a_2 c^2 + \cdots + a_h c^h \in \mathbf{R}.$$

Per esempio, se $p(t) = 2 - t - 3t^2$ allora $p(-1) = 0$, $p(2) = -12$.

Poiche' sia $\mathcal{M}(n, n)$ che $\mathrm{Hom}(V, V)$ sono algebre, ha senso sostituire in $p(t)$ sia una matrice quadrata A, che un endomorfismo $f : V \to V$. Cioe' possiamo definire

$$p(A) := a_0 I + a_1 A + a_2 A^2 + \cdots + a_h A^h \in \mathcal{M}(n, n),$$

e

$$p(f) := a_0 id_V + a_1 f + a_2 f^2 + \cdots + a_h f^h \in \mathrm{Hom}(V, V)$$

(con f^r si intende la composizione di f con se stesso $f \circ f \circ \cdots \circ f$ r volte). In altre parole, la funzione $p(f) : V \to V$ e' quell'operatore che al vettore $\mathbf{u} \in V$ associa il vettore $p(f)(\mathbf{u}) = a_0 \mathbf{u} + a_1 f(\mathbf{u}) + \cdots + a_h f^h(\mathbf{u})$, essendo $f^r := f \circ f \circ \cdots \circ f$ la composizione di f

con se' stesso r volte. In virtu' dell'isomorfismo di algebre descritto in precedenza, si ha anche
$$M_{\mathcal{B}}^{\mathcal{B}}(p(f)) = p(M_{\mathcal{B}}^{\mathcal{B}}(f))$$
dove $\mathcal{B}$ denota, come al solito, una qualunque base fissata di V.

Da questa uguaglianza deduciamo che se A e B sono matrici quadrate simili, allora, per ogni polinomio $p(t)$, lo sono anche $p(A)$ e $p(B)$. Infatti se A e B sono simili allora possiamo pensare ad A e B come matrici che rappresentano uno stesso operatore f, e la formula precedente ci dice che $p(A)$ e $p(B)$ rappresentano lo stesso operatore $p(f)$, e dunque $p(A)$ e $p(B)$ sono simili. Una dimostrazione diretta di questo fatto consiste nell'osservare che se $B = P^{-1}AP$ allora per ogni h si ha $B^h = P^{-1}A^hP$, e dunque $p(B) = P^{-1}p(A)P$.

Poiche' se due matrici sono simili hanno lo stesso rango (in quanto tale rango altro non e' che la dimensione dell'immagine dell'operatore f che rappresentano, che certo non dipende dalle coordinate) allora (se A e B sono simili) anche $p(A)$ e $p(B)$ hanno lo stesso rango. In particolare se A e' simile a B allora per ogni intero h il rango di A^h coincide con il rango di B^h, e, per un dato polinomio $p(t)$, $p(A) = \mathbf{0}$ se e solo se $p(B) = \mathbf{0}$.

Si osservi infine che comunque si assegnino polinomi $p(t)$ e $q(t)$, una matrice A ed un operatore f, valgono le seguenti proprieta' di calcolo:
$$(p+q)(A) = p(A) + q(A), \quad (p \cdot q)(A) = p(A) \cdot q(A) = q(A) \cdot p(A)$$
$$(p+q)(f) = p(f) + q(f), \quad (p \cdot q)(f) = p(f) \cdot q(f) = q(f) \cdot p(f)$$
(con il simbolo $p \cdot q$ si intende il prodotto dei due polinomi $p(t)$ e $q(t)$).

9. Matrici ed endomorfismi diagonalizzabili.

Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale V, $\dim V = n$. Al variare della base $\mathcal{B}$ di V, otteniamo le matrici rappresentative $M_{\mathcal{B}}^{\mathcal{B}}(f)$ di f, in generale tutte diverse tra loro. Una domanda che sorge naturale e' la seguente: tra tutte queste matrici e' possibile trovarne una semplice? Studieremo questo problema, in tutta generalita', nel prossimo capitolo. Adesso ci poniamo una domanda piu' precisa: tra tutte queste matrici e' possibile trovarne una che sia diagonale? Cioe', e' possibile rappresentare un endomorfismo con una matrice diagonale? Questo problema risulta di grande interesse sia in Matematica, che in Fisica ed in Ingegneria. Fra poco faremo qualche esempio. In generale la risposta non e' affermativa. In questo paragrafo studieremo sotto quali condizioni esiste una matrice rappresentativa diagonale, e come si trova. Alla fine perverremo ad un *algoritmo per la diagonalizzazione di un endomorfismo.*

Cominciamo con alcune definizioni, e degli esempi.

• *La definizione di endomorfismo diagonalizzabile.*

Un endomorfismo $f : V \to V$ si dice diagonalizzabile se esiste una base $\mathcal{A}$ di V tale che la matrice rappresentativa $M_{\mathcal{A}}^{\mathcal{A}}(f)$ di f sia una matrice diagonale.

Accanto alla definizione di endomorfismo diagonalizzabile, c'e' anche la definizione di matrice diagonalizzabile.

• *La definizione di matrice diagonalizzabile.*

Una matrice quadrata si dice diagonalizzabile se e' simile ad una matrice diagonale.

Osservazione. Le due definizioni precedenti sono in un certo senso "equivalenti". Infatti *se $f : V \to V$ e' un endomorfismo, ed $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ e' una matrice che rappresenta f*

rispetto ad una base $\mathcal{B}$ qualsiasi, allora A e' diagonalizzabile se e solo se lo e' f. In particolare, una matrice quadrata A di ordine n e' diagonalizzabile se e solo se lo e' l'operatore $L_A : \mathbf{x} \in \mathbf{R}^n \to A \cdot \mathbf{x} \in \mathbf{R}^n$. Percio' lo studio degli operatori diagonalizzabili e' equivalente a quello delle matrici diagonalizzabili. E' chiaro che se f e' diagonalizzabile allora lo e' anche A. Infatti sia $\mathcal{A}$ una base tale che $M^{\mathcal{A}}_{\mathcal{A}}(f) = D$ e' diagonale. Poiche'

$$M^{\mathcal{A}}_{\mathcal{A}}(f) = M^{\mathcal{B}}_{\mathcal{A}}(\mathrm{id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id}_V),$$

allora A e' simile a D e percio' A e' diagonalizzabile. Viceversa, se A e' diagonalizzabile esistono una matrice diagonale D ed una matrice invertibile P tali che $D = P^{-1} \cdot A \cdot P$. Sia $\mathcal{A}$ quella base di V formata dai vettori che hanno per coordinate rispetto alla base $\mathcal{B}$ le colonne della matrice P. Allora $D = M^{\mathcal{A}}_{\mathcal{A}}(f)$ e percio' f e' diagonalizzabile.

