III-2 Trasformazioni lineari e matrici

Indice

1 Trasformazioni lineari 1

2 Matrici 6

3 Immagine di una trasformazione lineare 12

4 Inversione di una trasformazione lineare 13

5 Soluzioni degli esercizi 15

In questa sezione vediamo alcune nozioni fondamentali sulle trasformazioni lineari tra due spazi vettoriali $\mathbb{R}^n$ e $\mathbb{R}^m$ e sulle matrici.

I vettori generalmente sono da intendere come *vettori colonna* (il motivo apparirà chiaro tra un po'). Talvolta però, quando la cosa non sarà del tutto indispensabile, per "esigenze tipografiche" saranno scritti in riga.

1 Trasformazioni lineari

Da un punto di vista generale teorico e nello stesso tempo da un punto di vista delle applicazioni è fondamentale il concetto di trasformazione (di uno spazio in un altro). Qui ci occuperemo di un particolare tipo di trasformazioni (le trasformazioni lineari) tra due spazi vettoriali (nel caso da noi considerato spazi $\mathbb{R}^n$). Successivamente parleremo di matrici, che non sono altro che strumenti per rappresentare le trasformazioni lineari.

Consideriamo una funzione f, definita nello spazio vettoriale $\mathbb{R}^n$, a valori nello spazio vettoriale $\mathbb{R}^m$. [1] Scriviamo naturalmente

$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

Qualora se ne presenti l'opportunità, per distinguere una tale funzione dalle usuali funzioni reali di variabile reale, cioè dalle funzioni definite in $\mathbb{R}$ a valori in $\mathbb{R}$, che lo studente ha visto in precedenza, si può dire che, se $n > 1$, la funzione è di variabile vettoriale e, se $m > 1$, la funzione è a valori vettoriali.

Possiamo ovviamente pensare anche a funzioni reali di variabile vettoriale ($f : \mathbb{R}^n \to \mathbb{R}$, con $n > 1$) e a funzioni vettoriali di variabile reale ($f : \mathbb{R} \to \mathbb{R}^m$, con $m > 1$).

Particolare importanza per i nostri scopi hanno le seguenti funzioni:

Definizione Se $f : \mathbb{R}^n \to \mathbb{R}^m$, diciamo che f è una **trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$** se f ha queste due proprietà:

1. $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ (proprietà di *additività*);
2. $\forall \boldsymbol{x} \in \mathbb{R}^n$ e $\forall \alpha \in \mathbb{R}$, $f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x})$ (proprietà di *omogeneità*).

Osservazione Si noti che nella definizione di trasformazione lineare si utilizza la struttura (le operazioni) di spazio vettoriale, sia nel dominio sia nel codominio di f.

Osservazione Immediata conseguenza della definizione data è che, se f è una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$ e se $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^k$ sono vettori di $\mathbb{R}^n$, allora per ogni c.l. $\alpha_1\boldsymbol{v}^1 + \alpha_2\boldsymbol{v}^2 + \ldots + \alpha_k\boldsymbol{v}^k$ dei k vettori si ha

$$f(\alpha_1\boldsymbol{v}^1 + \alpha_2\boldsymbol{v}^2 + \ldots + \alpha_k\boldsymbol{v}^k) = \alpha_1 f(\boldsymbol{v}^1) + \alpha_2 f(\boldsymbol{v}^2) + \ldots + \alpha_k f(\boldsymbol{v}^k).$$

Si può dire quindi che il trasformato di una c.l. è la c.l. dei trasformati.

[1] Si faccia attenzione quindi che in generale gli spazi tra cui la funzione è definita possono essere diversi: sono entrambi spazi tipo $\mathbb{R}^n$, ma con dimensioni che possono essere diverse. Naturalmente, come caso particolare, potremo avere funzioni definite nello stesso spazio.

Esempi

- Al solo scopo di mostrare come si deve procedere per provare che una trasformazione è lineare, in base alla definizione, fornisco il seguente esempio che, pur essendo contrassegnato dalla “doppia curva”, più che difficile è lungo. Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ definita da
$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_1 + x_2 \end{pmatrix}.$$
Verifichiamo che è una trasformazione lineare. Si tratta di verificare che f ha le due proprietà previste nella definizione.
Per quanto riguarda la proprietà 1, presi due generici elementi $\boldsymbol{x} = (x_1, x_2)$ e $\boldsymbol{y} = (y_1, y_2)$ in $\mathbb{R}^2$, abbiamo
$$\begin{aligned}
f(\boldsymbol{x} + \boldsymbol{y}) &= f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \\
\text{(per definizione di vettore somma)} \quad &= f \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \\
\text{(per la definizione di } f\text{)} \quad &= \begin{pmatrix} 2(x_1 + y_1) - (x_2 + y_2) \\ (x_1 + y_1) + (x_2 + y_2) \end{pmatrix} \\
\text{(per le proprietà di } \mathbb{R}\text{)} \quad &= \begin{pmatrix} 2x_1 - x_2 + 2y_1 - y_2 \\ x_1 + x_2 + y_1 + y_2 \end{pmatrix} \\
\text{(per definizione di vettore somma)} \quad &= \begin{pmatrix} 2x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} + \begin{pmatrix} 2y_1 - y_2 \\ y_1 + y_2 \end{pmatrix} \\
\text{(per la definizione di } f\text{)} \quad &= f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + f \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= f(\boldsymbol{x}) + f(\boldsymbol{y}).
\end{aligned}$$
L'additività è dunque dimostrata.
Vediamo l'omogeneità (proprietà 2). Preso un $\boldsymbol{x} \in \mathbb{R}^2$ e un numero reale α, abbiamo
$$\begin{aligned}
f(\alpha \boldsymbol{x}) &= f\left(\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \\
\text{(per definizione di moltiplicazione scalare)} \quad &= f \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix} \\
\text{(per la definizione di } f\text{)} \quad &= \begin{pmatrix} 2\alpha x_1 - \alpha x_2 \\ \alpha x_1 + \alpha x_2 \end{pmatrix} \\
\text{(per definizione di moltiplicazione scalare)} \quad &= \alpha \begin{pmatrix} 2x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} \\
\text{(per la definizione di } f\text{)} \quad &= \alpha f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \alpha f(\boldsymbol{x}).
\end{aligned}$$
Quindi f è una trasformazione lineare, essendo additiva ed omogenea.

- Consideriamo ora la funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ definita da
$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2^2 \end{pmatrix}.$$
Questa non è una trasformazione lineare. Infatti non è additiva.[2] Siano, ad esempio,
$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{e} \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

[2] Un commento è d'obbligo. Per provare che f non è additiva è sufficiente quello che si chiama un controesempio (ne basta uno solo): un controesempio è un caso particolare in cui la proprietà in questione non vale. Se possiamo trovare un controesempio significa che la proprietà in generale non vale. Si noti che invece, nell'esempio precedente, per provare che la trasformazione era lineare, abbiamo dovuto provarlo in generale, cioè con riferimento a vettori del tutto generici. Nel presente esempio la prova della non linearità può finire con la verifica della non additività (la definizione di linearità richiede l'additività *e* l'omogeneità, quindi se una trasformazione non è additiva non è nemmeno lineare).

Risulta allora

$$f(\boldsymbol{x}+\boldsymbol{y}) = f\begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}2\\9\end{pmatrix} \qquad \text{mentre} \qquad f(x)+f(y) = f\begin{pmatrix}1\\1\end{pmatrix} + f\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1\\4\end{pmatrix} = \begin{pmatrix}2\\5\end{pmatrix}.$$

Si vede facilmente che f non è nemmeno omogenea. Si potrebbe dire che f è lineare nella sua prima componente, ma non nella seconda.

Osservazione (importante, vedi anche la nota) Lo studente noti come, nel primo esempio, per dimostrare che f è lineare sia necessario farlo in generale, cioè con riferimento a generici vettori $\boldsymbol{x}$ e $\boldsymbol{y}$ per l'addizione e un generico vettore $\boldsymbol{x}$ e un generico numero reale α per la moltiplicazione scalare. In questo secondo esempio invece, per provare che f non è lineare, basta far vedere che anche una sola delle due proprietà non vale anche solo in un caso particolare.

- La funzione
$$f\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1 - x_2$$
è una trasformazione lineare di $\mathbb{R}^2$ in $\mathbb{R}$ (lo si verifichi).
- La funzione
$$f\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1 x_2$$
non è una trasformazione lineare di $\mathbb{R}^2$ in $\mathbb{R}$ (lo si verifichi).
- La funzione
$$f\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}x_1+1\\x_2\end{pmatrix}$$
non è una trasformazione lineare di $\mathbb{R}^2$ in $\mathbb{R}^2$ (lo si verifichi).

