Geometria e Algebra (II), 17.12.12

Funzioni lineari

1. Cosi' come il piano vettoriale geometrico $\mathfrak{P}_O$ e lo spazio vettoriale geometrico $\mathfrak{S}_O$ si possono identificare, tramite la scelta di un sistema di riferimento, con lo spazio vettoriale $\mathbb{R}^2$ e lo spazio vettoriale $\mathbb{R}^3$, allo stesso modo uno spazio vettoriale V di dimensione n si puo' identificare, tramite la scelta di una base, con lo spazio vettoriale $\mathbb{R}^n$. Precisamente, scelta una base $\mathfrak{V} = \{v_1, \ldots, v_n\}$ di V, ogni vettore $x \in V$ si puo' scrivere in uno ed un solo modo come combinazione lineare

   $$x = v_1x_1 + \ldots + v_nx_n$$

   dei vettori $v_1, \ldots, v_n$; gli scalari $x_1, \ldots, x_n$ sono le coordinate di x rispetto a $\mathfrak{V}$; possiamo associare ad x il vettore

   $$[x]_{\mathfrak{V}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

   Abbiamo cosi' una funzione "coordinata rispetto a $\mathfrak{V}$"

   $$[\ ]_{\mathfrak{V}} : V \to \mathbb{R}^n; \quad [\ ]_{\mathfrak{V}} : x \mapsto [x]_{\mathfrak{V}}.$$

   Questa funzione e' biunivoca, ed e' compatibile con le operazioni di somma e prodotto per scalari in V e $\mathbb{R}^n$:

   $$[x + y]_{\mathfrak{v}} = [x]_{\mathfrak{v}} + [y]_{\mathfrak{v}}$$

   $$[x\alpha]_{\mathfrak{v}} = [x]_{\mathfrak{v}}\alpha$$

   per ogni $x, y \in V$ ed ogni $\alpha \in \mathbb{R}$. (Lo si verifichi per esercizio).

2. **Funzioni lineari**

   Definizione 1 *Una funzione $f : V \to W$ fra due spazi vettoriali V e W su $\mathbb{R}$ si dice funzione lineare se e' compatibile con le operazioni di somma in V e W :*

   $$f(x + y) = f(x) + f(y), \qquad \forall x, y \in V,$$

   ed e' compatibile con le operazioni di prodotto per scalari in V e W :

   $$f(x\alpha) = f(x)\alpha, \qquad \forall x \in V, \alpha \in \mathbb{R}.$$

   Il termine "funzione lineare" ha vari sinomimi, fra i quali "omomorfismo;" in questa linea, a "funzione lineare biiettiva" corrisponde "isomorfismo". Dunque la funzione coordinata rispetto a una base $[\ ]_{\mathfrak{V}} : V \to \mathbb{R}^n$ da uno spazio vettoriale $n-$dimensionale V ad $\mathbb{R}^n$ e' un isomorfismo.

   Osserviamo che se $f : V \to W$ e' lineare, allora:

   f manda il vettore nullo di V nel vettore nullo di W, coe' $f(0_v) = 0_w$; infatti

$$f(0_v) = f(0_v\, 0) = f(0_v)\, 0 = 0_w;$$

f manda ciascuna combinazione lineare di vettori di V nella combinazione lineare dei corrispondenti vettori di W, con gli stessi coefficienti:

$$f\left(a_1 z_1 + \cdots + a_m z_m\right) = f\left(a_1\right) z_1 + \cdots + f\left(a_m\right) z_m, \qquad \forall a_i \in V,\ z_i \in \mathbb{R}.$$

3. La moltiplicazione di una matrice fissata $A \in \mathbb{R}^{m\times n}$ per vettori variabili in $\mathbb{R}^n$ e' una funzione $f_A = f$ da $\mathbb{R}^n$ verso $\mathbb{R}^m$:

$$f : \mathbb{R}^n \to \mathbb{R}^m, \qquad f : x \mapsto Ax,\ x \in \mathbb{R}^n;$$

per le proprieta' della moltiplicazione di matrici rispetto alla somma ed al prodotto per scalari, una tale funzione e' lineare:

$$f(x+y) = A(x+y) = Ax + Ay = f(x) + f(y),$$

$$f(x\alpha) = A(x\alpha) = (Ax)\alpha = f(x)\alpha,$$

per ogni $x, y \in R^n$ ed ogni $\alpha \in \mathbb{R}$. Piu' in dettaglio, ad una matrice