Esempi.

1) Nello spazio dei vettori geometrici del piano $\mathcal{V}_{O,\rho}$, consideriamo la riflessione $f : \mathcal{V}_{O,\rho} \to \mathcal{V}_{O,\rho}$ rispetto alla retta r. Se scegliamo una base $\mathcal{A} = \{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$ in modo tale che il vettore $\overrightarrow{OE_1}$ stia sulla retta r, ed il vettore $\overrightarrow{OE_2}$ sia ortogonale ad r, allora l'angolo con cui appare r rispetto all'asse $\overrightarrow{OE_1}$ e' $\alpha = 0$, e percio' la matrice rappresentativa di f e' diagonale:

$$M^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Quindi la riflessione rispetto ad una retta e' un endomorfismo diagonalizzabile del piano.

2) Consideriamo l'operatore $f : \mathbf{R}^3 \to \mathbf{R}^3$ definito ponendo

$$f(x, y, z) = (3x + 3y - z, -4x - 2y + 2z, -2x + 2y + 2z).$$

La matrice rappresentativa di f rispetto alla base canonica e'

$$M^{\mathcal{E}}_{\mathcal{E}}(f) = \begin{bmatrix} 3 & 3 & -1 \\ -4 & -2 & 2 \\ -2 & 2 & 2 \end{bmatrix}.$$

Ora consideriamo la seguente base di $\mathbf{R}^3$

$$\mathcal{A} := \{(-2, 1, -3), (1, 0, 2), (1, 1, 4)\}.$$

Poiche'

$$f(-2, 1 - 3) = (0, 0, 0), \quad f(1, 0, 2) = (1, 0, 2), \quad f(1, 1, 4) = 2(1, 1, 4)$$

allora

$$M^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

che e' una matrice diagonale. Percio' l'operatore f e' diagonalizzabile, ed $\mathcal{A}$ e' una base rispetto alla quale la matrice rappresentativa e' diagonale. L'espressione esplicita di f con le nuove coordinate e'

$$f(x', y', z') = (0, y', 2z'),$$

che e' molto piu' semplice dell'espressione esplicita rispetto alle coordinate canoniche.

3) La matrice

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

e' diagonalizzabile. Infatti, in base all'esempio precedente, si ha:

$$M_{\mathcal{A}}^{\mathcal{A}}(f) = M_{\mathcal{A}}^{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^3}) \cdot M_{\mathcal{E}}^{\mathcal{E}}(f) \cdot M_{\mathcal{E}}^{\mathcal{A}}(\mathrm{id}_{\mathbf{R}^3})$$

cioe'

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = P^{-1} \cdot \begin{bmatrix} 3 & 3 & -1 \\ -4 & -2 & 2 \\ -2 & 2 & 2 \end{bmatrix} \cdot P,$$

dove

$$P = M_{\mathcal{E}}^{\mathcal{A}}(\mathrm{id}_{\mathbf{R}^3}) = \begin{bmatrix} -2 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 2 & 4 \end{bmatrix}$$

e' la matrice del cambiamento delle coordinate.

4) Supponiamo che in un certo problema ci sia il bisogno di calcolare la potenza

$$A^h$$

della matrice A dell'esempio precedente. Posto

$$D := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

poiche' $A = P \cdot D \cdot P^{-1}$, allora

$$A^h = P \cdot D^h \cdot P^{-1}.$$

Questa formula ci dice che il calcolo di A^h si riconduce al calcolo della matrice P, e della potenza D^h, che, essendo D diagonale, si calcola immediatamente:

$$D^h = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^h \end{bmatrix}.$$

In conclusione

$$A^h = P \cdot D^h \cdot P^{-1} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^h \end{bmatrix} \cdot \begin{bmatrix} 2 & 2 & -1 \\ 7 & 5 & -3 \\ -2 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 7-2^{h+1} & 5-2^h & 2^h-3 \\ -2^{h+1} & -2^h & 2^h \\ 14-2^{h+3} & 10-2^{h+2} & 2^{h+2}-6 \end{bmatrix}.$$

Osserviamo di nuovo che per il calcolo di A^h non basta sapere che una certa matrice diagonale D e' simile ad A, ma e' necessario anche conoscere la matrice invertibile P che realizza la similitudine tra A e D. Ne' servirebbe diagonalizzare f con una matrice del tipo $M_{\mathcal{A}'}^{\mathcal{A}}(f)$, cioe' con basi differenti, perche' verrebbe meno la formula $A^h = P \cdot D^h \cdot P^{-1}$.

5) Come sara' chiaro in seguito, la matrice

$$\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

e' un esempio di matrice non diagonalizzabile.

Cio' premesso, ora andiamo a vedere sotto quali condizioni un operatore e' diagonalizzabile.

Supponiamo innanzitutto che $f : V \to V$ sia un operatore diagonalizzabile. Allora esiste una base $\mathcal{A} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ tale che la matrice rappresentativa $M_{\mathcal{A}}^{\mathcal{A}}(f)$ sia diagonale. Denotiamo con $(\lambda_1, \dots, \lambda_n)$ le componenti sulla diagonale principale di tale matrice, cioe' poniamo

$$M_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix}.$$

Per definizione di matrice rappresentativa abbiamo

$$\begin{cases} f(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1 \\ f(\mathbf{b}_2) = \lambda_2 \mathbf{b}_2 \\ \dots \\ f(\mathbf{b}_n) = \lambda_n \mathbf{b}_n. \end{cases}$$

Cio' motiva la seguente definizione.