Osservazione Se f è lineare, allora $f(\mathbf{0}) = \mathbf{0}$. Infatti

$$f(0\cdot\boldsymbol{x}) = 0\cdot f(\boldsymbol{x}) \quad \text{(omogeneità) e quindi} \quad f(\mathbf{0}) = \mathbf{0}.$$

Questo risultato fornisce un semplice metodo per dimostrare, in certi casi, che una trasformazione non è lineare. Se infatti una trasformazione f non trasforma $\mathbf{0}$ in $\mathbf{0}$, allora non può essere lineare. Ad esempio, nell'ultimo dei casi proposti qui sopra, si ha $f\binom{0}{0} = \binom{1}{0}$, e quindi f non è lineare. Si osservi anche che invece, nel secondo dei casi proposti, f non è lineare pur verificando la proprietà $f(\mathbf{0}) = \mathbf{0}$. Anche nel primo caso proposto si ha $f(\mathbf{0}) = \mathbf{0}$, ma questo non basta per dire che la trasformazione è lineare, trattandosi di una condizione necessaria ma non sufficiente per la linearità.

Sia f una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$. Si può pensare che sia

$$f(\boldsymbol{x}) = \begin{pmatrix}f_1(\boldsymbol{x})\\ \vdots \\ f_m(\boldsymbol{x})\end{pmatrix},$$

dove ciascuna f_j è una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}$. In pratica $f_j(\boldsymbol{x})$ è la j-esima componente del vettore $f(\boldsymbol{x})$.

Si può dimostrare facilmente che f è lineare se e solo se tutte le componenti f_j sono lineari.

Quindi, ad esempio per dimostrare che

$$f\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}2x_1 - x_2\\x_1 + x_2\end{pmatrix}$$

è lineare si potrebbe dimostrare che le sue singole componenti lo sono, e cioè che sono lineari le funzioni (a valori reali)

$$f_1\begin{pmatrix}x_1\\x_2\end{pmatrix} = 2x_1 - x_2 \qquad \text{e} \qquad f_2\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1 + x_2.$$

Parliamo ora dell'importante questione della rappresentazione di una trasformazione lineare.

Cominciamo con il caso particolare di una trasformazione lineare $f:\mathbb{R}^n \to \mathbb{R}$ (cioè a valori in $\mathbb{R}$). Se, dato $\boldsymbol{x}\in\mathbb{R}^n$, scriviamo $\boldsymbol{x} = x_1\boldsymbol{u}^1 + x_2\boldsymbol{u}^2 + \ldots + x_n\boldsymbol{u}^n$, dove $\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^n$ sono i vettori della base fondamentale di $\mathbb{R}^n$, la linearità di f permette di scrivere

$$f(\boldsymbol{x}) = f(x_1\boldsymbol{u}^1 + x_2\boldsymbol{u}^2 + \ldots + x_n\boldsymbol{u}^n) = x_1 f(\boldsymbol{u}^1) + x_2 f(\boldsymbol{u}^2) + \ldots + x_n f(\boldsymbol{u}^n)$$

(si noti che gli x_j e gli $f(\boldsymbol{u}^j)$ qui sono tutti numeri reali, dato che f ha valori in $\mathbb{R}$).

Se poniamo $\boldsymbol{a} = \big(f(\boldsymbol{u}^1), f(\boldsymbol{u}^2), \dots, f(\boldsymbol{u}^n)\big)$, allora possiamo scrivere $f(\boldsymbol{x}) = \langle \boldsymbol{a}, \boldsymbol{x} \rangle$. [3]

Il vettore $\boldsymbol{a}$ è sufficiente per rappresentare la trasformazione f. A titolo di esempio, se

$$f(\boldsymbol{x}) = f\big(x_1, x_2\big) = 3x_1 - 2x_2 = \langle (3, -2), (x_1, x_2) \rangle = \langle \boldsymbol{a}, \boldsymbol{x} \rangle,$$

è chiaro che per rappresentare f basta il vettore $\boldsymbol{a} = (3, -2)$.

A questo punto non è difficile capire che è possibile rappresentare una qualunque trasformazione lineare f attraverso una sorta di "tabella": basterà rappresentare ogni componente di f con un vettore (riga) come appena visto ed accostare tali vettori riga uno sull'altro.

Esempio Consideriamo ad esempio la trasformazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$ definita da

$$f(\boldsymbol{x}) = f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_2 + x_3 \end{pmatrix}.$$

Se rappresentiamo le due componenti del vettore $f(\boldsymbol{x})$ rispettivamente con i vettori riga

$$\begin{pmatrix} 2 & -1 & 0 \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 0 & 1 & 1 \end{pmatrix},$$

e accostiamo poi verticalmente questi due otteniamo

$$\begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Questa si chiama **matrice**. Possiamo scrivere il vettore $f(\boldsymbol{x})$ come prodotto tra la matrice suddetta e il vettore $\boldsymbol{x}$, cioè scrivere

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_2 + x_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

se pensiamo che tale prodotto matrice×vettore avvenga calcolando i prodotti interni delle **righe** della matrice per il vettore colonna $\boldsymbol{x}$.

Saremmo pervenuti allo stesso risultato ragionando anche così: la linearità di f, come prima, porta a dire che

$$f(\boldsymbol{x}) = f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = f(x_1\boldsymbol{u}^1 + x_2\boldsymbol{u}^2 + x_3\boldsymbol{u}^3) = x_1 f(\boldsymbol{u}^1) + x_2 f(\boldsymbol{u}^2) + x_3 f(\boldsymbol{u}^3) = x_1\begin{pmatrix} 2 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} -1 \\ 1 \end{pmatrix} + x_3\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$ [4]

Quello che abbiamo ottenuto è una c.l. dei vettori $f(\boldsymbol{u}^1), f(\boldsymbol{u}^2), f(\boldsymbol{u}^3)$, che sono le immagini dei vettori fondamentali di $\mathbb{R}^3$ attraverso la trasformazione f.

Pertanto la f può essere rappresentata da questi soli vettori, questa volta disposti in *colonna*. Se li affianchiamo in una tabella, otteniamo la stessa matrice di prima.

Riassumendo, la matrice

$$\begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{rappresenta la trasformazione } f \text{ e la scrittura} \quad \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

è un modo di scrivere il vettore $f(\boldsymbol{x})$ come prodotto tra la matrice ed il vettore $\boldsymbol{x}$. Tale vettore $f(\boldsymbol{x})$ può essere ottenuto in due modi:

- facendo i prodotti interni delle righe della matrice per la colonna $\boldsymbol{x}$
- facendo la c.l. delle colonne della matrice con coefficienti dati dalle componenti del vettore $\boldsymbol{x}$.

In generale allora, se f è una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$, dalla linearità segue che

$$\begin{aligned} f(\boldsymbol{x}) &= f(x_1\boldsymbol{u}^1 + x_2\boldsymbol{u}^2 + \ldots + x_n\boldsymbol{u}^n) \\ &= x_1 f(\boldsymbol{u}^1) + x_2 f(\boldsymbol{u}^2) + \ldots + x_n f(\boldsymbol{u}^n) \\ &= A\boldsymbol{x}, \end{aligned}$$

[3] Si tratta del prodotto interno tra i due vettori $\boldsymbol{a}$ e $\boldsymbol{x}$.

[4] Dato che $f(\boldsymbol{u}^1) = f\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right)$; $f(\boldsymbol{u}^2) = f\left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$; $f(\boldsymbol{u}^3) = f\left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$.

intendendo che A è la matrice ottenuta disponendo in colonna le immagini dei vettori fondamentali del dominio di f (cioè $\mathbb{R}^n$) e che $A\boldsymbol{x}$ rappresenta una c.l. delle colone di A con coefficienti dati dalle componenti di $\boldsymbol{x}$.

La matrice A si chiama **matrice di rappresentazione** della trasformazione lineare f.