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

corrisponde la funzione lineare $f_A = f : \mathbb{R}^n \to \mathbb{R}^m$ definita da

$$f(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Si noti che ciascuna delle m componenti del vettore cosi' ottenuto e' un polinomio nelle n variabili $x_1, \dots, x_n$ che e' omogeneo di primo grado, eventualmente ridotto al polinomio nullo. Cosi'

per $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$ si ha la funzione $f : \mathbb{R}^2 \to \mathbb{R}^3,$ definita da

$$f(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \\ a_{31}x_1 + a_{32}x_2 \end{bmatrix}.$$

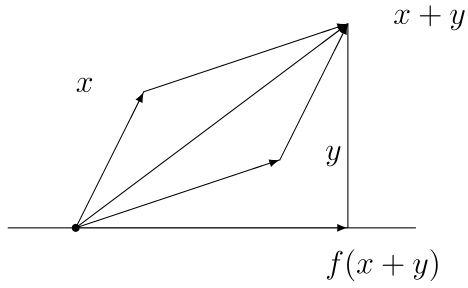
Ad esempio,

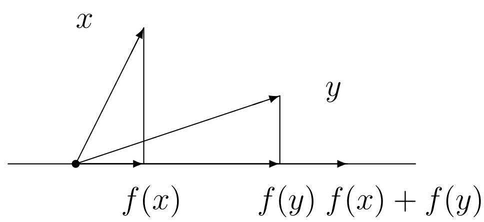
per $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ si ha $f : \mathbb{R}^2 \to \mathbb{R}^3, \quad f(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{bmatrix}.$

4. Nel piano vettoriale geometrico $\mathfrak{P}_O$ consideriamo una retta l passante per O, e la funzione

$$f = \mathrm{pr}_l : \mathfrak{P}_O \to \mathfrak{P}_O$$

che ad ogni vettore $x \in \mathfrak{P}_O$ associa il vettore $f(x) = \mathrm{pr}_l(x)$ proiezione ortogonale di x sulla retta l. Questa funzione e' lineare; per dare il senso di questa affermazione riportiamo di seguito le costruzioni di $f(x+y)$ e di $f(x)+f(y)$. Lasciamo per esercizio di descrivere le costruzioni di $f(x\alpha)$ e $f(x)\alpha$.





Identifichiamo il piano vettoriale geometrico $\mathfrak{P}_O$, mediante un sistema di riferimento cartesiano ortogonale monometrico, con il piano vettoriale $\mathbb{R}^2$. Rappresentata la retta l come la retta $l = \langle a \rangle$ generata da un vettore non nullo a, si ha

$$\mathrm{pr}_l(x) = a\frac{a \cdot x}{a \cdot a}.$$

Possiamo verificare la linearita' algebricamente:

$$\begin{aligned}\mathrm{pr}_l(x+y) = a\frac{a\cdot(x+y)}{a\cdot a} &= a\frac{a\cdot x + a\cdot y}{a\cdot a}\\ &= a\left(\frac{a\cdot x}{a\cdot a} + \frac{a\cdot y}{a\cdot a}\right) = a\frac{a\cdot x}{a\cdot a} + a\frac{a\cdot y}{a\cdot a} = \mathrm{pr}_l(x) + \mathrm{pr}_l(y);\end{aligned}$$

$$\mathrm{pr}_l(x\ \alpha) = a\frac{a\cdot(x\ \alpha)}{a\cdot a} = a\frac{(a\cdot x)\ \alpha}{a\cdot a} = a\frac{a\cdot x}{a\cdot a}\ \alpha = \mathrm{pr}_l(x)\ \alpha,$$

per ogni $x, y \in \mathfrak{P}_O$, ed ogni $\alpha \in \mathbb{R}$.

Consideriamo ora il caso specifico in cui $a = \left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$; si ha

$$\mathrm{pr}_l(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) \;=\; \begin{bmatrix} 2 \\ 1 \end{bmatrix} \frac{\begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}{\begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}} \;=\; \begin{bmatrix} 2 \\ 1 \end{bmatrix} \frac{2x_1 + x_2}{5} \;=\; \begin{bmatrix} \frac{4x_1+2x_2}{5} \\ \frac{2x_1+x_2}{5} \end{bmatrix};$$

Osserviamo che questa funzione e' la funzione lineare associata ad una matrice:

$$\mathrm{pr}_l(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} \frac{4x_1+2x_2}{5} \\ \frac{2x_1+x_2}{5} \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