• *La definizione di autovalore ed autovettore per un endomorfismo.*

Un numero $\lambda \in \mathbf{R}$ si dice autovalore per l'endomorfismo $f : V \to V$ se esiste un vettore non nullo $\mathbf{b} \neq \mathbf{0}$ di V tale che $f(\mathbf{b}) = \lambda \mathbf{b}$. Un siffatto vettore $\mathbf{b}$ si dice autovettore per f relativo all'autovalore λ.

La richiesta $\mathbf{b} \neq \mathbf{0}$ e' dovuta al fatto che altrimenti la definizione sarebbe banale, in quanto ogni numero sarebbe un autovalore poiche' $f(\mathbf{0}) = \mathbf{0} = \lambda \cdot \mathbf{0}$ per ogni λ. Inoltre abbiamo visto che se f e' diagonalizzabile, allora esiste una base di autovettori per f, che percio' devono essere vettori non nulli. Anzi, in base a quanto detto prima, e' chiaro che un *endomorfismo $f : V \to V$ e' diagonalizzabile se e solo se esiste una base di V formata da autovettori per f. In tal caso, detta $\mathcal{A}$ tale base, la matrice rappresentativa $M_{\mathcal{A}}^{\mathcal{A}}(f)$ e' diagonale, e sulla diagonale principale appaiono gli autovalori corrispondenti agli autovettori di $\mathcal{A}$.* In particolare, possiamo dire che condizione necessaria affinche' un endomorfismo sia diagonalizzabile e' che esso ammetta qualche autovettore. Gia' questa osservazione e' sufficiente per dire che una rotazione non e' diagonalizzabile, a meno che l'angolo di rotazione non sia $\alpha = 0$ oppure $\alpha = \pi$. Infatti un autovettore e' un vettore non nullo che viene trasformato in un vettore parallelo a se stesso, e questo non puo' accadere per una rotazione. In altre parole, in generale una rotazione non ha ne' autovalori ne' autovettori.

Ora andiamo a vedere come riconoscere la presenza di autovettori, e di autovalori, per un dato endomorfismo.

Fissiamo un endomorfismo $f : V \to V$, una base $\mathcal{B}$ qualsiasi, e sia $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ la matrice rappresentativa. Sia $\lambda \in \mathbf{R}$ un numero qualsiasi. Abbiamo detto che λ e' un autovalore per $f : V \to V$ se esiste un vettore non nullo $\mathbf{b} \neq \mathbf{0}$ di V tale che $f(\mathbf{b}) = \lambda \mathbf{b}$. Poiche'

$$[f(\mathbf{b})]_{\mathcal{B}} = A \cdot [\mathbf{b}]_{\mathcal{B}}$$

cio' equivale a dire che esiste un vettore numerico $\mathbf{x} \in \mathbf{R}^n$ non nullo (qui $n = \dim V$), tale che

$$A \cdot \mathbf{x} = \lambda \cdot \mathbf{x}.$$

Cioe' che esiste una soluzione $\mathbf{x}$ non nulla del sistema lineare omogeneo

$$(A - \lambda \cdot I) \cdot \mathbf{x} = \mathbf{0}.$$

Dalla teoria dei sistemi lineari sappiamo che lo spazio delle soluzioni di tale sistema ha dimensione pari a $n - rk(A - \lambda \cdot I)$. Quindi l'esistenza di una soluzione non nulla equivale a dire che tale dimensione e' positiva, cioe' che $rk(A - \lambda \cdot I) < n$. Tenuto conto che la matrice $A - \lambda \cdot I$ e' quadrata di ordine n, in definitiva dire che λ e' un autovalore per f equivale a dire che

$$\det(A - \lambda \cdot I) = 0.$$

Se consideriamo il numero λ come una variabile, il determinante $\det(A - \lambda \cdot I)$ e' un polinomio nella variabile λ. La discussione precedente motiva la seguente definizione.

- *La definizione di polinomio caratteristico di un endomorfismo.*

Sia $f : V \to V$ un endomorfismo, sia $\mathcal{B}$ una base di V, e sia $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ la matrice rappresentativa. Si definisce polinomio caratteristico di f il polinomio:

$$p_f(t) := \det(A - t \cdot I).$$

Per quanto appena detto possiamo dire che *gli autovalori di f sono esattamente le radici del suo polinomio caratteristico, cioe' sono tutti e soli quei numeri $\lambda \in \mathbf{R}$ per cui $p_f(\lambda) = 0$.*

Osservazione. (i) Se n e' la dimensione di V, allora il polinomio caratteristico $p_f(t)$ di f e' un polinomio di grado n, ed ha una forma del tipo:

$$p_f(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + (-1)^n t^n.$$

In particolare il polinomio caratteristico e' monico, nel senso che il suo coefficiente direttore e' ± 1. Naturalmente $a_0 = \det A$.

(ii) Il polinomio caratteristico di f non dipende dalla base fissata, in particolare i suoi coefficienti sono *invarianti* per cambiamento di coordinate. Infatti se A' e' un'altra matrice rappresentativa di f, sappiamo che A' ed A sono simili, cioe' esiste una matrice invertibile P tale che $A' = P^{-1} \cdot A \cdot P$. E allora, tenuto conto del Teorema di Binet e del fatto che $\det P^{-1} \cdot \det P = 1$, abbiamo:

$$\begin{aligned} \det(A' - t \cdot I) &= \det((P^{-1} \cdot A \cdot P) - t \cdot I) = \det(P^{-1} \cdot (A' - t \cdot I) \cdot P) \\ &= \det P^{-1} \cdot \det(A - t \cdot I) \cdot \det P = \det(A - t \cdot I). \end{aligned}$$

(iii) Il polinomio $\det(A - t \cdot I)$ si chiama anche polinomio caratteristico della matrice A, e lo denoteremo anche con $p_A(t)$. Abbiamo appena dimostrato che matrici simili hanno lo stesso polinomio caratteristico. E' chiaro che $p_A(t)$ coincide con il polinomio caratteristico di un qualunque endomorfismo rappresentabile con la matrice A.

(iv) L'insieme $\mathrm{spec}(f)$ degli autovalori di f si chiama *lo spettro di f*. Quindi abbiamo:

$$\mathrm{spec}(f) = \{\lambda \in \mathbf{R} : \lambda \text{ e' un autovalore per } f\} = \{\lambda \in \mathbf{R} : p_f(\lambda) = 0\}.$$

(v) Poiche' il polinomio caratteristico ha grado $n = \dim V$, allora f puo' avere al massimo n autovalori distinti, perche' un polinomio di grado n puo' avere al massimo n radici.