Osservazione Si noti anche che, viceversa, data una generica matrice A di m righe ed n colonne, se diamo alla scrittura $A\boldsymbol{x}$ il significato appena detto, e cioè

$$A\boldsymbol{x} = x_1\boldsymbol{a}^1 + x_2\boldsymbol{a}^2 + \ldots + x_n\boldsymbol{a}^n,$$

dove gli $\boldsymbol{a}^j$ sono le colonne della matrice A, allora la scrittura $f(\boldsymbol{x}) = A\boldsymbol{x}$, per ogni $\boldsymbol{x} \in \mathbb{R}^n$ definisce una funzione di $\mathbb{R}^n$ in $\mathbb{R}^m$. Tale funzione è una trasformazione lineare.[5] Pertanto possiamo dire che c'è una corrispondenza biunivoca tra l'insieme delle trasformazioni lineari di $\mathbb{R}^n$ in $\mathbb{R}^m$ e le matrici di m righe ed n colonne. Questo risultato viene talvolta indicato come *teorema di rappresentazione delle trasformazioni lineari*.

Osservazione Si noti che se f è una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$, allora la sua matrice di rappresentazione ha m righe (come la dimensione di $\mathbb{R}^m$) ed n colonne (come la dimensione di $\mathbb{R}^n$).

Osservazione Un aspetto importante, sul quale d'altro canto dico subito che non mi soffermerò, è il seguente: non appena c'è l'esigenza di dare ad una trasformazione lineare una qualche rappresentazione, occorre scrivere i vettori rispetto ad una qualche base dei rispettivi spazi vettoriali. In quanto fatto finora si è implicitamente inteso che tale rappresentazione sia rispetto alle basi fondamentali (di $\mathbb{R}^n$ e di $\mathbb{R}^m$). La trasformazione *non dipende* dal sistema di riferimento, in quanto stabilisce una corrispondenza tra vettori di due spazi, ma la matrice che rappresenta la trasformazione *dipende* invece dalla base che viene utilizzata.[6] Possiamo dire che ogni trasformazione lineare può essere rappresentata da molte matrici, a seconda della base, e che una matrice può rappresentare molte trasformazioni lineari, sempre a seconda della base scelta. Non studieremo le questioni legate al cambio di base. Per noi si tratterà sempre di rappresentare rispetto alle basi fondamentali.

Esercizio 1.1 Scrivere la matrice di rappresentazione delle seguenti trasformazioni lineari, indicando poi tra quali spazi è definita la trasformazione stessa.

$$\text{(a)}\quad f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_2 + 5x_3 \\ x_2 - x_3 \end{pmatrix} \qquad \text{(b)}\quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ 3x_1 + 2x_2 \\ x_1 \end{pmatrix} \qquad \text{(c)}\quad f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_1 \\ x_2 \end{pmatrix}$$

Esercizio 1.2 Si scriva l'espressione analitica della trasformazione lineare rappresentata dalle seguenti matrici A, precisando poi tra quali spazi è definita la trasformazione stessa.

$$\text{(a)}\quad A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{(b)}\quad A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \qquad \text{(c)}\quad A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Esercizio 1.3 Verificare che le seguenti trasformazioni *non* sono lineari.

$$\text{(a)}\quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 1 \\ x_2 - 1 \end{pmatrix} \qquad \text{(b)}\quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1x_2 \\ x_1 + x_2 \end{pmatrix} \qquad \text{(c)}\quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{x_1x_2}{x_1^2 + x_2^2 + 1}$$

[5] Infatti

$$\begin{array}{rcl}
\text{(per definizione di } f\text{)}\quad f(\boldsymbol{x}+\boldsymbol{y}) & = & A(\boldsymbol{x}+\boldsymbol{y}) \\
\text{(per definizione di } A\boldsymbol{x}\text{)} & = & (x_1+y_1)\boldsymbol{a}^1 + \ldots + (x_n+y_n)\boldsymbol{a}^n \\
\text{(prop. distributiva della molt. scalare in } \mathbb{R}^m\text{)} & = & x_1\boldsymbol{a}^1 + y_1\boldsymbol{a}^1 + \ldots + x_n\boldsymbol{a}^n + y_n\boldsymbol{a}^n \\
\text{(per definizione di } A\boldsymbol{x}\text{)} & = & A\boldsymbol{x} + A\boldsymbol{y} \\
\text{(per definizione di } f\text{)} & = & f(\boldsymbol{x}) + f(\boldsymbol{y}).
\end{array}$$

Così è dimostrata l'additività di f.

Analogamente (lo studente rifletta su quali proprietà vengono utilizzate per ottenere i passaggi indicati)

$$f(\alpha\boldsymbol{x}) = A(\alpha\boldsymbol{x}) = (\alpha x_1)\boldsymbol{a}^1 + \ldots + (\alpha x_n)\boldsymbol{a}^n = \alpha(x_1\boldsymbol{a}^1) + \ldots + \alpha(x_n\boldsymbol{a}^n) = \alpha A\boldsymbol{x} = \alpha f(\boldsymbol{x}),$$

e questo prova l'omogeneità.

[6] Una qualche analogia con questo aspetto la possiamo rilevare ad esempio nella seguente osservazione: la distanza che separa due località non dipende da come la esprimiamo, ma la sua misura sì, cambia cioè se la esprimiamo in metri o in millimetri.

2 Matrici

Una matrice quindi non è che un modo per rappresentare una trasformazione lineare.

Solitamente le matrici si indicano con lettere maiuscole. Come già osservato, la matrice che rappresenta una trasformazione di $\mathbb{R}^n$ in $\mathbb{R}^m$ ha m righe ed n colonne: diremo che è una matrice $m \times n$.

Possiamo vedere un vettore riga come una matrice $1 \times n$, cioè come la rappresentazione di una $f : \mathbb{R}^n \to \mathbb{R}$, e un vettore colonna come una matrice $m \times 1$, e quindi come la rappresentazione di una $f : \mathbb{R} \to \mathbb{R}^m$. In alcuni casi potrà essere utile vedere in una matrice solo l'aspetto generale di rappresentazione di una trasformazione lineare, in altri sarà utile vedere aspetti più particolari, come ad esempio i singoli elementi della matrice. Chiameremo *elemento di posto* (i, j) il numero che si trova nella riga i e colonna j della matrice. Se la matrice viene indicata con A, è consuetudine indicare con a_{ij} il suo elemento di posto (i, j).

$$
\begin{array}{cc}
 & \begin{array}{ccccc} \phantom{a_{11}} & & \text{colonna } j & & \phantom{a_{1n}} \\ & & \downarrow & & \end{array} \\
\text{riga } i \to & \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}
\end{array}
$$

Osservazioni

- Gli elementi della 1ª riga sono caratterizzati dal fatto che il primo indice è per tutti 1. Quindi sono gli elementi del tipo a_{1j}, con $j = 1, 2, \dots, n$.
- Gli elementi della 1ª colonna sono caratterizzati dal fatto che il secondo indice è per tutti 1. Quindi sono gli elementi del tipo a_{i1}, con $i = 1, 2, \dots, m$.
- Gli elementi della i-esima riga sono quelli del tipo a_{ij}, con $j = 1, 2, \dots, n$.
- Gli elementi della j-esima colonna sono quelli del tipo a_{ij}, con $i = 1, 2, \dots, m$.

La matrice di rappresentazione della trasformazione nulla[7] è la *matrice nulla* e tutti i suoi elementi sono nulli.

Se A è la matrice di rappresentazione della trasformazione f, allora la matrice che rappresenta $-f$ si indica con $-A$ e si chiama l'*opposta* di A: il suo elemento di posto (i, j) è l'opposto del corrispondente elemento di A.

Una matrice si dice *quadrata di ordine* n se rappresenta una trasformazione $f : \mathbb{R}^n \to \mathbb{R}^n$. In una matrice quadrata il numero delle righe è uguale al numero delle colonne. Tale numero è detto appunto l'*ordine* della matrice quadrata. Quindi, dicendo ad esempio che A è una matrice quadrata di ordine 3, si intende che A è una matrice 3×3, e rappresenta cioè una trasformazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$.

In una matrice quadrata gli elementi di posto (i, j) con $i = j$ formano la *diagonale principale* della matrice. Ad esempio, nella matrice quadrata 3×3

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \text{la diagonale principale è data dagli elementi } a_{11} = 1,\ a_{22} = 5,\ a_{33} = 9.$$

Definizione Se A è una matrice $m \times n$, si chiama matrice **trasposta** di A, e si indica con A^T, la matrice di n righe ed m colonne il cui elemento di posto (i, j) è a_{ji}.

Osservazione Quindi l'elemento di posto (i, j) della trasposta si ottiene prendendo l'elemento di indici scambiati della matrice A. Si può anche dire che A^T è la matrice che si ottiene da A scambiando le righe con le colonne.

Esempio La matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{ha come matrice trasposta la matrice} \quad \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Non si tratta, come forse si potrebbe pensare, della matrice che rappresenta la trasformazione inversa.[8]

Definizione Una matrice quadrata A si dice *simmetrica* se $A^T = A$, e quindi se $a_{ij} = a_{ji}$, per ogni i, j.

Esempio E' simmetrica la matrice

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

[7] La trasformazione nulla è quella che trasforma ogni vettore nel vettore nullo, cioè quella tale che $f(\boldsymbol{x}) = \mathbf{0}$ per ogni $\boldsymbol{x} \in \mathbb{R}^n$.

[8] Di trasformazione inversa parleremo alla fine di questa dispensa.

Osservazione Per essere simmetrica una matrice deve essere quadrata. Si tratta poi di una simmetrica rispetto alla diagonale principale.

Definizione Una matrice quadrata A è *diagonale* se risulta $a_{ij} = 0$ per ogni i, j, con $i \neq j$.

Quindi una matrice è diagonale se sono nulli gli elementi che *non* stanno sulla diagonale principale. Faccio osservare che nella definizione non si fa riferimento agli elementi che stanno sulla diagonale principale, i quali possono essere nulli oppure no.

Esempi Sono ad esempio matrici diagonali

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Risulta chiaro dalle definizioni che una matrice diagonale è anche simmetrica.

Un particolare ed importante esempio di matrice diagonale è la matrice di rappresentazione della trasformazione lineare *identità* da $\mathbb{R}^n$ ad $\mathbb{R}^n$. Si chiama identità da $\mathbb{R}^n$ ad $\mathbb{R}^n$ la trasformazione che associa ad ogni $\boldsymbol{x}$ lo stesso $\boldsymbol{x}$, cioè la funzione $I_n(\boldsymbol{x}) = \boldsymbol{x}$, per ogni $\boldsymbol{x} \in \mathbb{R}^n$. È un semplice esercizio provare che tale trasformazione è lineare. Cerchiamo di capire com'è fatta la matrice che rappresenta tale trasformazione. Dato che

$$I_n(\boldsymbol{x}) = \boldsymbol{x} = x_1\boldsymbol{u}^1 + x_2\boldsymbol{u}^2 + \ldots + x_n\boldsymbol{u}^n,$$

la matrice di rappresentazione è la matrice che ha come colonne i vettori fondamentali. Quindi questa matrice ha 1 sulla diagonale principale e 0 fuori di questa. Tale matrice si chiama ovviamente **matrice identità** e si indica di solito con I (oppure I_n se si vuole precisare che si tratta dell'identità in $\mathbb{R}^n$). Alcuni esempi sono

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

È naturale a questo punto definire nell'insieme delle matrici alcune **operazioni**.

Le prime due sono immediate, perché derivano direttamente dalle operazioni definite sulle trasformazioni lineari, di cui le matrici sono rappresentazioni.

Nell'insieme delle matrici $m \times n$, se A, B rappresentano rispettivamente due trasformazioni f, g di $\mathbb{R}^n$ in $\mathbb{R}^m$, la matrice che rappresenta la trasformazione somma $f + g$ deve necessariamente essere quella che si ottiene da A e B sommando gli elementi di posto corrispondente. Tale matrice verrà indicata con $A + B$.

Analogamente, se A rappresenta la trasformazione f e $\alpha \in \mathbb{R}$, la matrice che rappresenta la trasformazione αf deve necessariamente essere quella che si ottiene da A moltiplicando tutti i suoi elementi per α. Si indica naturalmente con αA.

Abbiamo quindi definito nell'insieme delle matrici $m \times n$ un'addizione tra matrici e una moltiplicazione di una matrice per uno scalare. Lo 0 (elemento neutro dell'addizione) per le matrici $m \times n$ è chiaramente la matrice nulla $m \times n$.

Esercizio Dimostrare che la somma di matrici diagonali è una matrice diagonale.

A questo punto ci si aspetta un prodotto tra matrici e la prima cosa che viene in mente è la matrice che rappresenta il prodotto delle due trasformazioni. Potremmo vedere che questa strada non si rivela la più interessante, semplicemente perché non disponiamo di un "interessante" prodotto di trasformazioni. Il problema è che se f e g sono due trasformazioni, $f(\boldsymbol{x})$ e $g(\boldsymbol{x})$ non sono in genere numeri reali ma vettori. La via da seguire è quella di pensare al prodotto di due trasformazioni in termini di composizione delle due.[9] Ricordo intanto che, se $f, g : \mathbb{R}^n \to \mathbb{R}^m$, allora non si può in generale parlare di funzione composta $f \circ g$ (oppure $g \circ f$) a meno che non sia $m = n$. [10]

Il caso più generale in cui esiste la trasformazione composta è quello in cui $g : \mathbb{R}^n \to \mathbb{R}^m$ e $f : \mathbb{R}^m \to \mathbb{R}^p$. Infatti allora esiste la trasformazione composta $f \circ g : \mathbb{R}^n \to \mathbb{R}^p$, dato che il codominio di g coincide con il dominio di f. Parlando di composizione delle trasformazioni lineari vale il seguente risultato:

Teorema Se $g : \mathbb{R}^n \to \mathbb{R}^m$ e $f : \mathbb{R}^m \to \mathbb{R}^p$ sono due trasformazioni lineari, allora $f \circ g$ è una trasformazione lineare da $\mathbb{R}^n$ ad $\mathbb{R}^p$.

[9] Qui sarebbe forse opportuno che lo studente, prima di continuare, andasse a rivedere le questioni legate alla composizione delle funzioni, cioè alla definizione di funzione composta, vista nella seconda parte del corso.

[10] Si ricordi che $(f \circ g)(\boldsymbol{x})$ significa $f(g(\boldsymbol{x}))$ e che quindi occorre che $g(\boldsymbol{x})$ appartenga al dominio di f. Pertanto se le due trasformazioni sono definite tra gli stessi spazi deve essere $m = n$.

Osservazione La tesi importante del teorema non è che $f \circ g$ è una trasformazione da $\mathbb{R}^n$ a $\mathbb{R}^p$ (questo è implicito nella definizione di trasformazione composta), ma è piuttosto che tale trasformazione composta è lineare, cosa che non è scontata e che va quindi dimostrata. E la dimostrazione non è difficile.

Dimostrazione Occorre dimostrare l'additività e l'omogeneità della funzione composta. Si ha

$$(f \circ g)(\boldsymbol{x} + \boldsymbol{y}) = f(g(\boldsymbol{x} + \boldsymbol{y})) = f(g(\boldsymbol{x}) + g(\boldsymbol{y})) = f(g(\boldsymbol{x})) + f(g(\boldsymbol{y})) = (f \circ g)(\boldsymbol{x}) + (f \circ g)(\boldsymbol{y})$$

e

$$(f \circ g)(\alpha\boldsymbol{x}) = f(g(\alpha\boldsymbol{x})) = f(\alpha g(\boldsymbol{x})) = \alpha f(g(\boldsymbol{x})) = \alpha(f \circ g)(\boldsymbol{x}).$$

Quindi la linearità è dimostrata.

La cosa interessante è ora capire come si ottiene la matrice di rappresentazione della trasformazione composta. Non vediamo i dettagli formali, do semplicemente il risultato: se la matrice A rappresenta la trasformazione f e la matrice B rappresenta la trasformazione g, allora la trasformazione composta $f \circ g$ è rappresentata dalla matrice AB, che si dice matrice **prodotto righe per colonne** di A per B e che è la matrice che ha come elemento di posto (i, j) il prodotto interno della i-esima riga della prima (cioè A) per la j-esima colonna della seconda (cioè B).

Esempio Siano

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

La matrice A rappresenta una trasformazione lineare f da $\mathbb{R}^3$ a $\mathbb{R}^2$; la matrice B rappresenta una trasformazione lineare g da $\mathbb{R}^2$ a $\mathbb{R}^3$.

È quindi possibile costruire la trasformazione composta $f \circ g$ che, attenzione, va da $\mathbb{R}^2$ a $\mathbb{R}^2$. Il prodotto righe per colonne AB è quindi una matrice 2×2. Risulta

$$AB = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 3 & 2 \end{pmatrix}.$$ [11]

In questo caso è possibile costruire anche la trasformazione composta $g \circ f$, che sarà una trasformazione da $\mathbb{R}^3$ a $\mathbb{R}^3$. Il prodotto righe per colonne BA è una matrice 3×3. Risulta

$$BA = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & -1 \end{pmatrix}.$$

Seguono ora alcune osservazioni.