(vi) Se f e' diagonalizzabile, denotata con $\mathcal{A}$ una base di autovettori per f, per calcolare il polinomio caratteristico di f possiamo adoperare la matrice diagonale $M_{\mathcal{A}}^{\mathcal{A}}(f)$. Percio', se $\lambda_1, \dots, \lambda_n$ sono le componenti della diagonale di $M_{\mathcal{A}}^{\mathcal{A}}(f)$, ricordando che il determinante di una matrice diagonale coincide con il prodotto delle componenti sulla diagonale principale, allora

$$p_f(t) = (-1)^n (t - \lambda_1) \cdot (t - \lambda_2) \cdot \cdots \cdot (t - \lambda_n).$$

In particolare osserviamo che *se f e' diagonalizzabile allora il polinomio caratteristico di f si decompone completamente nel prodotto di polinomi di primo grado.* Inoltre *se f e' diagonalizzabile, lo spettro di f e' costituito esattamente dalle componenti sulla diagonale principale di una qualunque matrice diagonale che rappresenti f.*

Ora andiamo a vedere che gli autovalori hanno una molteplicita'.

• *La definizione di molteplicita' algebrica e di molteplicita' geometrica di un autovalore.*

Sia $f : V \to V$ un endomorfismo, e sia λ un autovalore per f. Abbiamo visto che λ e' una radice del polinomio caratteristico $p_f(t)$ di f. Per il Teorema di Ruffini allora il polinomio $p_f(t)$ e' divisibile per $t-\lambda$, cioe' esiste un polinomio $q(t)$ tale che

$$p_f(t) = (t-\lambda)\cdot q(t).$$

Potrebbe accadere che anche il polinomio quoziente $q(t)$ ammetta λ come radice. In tal caso allora $(t-\lambda)^2$ divide il polinomio caratteristico. In generale ci sara' la massima potenza con cui $t-\lambda$ divide $p_f(t)$. Tale numero si dice *la molteplicita' algebrica dell'autovalore λ per f.* Denoteremo la molteplicita' algebrica col simbolo $m_a(\lambda)$.

Esempio. 1) Se $p_f(t) = t^3(t-8)^5$ allora f possiede esattamente due autovalori, cioe' 0 e 8, con molteplicita' algebrica $m_a(0) = 3$ e $m_a(8) = 5$.

2) Supponiamo che $p_f(t) = -t-2t^2-t^3$. Questo e' un caso molto diverso dal precedente perche' adesso il polinomio non appare fattorizzato. Tuttavia non e' difficile trovarne le radici, perche' possiamo mettere t in evidenza, riconducendo il calcolo degli autovalori ad un'equazione di secondo grado. Cioe' possiamo scrivere:

$$p_f(t) = -t(1+2t+t^2) = -t(t+1)^2.$$

Quindi in questo caso lo spettro di f e' formato dagli autovalori 0 e -1, e $m_a(0) = 1$, mentre $m_a(-1) = 2$.

3) Supponiamo che f sia rappresentabile con una matrice del tipo

$$D = \begin{bmatrix} -5 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Quindi f e' diagonalizzabile, e

$$p_f(t) = (t+5)(t+2)^2(t-1).$$

In particolare $\mathrm{spec}(f) = \{-5,-2,1\}$ e $m_a(-5) = 1$, $m_a(-2) = 2$, $m_a(1) = 1$. Si osservi che sulla diagonale di D troviamo tutti gli autovalori di f, e, per ciascun autovalore, la molteplicita' algebrica coincide con il numero di volte con cui l'autovalore appare sulla diagonale di D. E' chiaro che questo e' un fatto di carattere generale. Cioe': *se f e' un endomorfismo diagonalizzabile allora il polinomio caratteristico di f si decompone completamente nel prodotto di polinomi di primo grado; inoltre lo spettro di f e' costituito esattamente dalle componenti sulla diagonale principale di una qualunque matrice diagonale D che rappresenti f, e, per ciascun autovalore, la molteplicita' algebrica coincide con il numero di volte con cui l'autovalore appare sulla diagonale di D.*

Accanto alla molteplicita' algebrica di un autovalore, si definisce anche la molteplicita' geometrica. La molteplicita' geometrica $m_g(\lambda)$ di un autovalore λ per un operatore f e' il massimo numero di autovettori di f linearmente indipendenti relativi a λ. Piu' precisamente, possiamo innanzitutto definire *l'autospazio V_λ di f relativo all'autovalore λ* come

$$V_\lambda = \ker(f - \lambda \cdot \mathrm{id}_V).$$

Si osservi che

$$V_\lambda = \{\mathbf{u} \in V : f(\mathbf{u}) = \lambda \cdot \mathbf{u}\}.$$

Cioe' l'autospazio V_λ e' il sottospazio di V costituito da tutti gli autovettori di f relativi a λ, con l'aggiunta del vettore nullo. Allora si definisce:

$$m_g(\lambda) := \dim V_\lambda.$$

Osserviamo anche che

$$m_g(\lambda) = n - rk(A - \lambda \cdot I),$$

dove $n = \dim V$, ed A e' una qualunque matrice rappresentativa di f. Inoltre, poiche' λ e' un autovalore, allora

$$m_g(\lambda) \geq 1.$$

- *Confronto tra la molteplicita' algebrica e la molteplicita' geometrica di un autovalore.*

Questi due numeri che abbiamo appena definito sono confrontabili, nel seguente senso.