Osservazioni

- Si intuisce che, nel prodotto di due matrici, come nella composizione di due trasformazioni, l'ordine è importante.
- Date due matrici A e B, per poter effettuare il prodotto righe per colonne AB è necessario e sufficiente che il numero di colonne di A sia uguale al numero di righe di B.
- Il prodotto righe per colonne AB (quando si può fare) è una matrice che ha tante righe quante ne ha A e tante colonne quante ne ha B (si dia una motivazione a questa regola pensando al significato delle due matrici).
- Il prodotto tra matrici ha come caso particolare il prodotto di una matrice per un vettore, o di un vettore per una matrice, o anche di un vettore per un vettore.

  Ad esempio, se

  $$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \qquad \text{e} \qquad x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

  possiamo scrivere

  $$Ax = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}.$$ [12]

[11] L'elemento di posto $(1, 1)$ (lo 0) è il prodotto interno della 1^{a} riga di A per la 1^{a} colonna di B, l'elemento di posto $(1, 2)$ (il 2) è il prodotto interno della 1^{a} riga di A per la 2^{a} colonna di B, e così via.

Si considerino anche questi altri esempi:

$$(1\ \ -1)\begin{pmatrix}1 & 0 & -1\\ 1 & 2 & 0\end{pmatrix} = (0\ \ -2\ \ -1) \quad , \quad (1\ \ 2\ \ 3)\begin{pmatrix}4\\5\\6\end{pmatrix} = 32 \quad , \quad \begin{pmatrix}1\\2\end{pmatrix}(3\ \ 4) = \begin{pmatrix}3 & 4\\ 6 & 8\end{pmatrix}.$$

- Un caso particolare in cui il prodotto tra due matrici si può fare è quando le matrici sono quadrate (ovviamente devono essere anche dello stesso ordine). Questo perché la composizione di due trasformazioni lineari di uno spazio in sé ovviamente è sempre possibile. Si verifica facilmente su qualche esempio che in questo caso il prodotto *non è commutativo*. Quindi, se A e B sono matrici quadrate, può succedere che AB sia diversa da BA.

  Consideriamo il seguente esempio:

  $$A = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

  Risulta

  $$AB = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}2 & 1\\ 1 & 0\end{pmatrix}$$

  mentre

  $$BA = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 2\end{pmatrix}.$$

  Nel prodotto tra matrici occorre dunque fare attenzione all'ordine in cui si effettua l'operazione: quindi moltiplicare A *a destra* per B non equivale in genere a moltiplicare A *a sinistra* per B. [13]

- Non vale, nel prodotto tra matrici, la *legge di annullamento del prodotto*, alla quale siamo abituati in $\mathbb{R}$. [14]

  Può cioè succedere che il prodotto di due matrici non nulle sia la matrice nulla.

  Ad esempio, questo succede con

  $$A = \begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}.$$

- Nell'insieme delle matrici quadrate $n\times n$ la matrice identità I_n è elemento neutro del prodotto. È facile dimostrare infatti che si ha $AI_n = I_nA = A$, qualunque sia la matrice A $n \times n$.

Osservazione Si può dimostrare in generale che, dovendo fare la matrice trasposta di un prodotto di due matrici A e B, si può fare il prodotto delle trasposte, ma occorre scambiare l'ordine. Vale cioè la formula generale

$$(AB)^T = B^TA^T.$$

Non è così difficile da dimostrare: l'elemento di posto (i, j) di $(AB)^T$ è l'elemento di posto (j, i) di AB (la trasposta scambia le righe con le colonne); quindi si tratta del prodotto interno

$$\big\langle j\text{-esima riga di } A\ ,\ i\text{-esima colonna di } B\big\rangle.$$

L'elemento di posto (i, j) di B^TA^T è il prodotto interno

$$\big\langle i\text{-esima riga di } B^T\ ,\ j\text{-esima colonna di } A^T\big\rangle,$$

[12] Quindi la scrittura $A\boldsymbol{x}$, utilizzata già da un po' per indicare una c.l. delle colonne di A con coefficienti in x, non è altro che un prodotto righe per colonne di due matrici, delle quali la seconda è un vettore colonna. Si ricordi che prima avevamo scritto

$$Ax = \begin{pmatrix}0 & 1 & -1\\ 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}\cdot 1 + \begin{pmatrix}1\\2\end{pmatrix}\cdot 2 + \begin{pmatrix}-1\\0\end{pmatrix}\cdot 3 = \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}-3\\0\end{pmatrix} = \begin{pmatrix}-1\\5\end{pmatrix}.$$

[13] Naturalmente, come già detto prima, se le matrici non sono quadrate, può succedere che sia possibile moltiplicare A a destra per B ma non a sinistra.

[14] La legge di annullamento del prodotto dice che il prodotto di due numeri reali è nullo se e solo se almeno uno dei due numeri è nullo. Con le matrici non vale.

cioè il

$$\langle i\text{-esima colonna di } B\ ,\ j\text{-esima riga di } A\rangle.$$

Per la proprietà commutativa del prodotto interno si tratta dello stesso valore. Quindi le due matrici sono uguali.

Esempio Verifichiamo questo risultato su di un semplice esempio: siano

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

Abbiamo

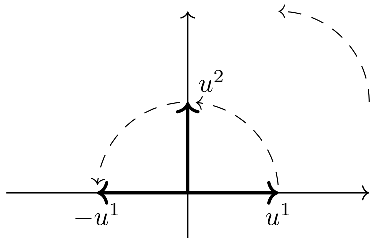
$$AB = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \quad \text{e quindi} \quad (AB)^T = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

Ora calcoliamo $B^T A^T$:

$$B^T A^T = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}. \quad \text{Si ha invece} \quad A^T B^T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Chiudiamo il paragrafo con un paio di esempi importanti. In precedenza abbiamo visto come dalla definizione analitica di una trasformazione lineare (ciè dall'espressione di $f(\boldsymbol{x})$) si possa facilmente trovare la sua matrice di rappresentazione. Ora vediamo come si possa ottenere la matrice di rappresentazione di una trasformazione definita invece attraverso alcuni suoi aspetti geometrici.

Esempio Consideriamo la trasformazione del piano in sé (quindi di $\mathbb{R}^2$ in $\mathbb{R}^2$) individuata dalla rotazione del piano di un angolo di $\pi/2$ radianti.[15]
Le questioni aperte sono due: constatare che si tratta di una trasformazione lineare e trovare poi la matrice di rappresentazione.
Lasciando da parte gli aspetti formali, non è difficile capire che la trasformazione è in effetti lineare: per quanto riguarda l'additività, se abbiamo due vettori $\boldsymbol{x}, \boldsymbol{y}$ e la loro somma $\boldsymbol{x} + \boldsymbol{y}$ (si pensi al parallelogramma), allora sottoponendo $\boldsymbol{x} + \boldsymbol{y}$ alla rotazione si ottiene lo stesso vettore che si ottiene sottoponendo $\boldsymbol{x}$ e $\boldsymbol{y}$ alla rotazione e poi calcolandone la somma. Ancora più evidente è l'omogeneità: la rotazione di un vettore $\alpha\boldsymbol{x}$ porta allo stesso vettore che si ottiene ruotando prima $\boldsymbol{x}$ e moltiplicando poi quest'ultimo per α.



Troviamo ora la matrice di rappresentazione di questa rotazione del piano in sé. Qui occorre ricordare che la matrice ha per colonne i trasformati dei vettori fondamentali. Quindi, se indichiamo con r la trasformazione in questione ($r : \mathbb{R}^2 \to \mathbb{R}^2$), e con R la sua matrice di rappresentazione, allora le due colonne di R sono

$$r(\boldsymbol{u}^1) = r\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \boldsymbol{u}^2 \qquad \text{e} \qquad r(\boldsymbol{u}^2) = r\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\boldsymbol{u}^1.$$

La matrice di rappresentazione della rotazione è quindi la matrice

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Moltiplicando per questa matrice un qualunque vettore di $\mathbb{R}^2$ si ottiene il vettore trasformato, cioè quello ottenuto dalla sua rotazione di 90 gradi.[16]

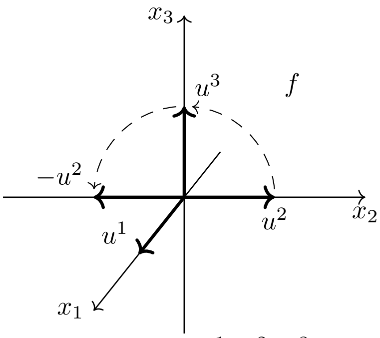
Esempio La non commutatività del prodotto tra matrici non è che la conseguenza della non commutatività della composizione di trasformazioni lineari. Questo significa ad esempio che se noi operiamo due rotazioni dello spazio $\mathbb{R}^3$ in sé, la trasformazione risultante dipende dall'ordine con cui operiamo le due trasformazioni. Verifichiamo per esercizio il tutto su di un esempio particolare.