Sia $f : V \to V$ un endomorfismo, e sia $\lambda \in \mathbf{R}$ un autovalore per f. Allora:

1) *la molteplicita' geometrica e' minore o uguale alla molteplicita' algebrica, cioe'*

$$1 \leq m_g(\lambda) \leq m_a(\lambda);$$

2) *se inoltre f e' diagonalizzabile, vale l'uguaglianza*

$$m_g(\lambda) = m_a(\lambda).$$

Dimostrazione. 1) Sia $\mathcal{A}_\lambda = \{\mathbf{b}_1, \ldots, \mathbf{b}_h\}$ una base per l'autospazio V_λ, quindi $h = m_g(\lambda)$. Aggiungendo opportuni $n - h$ vettori di V possiamo formare una base $\mathcal{B}$ di V:

$$\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_h, \mathbf{c}_{h+1} \ldots, \mathbf{c}_n\}.$$

Poiche' $f(\mathbf{b}_1) = \lambda \cdot \mathbf{b}_1$, allora la prima colonna della matrice rappresentativa $M_{\mathcal{B}}^{\mathcal{B}}(f)$ di f e' il vettore numerico di lunghezza n dato da $(\lambda, 0, \ldots, 0)^T$. La seconda colonna e' il vettore $(0, \lambda, 0, \ldots, 0)^T$, ..., la colonna di posto h e' il vettore $(0, \ldots, 0, \lambda, 0, \ldots, 0)^T$, dove λ occupa il posto di indice $(i, j) = (h, h)$. Quindi la prima colonna della matrice $M_{\mathcal{B}}^{\mathcal{B}}(f) - t \cdot I$ e' il vettore numerico $(\lambda - t, 0, \ldots, 0)^T$, la seconda colonna e' il vettore $(0, \lambda - t, 0, \ldots, 0)^T$, ..., la colonna di posto h e' il vettore $(0, \ldots, 0, \lambda - t, 0, \ldots, 0)^T$. Percio', nel calcolare il polinomio caratteristico $p_f(t) = \det(M_{\mathcal{B}}^{\mathcal{B}}(f) - t \cdot I)$ di f con lo sviluppo di Laplace rispetto alle prime h colonne, riusciremo a mettere in evidenza per h volte il fattore $\lambda - t$. Cioe' potremo scrivere

$$p_f(t) = (t - \lambda)^h \cdot q(t)$$

dove $q(t)$ e' un opportuno polinomio. Nella matrice seguente, per maggiore chiarezza, abbiamo fatto un esempio con $n = 5$ ed $h = 3$. Le componenti a_{ij} denotano numeri qualsiasi.

$$M_{\mathcal{B}}^{\mathcal{B}}(f) - t \cdot I = \begin{bmatrix} \lambda - t & 0 & 0 & a_{14} & a_{15} \\ 0 & \lambda - t & 0 & a_{24} & a_{25} \\ 0 & 0 & \lambda - t & a_{34} & a_{35} \\ 0 & 0 & 0 & a_{44} - t & a_{45} \\ 0 & 0 & 0 & a_{54} & a_{55} - t \end{bmatrix},$$

$$\det(M_{\mathcal{B}}^{\mathcal{B}}(f) - t \cdot I) = (t - \lambda)^3[-t^2 - (a_{44} + a_{55})t - (a_{44}a_{55} - a_{45}a_{54})].$$

Cio' prova che la massima potenza con cui $t-\lambda$ divide $p_f(t)$, cioe' $m_a(\lambda)$, e' almeno $h=m_g(\lambda)$.

2) Supponiamo ora che f sia diagonalizzabile, e sia $\mathcal{A}$ una base di autovettori per f, in modo tale che la matrice $M_{\mathcal{A}}^{\mathcal{A}}(f)$ sia diagonale. Sia k la molteplicita' algebrica di λ. Abbiamo visto che λ appare esattamente k volte lungo la diagonale di $M_{\mathcal{A}}^{\mathcal{A}}(f)$. In corrispondenza delle k componenti di $M_{\mathcal{A}}^{\mathcal{A}}(f)$ uguali a λ, ci saranno k autovettori relativi a λ nella base $\mathcal{A}$. Quindi nell'autospazio V_λ ci sono almeno k vettori linearmente indipendenti. Ne consegue che $k\leq \dim V_\lambda$, cioe' che $m_a(\lambda)=k\leq \dim V_\lambda = m_g(\lambda)$. Cio' conclude la dimostrazione, perche' gia' sappiamo che in generale $m_g(\lambda)\leq m_a(\lambda)$. $\square$

Riassumendo brevemente quanto detto sinora, possiamo dire che condizione necessaria affinche' un operatore $f:V\to V$ sia diagonalizzabile e' che siano verificate le seguenti due condizioni.

1) Il polinomio caratteristico di f si decompone completamente nel prodotto di polinomi a coefficienti reali di primo grado.

2) Per ciascun autovalore λ di f la molteplicita' algebrica coincide con quella geometrica.

Ora andremo a provare che tali condizioni sono anche sufficienti, pervenendo cosi' all'annunciato algoritmo per la diagonalizzazione di un endomorfismo.

• *Algoritmo per la diagonalizzazione di un endomorfismo.*

Sia $f:V\to V$ un endomorfismo di uno spazio vettoriale reale V di dimensione finita $\dim V=n$. *L'operatore f e' diagonalizzabile se e solo se valgono le seguenti due proprieta'.*

1) *Il polinomio caratteristico di f si decompone completamente nel prodotto di polinomi di primo grado a coefficienti reali.*

2) *Per ogni autovalore λ di f si ha $m_g(\lambda)=m_a(\lambda)$.*

In tal caso, denotata con $\mathcal{A}_\lambda$ una base per l'autospazio V_λ, l'unione

$$\mathcal{A}:=\bigcup_{\lambda\in Spec(f)}\mathcal{A}_\lambda$$

e' una base di V tale che $M_{\mathcal{A}}^{\mathcal{A}}(f)$ e' una matrice diagonale (in altre parole $\mathcal{A}$ e' una base di V formata da autovettori per f). Sulla diagonale principale di tale matrice appaiono gli autovalori di f, ciascuno tante volte quant'e' la sua molteplicita' algebrica. Tali autovalori appaiono inoltre in ordine corrispondente all'ordine con cui figurano gli autovettori di $\mathcal{A}$. In particolare, se $f=L_A$ e' un operatore del tipo $\mathbf{x}\in\mathbf{R}^n\to A\mathbf{x}\in\mathbf{R}^n$, detta P la matrice le cui colonne sono i vettori di $\mathcal{A}$, allora $P^{-1}AP$ e' una matrice diagonale simile ad A.

Dimostrazione. Occorre solo provare che le due proprieta' sono una condizione sufficiente.

A tale proposito, denotiamo con $\lambda_1,\dots,\lambda_h$ gli autovalori (distinti) di f. Poiche' per ipotesi il polinomio caratteristico si fattorizza nel prodotto di polinomi di primo grado, raccogliendo gli eventuali fattori uguali, potremo mettere $p_f(t)$ sotto la forma (a meno del segno):

$$p_f(t)=(t-\lambda_1)^{m_a(\lambda_1)}\cdot\dots\cdot(t-\lambda_h)^{m_a(\lambda_h)}.$$

E poiche' $p_f(t)$ ha grado n deve essere:

$$m_a(\lambda_1)+\dots+m_a(\lambda_h)=n.$$

Ora per ciascun λ_i denotiamo con $\mathcal{A}_{\lambda_i}$ una base per l'autospazio V_{λ_i}. In ciascun $\mathcal{A}_{\lambda_i}$ ci sono esattamente $m_g(\lambda_i)$ vettori. Utilizzeremo il seguente lemma che dimostreremo dopo, una volta conclusa la dimostrazione dell'algoritmo.

Lemma. Siano $\lambda_1, \dots, \lambda_r$ *autovalori distinti per* f, *e per ogni* $i = 1, \dots, r$, *siano* $\mathbf{b}_{1i}, \dots, \mathbf{b}_{l_i i}$ l_i *autovettori linearmente indipendenti relativi all'autovalore* λ_i. *Allora il sistema che si ottiene riunendo tutti i vettori* $\mathbf{b}_{ji}$, $i = 1, \dots, r$, $j = 1, \dots, l_i$ *e' linearmente indipendente.*

Grazie a questa proprieta', riunendo le basi $\mathcal{A}_{\lambda_i}$ si ottiene un sistema $\mathcal{A}$ linearmente indipendente di autovettori per f. La dimostrazione sara' conclusa se faremo vedere che in $\mathcal{A}$ ci sono n vettori. In tal caso infatti potremo dire che $\mathcal{A}$ e' una base di V costituita da autovettori, e sappiamo che cio' equivale a dire che f e' diagonalizzabile.

Se contiamo i vettori che stanno in $\mathcal{A}$, troviamo:

$$\#\mathcal{A} = m_g(\lambda_1) + \dots + m_g(\lambda_h).$$

Ma per ipotesi la molteplicita' geometrica coincide con quella algebrica, e percio':

$$m_g(\lambda_1) + \dots + m_g(\lambda_h) = m_a(\lambda_1) + \dots + m_a(\lambda_h) = n. \quad \square$$

Avendo concluso la dimostrazione dell'algoritmo, andiamo a dimostrare il lemma.

Dimostrazione del Lemma. Ragioniamo per induzione sul numero di autovalori r. Se $r = 1$ non dobbiamo dimostrare niente, perche' gia' sappiamo che i vettori $\mathbf{b}_{11}, \dots, \mathbf{b}_{l_1 1}$ sono liberi. Supponiamo allora $r > 1$, e sia

$$a_{11}\mathbf{b}_{11} + \dots + a_{l_1 1}\mathbf{b}_{l_1 1} + a_{12}\mathbf{b}_{12} + \dots + a_{l_2 2}\mathbf{b}_{l_2 2} + \dots + a_{1r}\mathbf{b}_{1r} + \dots + a_{l_r r}\mathbf{b}_{l_r r} = \mathbf{0}$$

una relazione tra i vettori $\mathbf{b}_{ji}$. Andiamo a provare che tutti i pesi a_{ji} devono essere nulli. Moltiplicando tutto per λ_1 otteniamo ancora una relazione:

$$a_{11}\lambda_1\mathbf{b}_{11} + \dots + a_{l_1 1}\lambda_1\mathbf{b}_{l_1 1} + a_{12}\lambda_1\mathbf{b}_{12} + \dots + a_{l_2 2}\lambda_1\mathbf{b}_{l_2 2} + \dots + a_{1r}\lambda_1\mathbf{b}_{1r} + \dots + a_{l_r r}\lambda_1\mathbf{b}_{l_r r} = \mathbf{0}.$$

Applicando alla prima relazione l'operatore f, tenuto conto che i vettori in esame sono autovettori, otteniamo:

$$a_{11}\lambda_1\mathbf{b}_{11} + \dots + a_{l_1 1}\lambda_1\mathbf{b}_{l_1 1} + a_{12}\lambda_2\mathbf{b}_{12} + \dots + a_{l_2 2}\lambda_2\mathbf{b}_{l_2 2} + \dots + a_{1r}\lambda_r\mathbf{b}_{1r} + \dots + a_{l_r r}\lambda_r\mathbf{b}_{l_r r} = \mathbf{0}.$$

Sottraendo la seconda relazione, gli autovettori relativi all'autovalore λ_1 si elidono, e mettendo in evidenza, abbiamo:

$$a_{12}(\lambda_2 - \lambda_1)\mathbf{b}_{12} + \dots + a_{l_2 2}(\lambda_2 - \lambda_1)\mathbf{b}_{l_2 2} + \dots + a_{1r}(\lambda_r - \lambda_1)\mathbf{b}_{1r} + \dots + a_{l_r r}(\lambda_r - \lambda_1)\mathbf{b}_{l_r r} = \mathbf{0}.$$

Questa adesso e' una relazione che coinvolge solo autovettori relativi agli $r - 1$ autovalori $\lambda_2, \dots, \lambda_r$. Per ipotesi induttiva tale relazione deve essere banale. Poiche' gli autovalori sono tutti distinti tra loro, $\lambda_i - \lambda_1 \neq 0$ per ogni $i = 2, \dots, r$. Cio' implica che tutti i pesi a_{ji} sono nulli per ogni $i = 2, \dots, r$ ed ogni $j = 1, \dots, l_i$. Cio' implica anche che la prima relazione si riduce ad una relazione tra gli autovettori relativi a λ_1, che sono liberi per ipotesi. In definitiva tutti i pesi a_{ji} sono nulli, ed il lemma e' dimostrato. $\square$

Osservazione. (i) Il termine "algoritmo" e' un po' impreciso, perche' il calcolo delle radici di un polinomio, in generale, non puo' essere eseguito da un computer. Tutti gli altri passi invece sono implementabili.