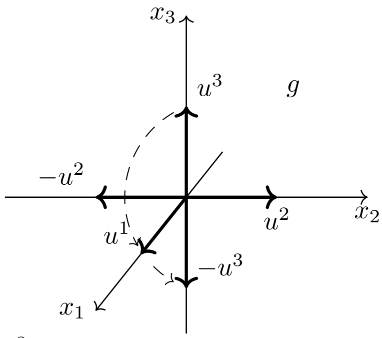
Consideriamo la trasformazione lineare f nello spazio tridimensionale (vedi figura sotto a sinistra) che mantiene fisso l'asse x_1 e fa ruotare in senso antiorario gli assi x_2 e x_3 di un angolo di $\pi/2$. [17] Consideriamo poi la trasformazione g che mantiene fisso l'asse x_2 e fa ruotare in senso antiorario gli assi x_1 e x_3 di un angolo di $\pi/2$ (a destra).

[15] Si tratta quindi di una rotazione di 90 gradi. Si faccia anche attenzione, per fissare le idee, al fatto che la rotazione è di un angolo positivo e quindi è una rotazione in senso antiorario.

[16] Ad esempio, il trasformato del vettore $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ è il vettore $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, come si verifica facilmente sul grafico.

[17] Si rifletta sul fatto che tale trasformazione lineare è univocamente individuata da queste condizioni.





Indicando come sempre con $\boldsymbol{u}^1, \boldsymbol{u}^2, \boldsymbol{u}^3$ i vettori fondamentali di $\mathbb{R}^3$, possiamo dire che le due trasformazioni sono univocamente individuate dalle seguenti condizioni (si verifichi sulle figure qui sopra):

$$\begin{aligned} f(\boldsymbol{u}^1) &= \boldsymbol{u}^1 \\ f(\boldsymbol{u}^2) &= \boldsymbol{u}^3 \\ f(\boldsymbol{u}^3) &= -\boldsymbol{u}^2 \end{aligned} \qquad \text{da cui la matrice di rappresentazione di } f \text{ è} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix};$$

$$\begin{aligned} g(\boldsymbol{u}^1) &= -\boldsymbol{u}^3 \\ g(\boldsymbol{u}^2) &= \boldsymbol{u}^2 \\ g(\boldsymbol{u}^3) &= \boldsymbol{u}^1 \end{aligned} \qquad \text{da cui la matrice di rappresentazione di } g \text{ è} \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Costruiamo ora le trasformazioni composte: con $f \circ g$ si ha

$$\begin{aligned} (f\circ g)(\boldsymbol{u}^1) &= f(g(\boldsymbol{u}^1)) = f(-\boldsymbol{u}^3) = \boldsymbol{u}^2 \\ (f\circ g)(\boldsymbol{u}^2) &= f(g(\boldsymbol{u}^2)) = f(\boldsymbol{u}^2) = \boldsymbol{u}^3 \\ (f\circ g)(\boldsymbol{u}^3) &= f(g(\boldsymbol{u}^3)) = f(\boldsymbol{u}^1) = \boldsymbol{u}^1 \end{aligned} \qquad \text{con matrice} \quad AB = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Con $g \circ f$ invece si ha

$$\begin{aligned} (g\circ f)(\boldsymbol{u}^1) &= g(f(\boldsymbol{u}^1)) = g(\boldsymbol{u}^1) = -\boldsymbol{u}^3 \\ (g\circ f)(\boldsymbol{u}^2) &= g(f(\boldsymbol{u}^2)) = g(\boldsymbol{u}^3) = \boldsymbol{u}^1 \\ (g\circ f)(\boldsymbol{u}^3) &= g(f(\boldsymbol{u}^3)) = g(-\boldsymbol{u}^2) = -\boldsymbol{u}^2 \end{aligned} \qquad \text{con matrice} \quad BA = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}.$$

Lo studente verifichi che le matrici prodotto sono quelle che si ottengono eseguendo il prodotto righe per colonne.

Esercizio 2.1 Date le seguenti matrici, si scriva per ciascuna la matrice trasposta e poi si calcolino i prodotti a due a due, nei casi in cui è possibile.

$$A = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

Esercizio 2.2 Date le due trasformazioni lineari

$$f: \mathbb{R}^2 \to \mathbb{R}^3 \quad \text{con} \quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_1 - 2x_2 \end{pmatrix} \qquad \text{e} \qquad g: \mathbb{R}^3 \to \mathbb{R}^2 \quad \text{con} \quad g\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} y_2 + y_3 \\ y_1 - y_2 \end{pmatrix},$$

si scriva l'espressione analitica delle trasformazioni composte $g \circ f$ e $f \circ g$.

3 Immagine di una trasformazione lineare

Sia f una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$.

Definizione Definiamo **immagine** di f l'insieme

$$\operatorname{Im} f = \{ f(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^n \}.$$ [18]

Osservazione Ovviamente $\operatorname{Im} f \subseteq \mathbb{R}^m$. Inoltre, dal fatto che $f(\mathbf{0}) = \mathbf{0}$ segue che l'immagine di f non può essere vuota.

Teorema $\operatorname{Im} f$ è un sottospazio di $\mathbb{R}^m$.

Osservazione Allo studente non sfugga che la definizione di immagine di f definisce un particolare sottoinsieme di $\mathbb{R}^m$. Il teorema prova che questo sottoinsieme ha una struttura ben precisa, quella di *sottospazio*. Dire sottospazio è ben diverso dal dire sottoinsieme.

Dimostrazione Se $\boldsymbol{y}^1, \boldsymbol{y}^2 \in \operatorname{Im} f$, significa che esistono $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathbb{R}^n$ tali che $\boldsymbol{y}^1 = f(\boldsymbol{x}^1)$ e $\boldsymbol{y}^2 = f(\boldsymbol{x}^2)$. Allora, per la linearità di f, se $\alpha_1, \alpha_2 \in \mathbb{R}$, si ottiene

$$\alpha_1 \boldsymbol{y}^1 + \alpha_2 \boldsymbol{y}^2 = \alpha_1 f(\boldsymbol{x}^1) + \alpha_2 f(\boldsymbol{x}^2) = f(\alpha_1 \boldsymbol{x}^1 + \alpha_2 \boldsymbol{x}^2),$$

e questo significa che $\alpha_1 \boldsymbol{y}^1 + \alpha_2 \boldsymbol{y}^2$ è immagine di un vettore di $\mathbb{R}^n$ e cioè appartiene a $\operatorname{Im} f$. Abbiamo quindi provato che facendo una c.l. di due elementi di $\operatorname{Im} f$ si ottiene ancora un elemento di $\operatorname{Im} f$, e quindi che $\operatorname{Im} f$ è un sottospazio.

Osservazione Dato che ogni $\boldsymbol{y} \in \operatorname{Im} f$ è c.l. dei vettori $f(\boldsymbol{u}^j)$, cioè i trasformati dei vettori fondamentali di $\mathbb{R}^n$,[19] allora il sottospazio $\operatorname{Im} f$ è generato dai vettori $f(\boldsymbol{u}^j)$. E, dato che tali vettori sono le colonne della matrice di rappresentazione di f, possiamo anche dire che $\operatorname{Im} f$ è generato dalle colonne di tale matrice.

Esempio Si consideri la matrice

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

Essa rappresenta una trasformazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$. Sappiamo già come ottenere l'espressione di questa f. [20]

Ora possiamo dire che l'immagine di f è un sottospazio di $\mathbb{R}^2$ e che questo sottospazio è generato dai vettori $\binom{1}{-1}$, $\binom{-1}{1}$, $\binom{2}{-2}$, cioè che tutti gli elementi dell'immagine di f si possono scrivere come un'opportuna c.l. di questi tre. La domanda che può nascere è: l'immagine di f è in realtà tutto lo spazio $\mathbb{R}^2$ (cioè si ha $\dim \operatorname{Im} f = 2$) oppure è un sottospazio non banale (cioè si ha $\dim \operatorname{Im} f < 2$)? In realtà sappiamo già dare una risposta: l'immagine di f è tutto $\mathbb{R}^2$ se e soltanto se tra i vettori colonna della matrice A ce ne sono almeno due che risultano linearmente indipendenti, cosa che in questo caso è falsa (provare a dimostrarlo). In questo caso si ha $\dim \operatorname{Im} f = 1$.

Definizione Se f è una trasformazione lineare di $\mathbb{R}^n$ in $\mathbb{R}^m$, si definisce **rango** di f e si indica con $r(f)$ la dimensione del sottospazio $\operatorname{Im} f$.

Esempio Data la trasformazione lineare $f : \mathbb{R}^4 \to \mathbb{R}^3$ definita da

$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x - y + t \\ 0 \\ y + z - t \end{pmatrix},$$

determiniamo la dimensione di $\operatorname{Im} f$. Troviamo poi una base di $\operatorname{Im} f$.

Dall'identità

$$\begin{pmatrix} x - y + t \\ 0 \\ y + z - t \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

[18] Non è una novità: si tratta di quello che abbiamo sempre chiamato immagine di una funzione. Si noti che si può scrivere in modo equivalente

$$\operatorname{Im} f = \{ \boldsymbol{y} \in \mathbb{R}^m : \exists \boldsymbol{x} \in \mathbb{R}^n \text{ tale che } \boldsymbol{y} = f(\boldsymbol{x}) \}.$$

Si ha anche, con altra notazione, $\operatorname{Im} f = f(\mathbb{R}^n)$.

[19] Ogni $\boldsymbol{y}$ che sta in $\operatorname{Im} f$ è un $f(\boldsymbol{x})$, per qualche $\boldsymbol{x}$ in $\mathbb{R}^n$, e si può sempre scrivere $f(\boldsymbol{x}) = x_1 f(\boldsymbol{u}^1) + \ldots x_n f(\boldsymbol{u}^n)$.

[20] Ricordo che l'espressione di f è l'espressione analitica di $f(\boldsymbol{x})$. In questo caso è

$$f(\boldsymbol{x}) = A\boldsymbol{x} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 + 2x_3 \\ -x_1 + x_2 - 2x_3 \end{pmatrix}.$$

si ricava che i vettori

$$\boldsymbol{v}^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{v}^2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \ \boldsymbol{v}^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \ \boldsymbol{v}^4 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

sono generatori di $\operatorname{Im} f$ (sono le colonne della matrice che rappresenta f). Tali vettori sono certamente linearmente dipendenti, dato che sono 4 vettori in uno spazio di dimensione 3. Possiamo osservare poi che $\boldsymbol{v}^4 = -\boldsymbol{v}^2$, quindi $\boldsymbol{v}^4$ dipende linearmente dagli altri 3 e lo possiamo eliminare. Ma si ha anche $\boldsymbol{v}^3 = \boldsymbol{v}^1 + \boldsymbol{v}^2$, e quindi anche $\boldsymbol{v}^3$ dipende linearmente dagli altri 2 e può essere eliminato. Restano $\boldsymbol{v}^1$ e $\boldsymbol{v}^2$, che sono indipendenti. Quindi $r(f) = \dim \operatorname{Im} f = 2$ e una base di $\operatorname{Im} f$ è formata appunto da $\boldsymbol{v}^1$ e $\boldsymbol{v}^2$.

La nozione di rango si estende alle matrici con la seguente naturale

Definizione Data una matrice $m \times n$ A, che rappresenta una trasformazione $f : \mathbb{R}^n \to \mathbb{R}^m$, si definisce **rango** di A e si indica con rA il rango della trasformazione f.

Osservazione Dato che il rango della matrice A è la dimensione di $\operatorname{Im} f$ e che questo spazio è generato dalle colonne di A, allora rA coincide con il *massimo numero di colonne l.i. di* A. Il rango è un concetto fondamentale, sia da un punto di vista teorico sia nelle applicazioni.

Concludiamo il paragrafo con un altro risultato importante, di cui non do dimostrazione.

Teorema Data una qualunque matrice A, risulta

$$rA = rA^T.$$

Osservazione Mettendo insieme i risultati e le osservazioni precedenti, possiamo allora dire che in una qualunque matrice il massimo numero di colonne indipendenti coincide con il massimo numero di righe indipendenti. Si noti ancora una volta che le righe e le colonne sono vettori appartenenti in genere a spazi diversi.

Non disponiamo ancora di un metodo "comodo" per il calcolo del rango di una matrice. Arriveremo a questo tra un po'. Rinvio anche ulteriori esercizi sulla determinazione dell'immagine di una trasformazione lineare a quando disporremo di ulteriori strumenti.

4 Inversione di una trasformazione lineare

In questo paragrafo ci poniamo il problema di studiare l'invertibilità di una trasformazione lineare. Dato che, per motivi che via via risulteranno chiari, l'unico caso significativo a questo proposito è quello delle trasformazioni lineari di uno spazio in sé (di $\mathbb{R}^n$ in $\mathbb{R}^n$), per la definizione seguente consideriamo una trasformazione di questo tipo.

Definizione La trasformazione $f : \mathbb{R}^n \to \mathbb{R}^n$ si dice **invertibile** se esiste $g : \mathbb{R}^n \to \mathbb{R}^n$ tale che

$$f \circ g = g \circ f = I_n,$$

avendo indicato con I_n la trasformazione identità su $\mathbb{R}^n$. [21]

Se f è invertibile la trasformazione g si chiama **inversa** di f e si indica con f^{-1}. Si può dimostrare che vale il seguente

Teorema Se f è una trasformazione lineare invertibile, allora anche f^{-1} è lineare.[22]

Parlando di funzioni in generale, nella seconda parte del corso, avevamo visto che una funzione è invertibile se e soltanto se è contemporaneamente iniettiva e suriettiva (si vada eventualmente a rivedere questi concetti). Il risultato naturalmente vale ancora parlando di funzioni che sono in particolare trasformazioni lineari. Per queste ultime le due proprietà si possono esprimere nel modo che ora vediamo.

Torniamo a considerare una generica trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$. La trasformazione f *può* avere queste due proprietà:

(i) Per ogni $\boldsymbol{x} \in \mathbb{R}^n$, se $f(\boldsymbol{x}) = \mathbf{0}$, allora $\boldsymbol{x} = \mathbf{0}$ (equivale a dire che l'unico vettore che viene trasformato nel vettore nullo è il vettore nullo);

(ii) Per ogni $\boldsymbol{y} \in \mathbb{R}^m$ esiste un $\boldsymbol{x} \in \mathbb{R}^n$ tale che $f(\boldsymbol{x}) = \boldsymbol{y}$ (cioè ogni vettore $\boldsymbol{y}$ di $\mathbb{R}^m$ è il trasformato di qualche $\boldsymbol{x}$).

[21] Lo studente noti che il concetto di invertibilità è lo stesso che abbiamo già incontrato per una funzione reale.
[22] Si noti che senza questo risultato non si poteva dare per scontato che l'inversa di una trasformazione lineare fosse anch'essa lineare.

Esercizio Si dimostri che la proprietà (i) equivale a dire che per ogni $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, se $f(\boldsymbol{x}) = f(\boldsymbol{y})$, allora $\boldsymbol{x} = \boldsymbol{y}$ e che questa a sua volta equivale a dire che per ogni $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, se $\boldsymbol{x} \neq \boldsymbol{y}$, allora $f(\boldsymbol{x}) \neq f(\boldsymbol{y})$.

Osservazione La proprietà (i) esprime quindi il fatto che f è *iniettiva.* La proprietà (ii) invece esprime il fatto che f è *suriettiva.* Possiamo dire che f è suriettiva se e solo se $\operatorname{Im} f = \mathbb{R}^m$ (mentre in generale si ha $\operatorname{Im} f = f(\mathbb{R}^n) \subset \mathbb{R}^m$), e cioè se e solo se il rango di f è m.

Sull'invertibilità di una trasformazione lineare sussiste sempre lo stesso risultato (valido per le funzioni in generale):

Teorema Una trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^n$ è invertibile se e solo se valgono le proprietà (i) e (ii) (contemporaneamente), cioè f è invertibile se e solo se è iniettiva *e* suriettiva.

Esempi Una trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$ può essere iniettiva ma non suriettiva, oppure suriettiva ma non iniettiva.