(ii) Sia $f : V \to V$ un endomorfismo, e siano $\lambda_1, \dots, \lambda_h$ gli autovalori distinti di f. Per ogni λ_i sia $\mathcal{A}_{\lambda_i}$ una base per l'autospazio V_{λ_i}. Abbiamo appena dimostrato nel lemma precedente che la riunione delle basi $\mathcal{A}_{\lambda_i}$ e' ancora un sistema linearmente indipendente. Percio' la somma dei sottospazi V_{λ_i} e' diretta, cioe'

$$V_{\lambda_1} + \cdots + V_{\lambda_h} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_h}.$$

Ed inoltre possiamo dire che $f : V \to V$ e' diagonalizzabile se e solo se tutto lo spazio V e' la somma diretta dei suoi autospazi:

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_h}.$$

In tal caso ogni vettore $\mathbf{u} \in V$ si puo' esprimere, in unico modo, come somma di autovettori. Ne consegue, e gia' lo sappiamo, che un operatore non e' diagonalizzabile se e solo se non ha un numero sufficiente di autovettori che consenta di generare V.

(iii) Se λ e' un autovalore per f con molteplicita' algebrica 1 allora necessariamente avremo $m_g(\lambda) = m_a(\lambda) = 1$. Percio' se f e' definito su uno spazio V di dimensione n e possiede n autovalori distinti, allora f e' diagonalizzabile.

(iv) Osserviamo che 0 e' un autovalore per f se e solo se $\ker f \neq \{\mathbf{0}\}$. In tal caso $\ker f$ coincide con l'autospazio V_0 relativo all'autovalore $\lambda = 0$.

(v) Ricordiamo che se λ e' un autovalore per $f : V \to V$, detta $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ una matrice rappresentativa di f, allora

$$m_g(\lambda) = \dim V - rk(A - \lambda \cdot I).$$

Esercizio. Sia $f : \mathbf{R}^3 \to \mathbf{R}^3$ l'operatore lineare definito ponendo

$$f(x, y, z) = (3x + 3y - z, -4x - 2y + 2z, -2x + 2y + 2z).$$

Provare che f e' diagonalizzabile, e trovare una base di $\mathbf{R}^3$ formata da autovettori.

Svolgimento. La matrice rappresentativa di f rispetto alla base canonica e':

$$A = M_{\mathcal{E}}^{\mathcal{E}}(f) = \begin{bmatrix} 3 & 3 & -1 \\ -4 & -2 & 2 \\ -2 & 2 & 2 \end{bmatrix}.$$

Quindi il polinomio caratteristico di f e':

$$p_f(t) = \det \begin{bmatrix} 3-t & 3 & -1 \\ -4 & -2-t & 2 \\ -2 & 2 & 2-t \end{bmatrix} = -t(t-1)(t-2).$$

Percio' f possiede tre autovalori distinti. Cio' prova che f e' diagonalizzabile. Per trovare una base di autovettori, andiamo a studiare ad uno ad uno i singoli autospazi.

L'autospazio V_0 altro non e' che il nucleo di f, la cui rappresentazione e' data dal sistema:

$$A \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

cioe' dal sistema

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

Risolvendo tale sistema si vede che una base per V_0 e' data dal vettore $(2, -1, 3)$.

L'autospazio V_1 e' rappresentato dal sistema:

$$(A - I) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

cioe' dal sistema

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

Risolvendo tale sistema si vede che una base per V_1 e' data dal vettore $(1, 0, 2)$.

Infine l'autospazio V_2 e' rappresentato dal sistema:

$$(A - 2I) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

cioe' dal sistema

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

Risolvendo tale sistema si vede che una base per V_1 e' data dal vettore $(1, 1, 4)$.

Mettendo insieme i tre vettori trovati, otteniamo una base di autovettori per f:

$$\mathcal{A} = \{(2, -1, 3), (1, 0, 2), (1, 1, 4)\}.$$

In particolare

$$M^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \quad \square$$

Esercizio. Provare che la seguente matrice A e' diagonalizzabile, e trovare una matrice invertibile P tale che $P^{-1} \cdot A \cdot P$ sia una matrice diagonale.

$$A = \begin{bmatrix} 3 & 3 & -1 \\ -4 & -2 & 2 \\ -2 & 2 & 2 \end{bmatrix}.$$

Svolgimento. E' una variante dell'esercizio precedente. Infatti nell'esercizio precedente abbiamo visto che l'operatore f, che coincide con L_A, e' diagonalizzabile. Cio' equivale a dire che A e' diagonalizzabile. Inoltre la matrice P cercata e' la matrice che ha per colonne i vettori della base di autovettori $\mathcal{A}$ trovata:

$$P = M^{\mathcal{A}}_{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^3}) = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}.$$

Infatti

$$M^{\mathcal{A}}_{\mathcal{A}}(f) = M^{\mathcal{E}}_{\mathcal{A}}(\mathrm{id}_{\mathbf{R}^3}) \cdot M^{\mathcal{E}}_{\mathcal{E}}(f) \cdot M^{\mathcal{A}}_{\mathcal{E}}(\mathrm{id}_{\mathbf{R}^3}),$$

e cioe'

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = P^{-1} \cdot \begin{bmatrix} 3 & 3 & -1 \\ -4 & -2 & 2 \\ -2 & 2 & 2 \end{bmatrix} \cdot P. \quad \square$$

Esercizio. Provare che la seguente matrice A e' diagonalizzabile, e trovare una matrice invertibile P tale che $P^{-1} \cdot A \cdot P$ sia una matrice diagonale.

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & -1 \end{bmatrix}.$$

Svolgimento. Il polinomio caratteristico di A e':

$$p_A(t) = -(t+1)^2(t-1).$$

Quindi lo spettro di A e' costituito dagli autovalori -1 ed 1. Poiche' $m_a(1) = 1$, allora $m_a(1) = m_g(1) = 1$. Invece $m_a(-1) = 2$ e

$$m_g(-1) = 3 - rk(A+I) = 3 - rk \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} = 2.$$

Percio' A e' diagonalizzabile.

L'autospazio V_1 e' rappresentato dal sistema:

$$(A - I) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

cioe' dal sistema

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

Risolvendo tale sistema si vede che una base per V_1 e' data dal vettore $(1, 2, 0)$.

L'autospazio V_{-1} e' rappresentato dal sistema:

$$(A + I) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

cioe' dal sistema

$$\begin{cases} y + 2z = 0 \\ 2y + 4z = 0. \end{cases}$$

Risolvendo tale sistema si vede che una base per V_{-1} e' data dai vettori $(1, 0, 0)$ e $(0, -2, 1)$.

Mettendo in colonna gli autovettori cosi' calcolati otteniamo la matrice cercata:

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

Infatti

$$P^{-1} \cdot A \cdot P = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \square$$

Esercizio. Provare che la seguente matrice A non e' diagonalizzabile.

$$A = \begin{bmatrix} 8 & 7 & -3 \\ -4 & -2 & 2 \\ 8 & 10 & -2 \end{bmatrix}.$$

Svolgimento. Il polinomio caratteristico e':

$$p_A(t) = -t(t-2)^2.$$

Ora l'autovalore 2 ha molteplicita' algebrica 2, ma la sua molteplicita' geometrica e' 1. Infatti:

$$m_g(2) = 3 - rk(A - 2I) = 3 - rk \begin{bmatrix} 6 & 7 & -3 \\ -4 & -4 & 2 \\ 8 & 10 & -4 \end{bmatrix} = 1. \quad \square$$

Esercizio. Dire se e' vero oppure no che la proiezione ortogonale e' un operatore diagonalizzabile.

Svolgimento. Si', lo e'. Infatti denotiamo con $f : \mathcal{V}_{O,\rho} \to V_{O,\rho}$ l'operatore di proiezione ortogonale su una retta r del piano ρ, passante per il punto di applicazione O. I vettori applicati in O ortogonali ad r avranno proiezione nulla. Quindi 0 e' un autovalore per f. Ed anche 1 e' un autovalore, perche' i vettori giacenti su r coincidono con la propria proiezione. Quindi f ha due autovalori distinti, e percio', avendo $V_{O,\rho}$ dimensione 2, la proiezione ortogonale e' diagonalizzabile. $\square$

Esercizio. Dire se e' vero oppure no che la rotazione di angolo $\alpha = 90^0$ e' un operatore diagonalizzabile.

Svolgimento. Abbiamo gia' osservato che una rotazione $f : \mathcal{V}_{O,\rho} \to V_{O,\rho}$ non e' diagonalizzabile (a meno che $\alpha \in \{0, \pi\}$). Una dimostrazione algebrica consiste nel considerare la matrice rappresentativa della rotazione rispetto ad una fissata base monometrica ortogonale $\mathcal{B}$:

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

ed osservare che il polinomio caratteristico

$$p_f(t) = 1 + t^2$$

non ha radici reali. Percio', non avendo la rotazione autovalori, certamente non e' diagonalizzabile. $\square$

La seguente osservazione, che conclude il presente capitolo, ci introduce al prossimo, sulla forma canonica di Jordan.

Osservazione. Come abbiamo visto nell'esercizio precedente, la matrice

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

non e' diagonalizzabile. Piu' precisamente, non esiste una matrice quadrata P di ordine 2, invertibile, con componenti *numeri reali* tale che $P^{-1} \cdot A \cdot P$ sia una matrice diagonale. Cio' e' dovuto al fatto che il polinomio caratteristico di A

$$p_A(t) = 1 + t^2$$

non possiede radici reali. Ma se noi ampliamo il campo degli scalari, introducendo i numeri complessi, allora tale polinomio e' decomponibile nel prodotto di polinomi di primo grado:

$$p_A(t) = (t - i)(t + i)$$

dove con $i = \sqrt{-1}$ denotiamo la radice immaginaria di -1. Poiche' tutta la teoria che abbiamo studiato sugli spazi vettoriali vale anche per gli spazi vettoriali a coefficienti complessi (per esempio $V = \mathbf{C}^n$), ne consegue che l'algoritmo per la diagonalizzazione si applica ad A, se consideriamo A come una matrice a coefficienti complessi. Cioe', poiche' la matrice A e' di ordine 2 ed ha due autovalori distinti:

$$\mathrm{spec}(A) = \{i, -i\},$$

allora A e' diagonalizzabile su $\mathbf{C}$, nel senso che esiste una matrice quadrata P di ordine 2, invertibile, con componenti *numeri complessi* tale che $P^{-1} \cdot A \cdot P$ sia una matrice diagonale. Infatti l'autospazio V_i e' rappresentato dal sistema

$$\begin{cases} -ix - y = 0 \\ x - iy = 0. \end{cases}$$

Da cui si deduce che

$$V_i = \mathrm{Span}((i, 1)) \subseteq \mathbf{C}^2.$$

Similmente si vede che

$$V_{-i} = \mathrm{Span}((-i, 1)) \subseteq \mathbf{C}^2.$$

Quindi disponendo in colonna gli autovettori appena calcolati otteniamo una matrice invertibile

$$P := \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$$

tale che $P^{-1} \cdot A \cdot P$ e' diagonale. Infatti:

$$P^{-1} = \begin{bmatrix} \frac{1}{2i} & \frac{1}{2} \\ \frac{-1}{2i} & \frac{1}{2} \end{bmatrix},$$

e percio'

$$P^{-1} \cdot A \cdot P = \begin{bmatrix} \frac{1}{2i} & \frac{1}{2} \\ \frac{-1}{2i} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Questo esempio e' un caso particolare di un fatto generale, cioe' che ogni polinomio a coefficienti complessi e' decomponibile nel prodotto di polinomi complessi di primo grado. Per cui la prima condizione dell'algoritmo diventa sovrabbondante per un endomorfismo di uno spazio a coefficienti complessi. Ne consegue una semplificazione nell'enunciato dell'algoritmo, e cioe' *un endomorfismo $f : V \to V$ di uno spazio a coefficienti complessi e' diagonalizzabile se e solo se per ogni autovalore di f la molteplicita' geometrica coincide con quella algebrica.* Come vedremo, quest'ultima "ostruzione" non si puo' risolvere ampliando il campo dei numeri. Anche sui numeri complessi non e' detto che una matrice sia diagonalizzabile, ma comunque sara' possibile ricondurla, via similitudine, ad una forma semplice, la cosiddetta "forma canonica di Jordan".

Indice dei paragrafi.