- Si consideri ad esempio la trasformazione lineare $f : \mathbb{R} \to \mathbb{R}^2$, definita da

$$f(\boldsymbol{x}) = \begin{pmatrix} x \\ 2x \end{pmatrix}, \quad \forall x \in \mathbb{R}.$$

  Essa è iniettiva, dato che $\begin{pmatrix} x \\ 2x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ può aversi solo con $x = 0$. Essa non è suriettiva, dato che ad esempio il vettore $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ non è immagine di nessun $\boldsymbol{x}$.

- Si consideri ora la trasformazione lineare $f : \mathbb{R}^2 \to \mathbb{R}$, definita da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = x, \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

  Essa non è iniettiva, dato che ad esempio i vettori $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ hanno la stessa immagine. Essa è invece suriettiva, dato che ogni $x \in \mathbb{R}$ è immagine, ad esempio, del vettore $\begin{pmatrix} x \\ 0 \end{pmatrix}$.

Ci sono anche esempi di trasformazioni che non sono né iniettive né suriettive:

- consideriamo la $f : \mathbb{R}^2 \to \mathbb{R}^2$, definita da

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

  Essa non è iniettiva, dato che ad esempio i vettori $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ hanno la stessa immagine. Essa non è nemmeno suriettiva, dato che ad esempio il vettore $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ non è immagine di nessun vettore.

Per capire meglio come opera una trasformazione invertibile, enuncio questo risultato generale:

Proposizione Una trasformazione f è invertibile se e solo se trasforma basi in basi, cioè se e solo se, data una qualunque base di $\mathbb{R}^n$ $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n\}$, allora $\{f(\boldsymbol{v}^1), f(\boldsymbol{v}^2), \ldots, f(\boldsymbol{v}^n)\}$ è anch'essa una base di $\mathbb{R}^n$.

Osservazione Come caso particolare di quanto appena detto abbiamo che f è invertibile se e solo se i vettori $f(\boldsymbol{u}^1), f(\boldsymbol{u}^2), \ldots, f(\boldsymbol{u}^n)$ sono una base di $\mathbb{R}^n$.

Visti i risultati sulle trasformazioni invertibili, possiamo dare la seguente

Definizione Se A è la matrice di rappresentazione della trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^n$ e se f è invertibile, allora si chiama **matrice inversa** di A la matrice di rappresentazione di f^{-1}.

Questo equivale a chiamare **invertibile** una matrice $n \times n$ A se esiste un'altra matrice $n \times n$ B tale che valga

$$AB = BA = I_n,$$

dove I_n è la matrice identità delle matrici $n \times n$.

Osservazione Per quanto visto in precedenza, possiamo dire che una matrice A è invertibile se e solo se le sue colonne sono linearmente indipendenti.

Esempio La matrice

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

è invertibile, dato che le sue colonne sono i vettori fondamentali di $\mathbb{R}^2$. Si trova facilmente che la matrice inversa è la matrice stessa.

Si intuisce però che con matrici di ordine 3 o più non è così facile trovare la matrice inversa (e nemmeno capire se la matrice è invertibile). Serve allora un procedimento di calcolo che, data una qualunque matrice quadrata, mi dica intanto se essa è invertibile e, in caso affermativo, mi permetta di trovare la matrice inversa. Arriveremo a tale procedimento nella prossima dispensa.

5 Soluzioni degli esercizi

Esercizio 1.1

(a) La matrice di rappresentazione di f è

$$\begin{pmatrix} 2 & -3 & 5 \\ 0 & 1 & -1 \end{pmatrix}, \quad \text{dato che} \quad f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & -3 & 5 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

La trasformazione f è definita in $\mathbb{R}^3$ a valori in $\mathbb{R}^2$, cioè possiamo scrivere $f : \mathbb{R}^3 \to \mathbb{R}^2$.

(b) La matrice di rappresentazione di f è

$$\begin{pmatrix} 1 & -1 \\ 3 & 2 \\ 1 & 0 \end{pmatrix}, \quad \text{dato che} \quad f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

La trasformazione f è definita in $\mathbb{R}^2$ a valori in $\mathbb{R}^3$, cioè possiamo scrivere $f : \mathbb{R}^2 \to \mathbb{R}^3$.

(c) La matrice di rappresentazione di f è

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{dato che} \quad f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

La trasformazione f è definita in $\mathbb{R}^3$ a valori in $\mathbb{R}^3$, cioè possiamo scrivere $f : \mathbb{R}^3 \to \mathbb{R}^3$.

Esercizio 1.2

(a) La trasformazione rappresentata da A è una funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ la cui espressione analitica è

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix}.$$

(b) La trasformazione rappresentata da A è una funzione $f : \mathbb{R}^3 \to \mathbb{R}^3$ la cui espressione analitica è

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_3 \\ x_2 + x_3 \\ -x_1 + x_2 \end{pmatrix}.$$

(c) La trasformazione rappresentata da A è una funzione $f : \mathbb{R} \to \mathbb{R}^3$ la cui espressione analitica è

$$f(x) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} x = \begin{pmatrix} x \\ 2x \\ 3x \end{pmatrix}.$$

Esercizio 1.3

(a) La trasformazione è da $\mathbb{R}^2$ a $\mathbb{R}^2$. Si può osservare che, banalmente, il trasformato del vettore nullo non è il vettore nullo. Infatti $f\binom{0}{0} = \binom{1}{-1}$. Lo studente provi a trovare un controesempio all'additività e all'omogeneità.

(b) La trasformazione è da $\mathbb{R}^2$ a $\mathbb{R}^2$. Si può osservare che questa volta il trasformato del vettore nullo è il vettore nullo. Però la trasformazione non è additiva. Infatti con $\boldsymbol{x} = \binom{1}{0}$ e $\boldsymbol{y} = \binom{0}{1}$ si ha che $f(\boldsymbol{x}+\boldsymbol{y}) = f\binom{1}{1} = \binom{1}{2}$, mentre $f(\boldsymbol{x}) + f(\boldsymbol{y}) = \binom{0}{1} + \binom{0}{1} = \binom{0}{2}$.

Lo studente provi a trovare un controesempio all'omogeneità.

(c) Anche in questo caso il trasformato del vettore nullo è lo zero (qui il numero reale zero). Però la trasformazione non è omogenea. Infatti con $\boldsymbol{x} = \binom{1}{1}$ si ha che $f(2\boldsymbol{x}) = f\binom{2}{2} = \frac{4}{9}$, mentre $2f(\boldsymbol{x}) = 2 \cdot \frac{1}{3} = \frac{2}{3}$.

Lo studente provi a trovare un controesempio all'additività.

Esercizio 2.1

Le matrici trasposte sono

$$A^T = \begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix} \qquad B^T = \begin{pmatrix} 0 & -1 \\ 1 & 1 \\ -2 & 0 \end{pmatrix} \qquad C^T = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \qquad D^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ -1 & 0 & 1 \end{pmatrix}.$$

I prodotti che si possono fare sono i seguenti:

$$AB = \begin{pmatrix} 1 & 2 & -6 \\ -2 & 3 & -2 \end{pmatrix} \qquad BC = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \qquad BD = \begin{pmatrix} -2 & 4 & -2 \\ -1 & 2 & 1 \end{pmatrix}.$$

Sono anche possibili, scambiando l'ordine, i prodotti DC, CB e CA. Lo studente provi a calcolarli.

Esercizio 2.2

Si potrebbe rispondere alla domanda operando soltanto sulle espressioni analitiche delle due trasformazioni, ma forse conviene utilizzare le rispettive matrici di rappresentazione. Troviamo intanto allora queste matrici, che indichiamo con A_f e con A_g. Si ha

$$A_f = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -2 \end{pmatrix} \qquad \text{e} \qquad A_g = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Per quanto visto nella lezione, la trasformazione composta $g \circ f$ è rappresentata dalla matrice prodotto (righe per colonne) di A_g per A_f, cioè da

$$A_g A_f = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}.$$

Pertanto l'espressione analitica di $g \circ f : \mathbb{R}^2 \to \mathbb{R}^2$ è

$$g \circ f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_2 \\ 2x_2 \end{pmatrix}.$$

La trasformazione composta $f \circ g$ è invece rappresentata dalla matrice prodotto di A_f per A_g, cioè da

$$A_f A_g = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ -2 & 3 & 1 \end{pmatrix}.$$

Quindi l'espressione analitica di $f \circ g : \mathbb{R}^3 \to \mathbb{R}^3$ è

$$f \circ g \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ -2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_3 \\ -x_1 + 2x_2 + x_3 \\ -2x_1 + 3x_2 + x_3 \end{pmatrix}.$$