5. Consideriamo la funzione

   $$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = 1, \ \forall x \in \mathbb{R}.$$

   Ci chiediamo se e' lineare. Osserviamo che l'uguaglianza $f(x + y) = f(x) + f(y)$ non e' sempre soddisfatta, anzi non e' mai soddisfatta, in quanto al primo membro si ha $f(x+y) = 1$, mentre al secondo membro si ha $f(x)+f(y) = 2$. Dunque questa funzione non e' lineare. Alla stessa conclusione si poteva giungere usando la prima osservazione sulle funzioni lineari (cfr. punto 2) ed osservando che $f(0) = 1 \neq 0$. In generale, si ha che una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ che assume un valore costante $f(x) = c$ e' lineare se e solo se $c = 0_m$.

   Consideriamo la funzione norma

   $$\| \ \| : \mathbb{R}^2 \to \mathbb{R}, \quad \|x\| = \sqrt{x_1^2 + x_2^2}, \ \forall x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2.$$

   Ci chiediamo se e' lineare. Osserviamo che l'uguaglianza $\|x+y\| = \|x\|+\|y\|$ non e' sempre soddisfatta, in quanto per $x = e_1$ ed $y = e_2$ al primo membro si ha $\|e_1 + e_2\| = \sqrt{2}$, mentre al secondo membro si ha $\|e_1\| + \|e_2\| = 2$. Dunque questa funzione non e' lineare. Allo stesso modo si verifica che la funzione norma

   $$\| \ \| : \mathbb{R}^n \to \mathbb{R}, \quad \|x\| = \sqrt{\textstyle\sum_1^n x_i^2} \ \ \forall x = [x_i]_1^n \in \mathbb{R}^n$$

   non e' lineare.

6. Abbiamo visto come si possa costruire una funzione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$ a partire da una qualsiasi matrice $A \in \mathbb{R}^{m \times n}$, ponendo $f(x) = Ax$, per ogni $x \in \mathbb{R}^n$. Ci chiediamo ora se in questo modo si ottengono tutte le funzioni lineari $\mathbb{R}^n \to \mathbb{R}^m$, piu' in generale ci chiediamo come sono fatte le funzioni lineari $\mathbb{R}^n \to \mathbb{R}^m$.

   Consideriamo il caso piu' semplice. Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione lineare. Osserviamo che $f(x) = f(1\,x) = f(1)\,x$, per ogni $x \in \mathbb{R}$; dunque si ha

   $$f(x) = mx, \quad x \in \mathbb{R},$$

   dove $m = f(1)$.

   Sia ora $f : \mathbb{R}^n \to \mathbb{R}^m$ una funzione lineare. Osserviamo che, esprimendo il generico vettore $x \in \mathbb{R}^n$ come combinazione lineare dei vettori della base canonica di $\mathbb{R}^n$

   $$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = e_1x_1 + \cdots e_nx_n,$$

   per la linearita' di f si ha

$$f(x) = f(e_1x_1 + \cdots + e_nx_n) = f(e_1)x_1 + \cdots + f(e_n)x_n;$$

ora

$$f(e_1)x_1 + \cdots + f(e_n)x_n = [f(e_1),\ \ldots,\ f(e_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix};$$

dunque si ha

$$f(x) = M(f)\, x, \quad x \in \mathbb{R}^n,$$

dove

$$M(f) = [f(e_1),\ \ldots,\ f(e_n)]\,.$$

Diciamo che $M(f)$ e' la matrice che rappresenta f.

7. **Esempio** Consideriamo di nuovo la funzione lineare $f = \mathrm{pr}_l : \mathbb{R}^2 \to \mathbb{R}^2$ proiezione ortogonale sulla retta $l = \langle a \rangle$ generata dal vettore $a = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$, gia' considerata al punto 4. Per quanto appena visto si ha

$$\mathrm{pr}_l(x) = M(f)\, x \quad x \in \mathbb{R}^2,$$

dove

$$M(f) = [f(e_1),\ f(e_2)]\,.$$

Ora,

$$f(e_1) = a\frac{a \cdot e_1}{a \cdot a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \frac{2}{5} = \begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \end{bmatrix};$$

$$f(e_2) = a\frac{a \cdot e_2}{a \cdot a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \frac{1}{5} = \begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \end{bmatrix},$$

$$M(f) = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}.$$

8. Ci chiediamo ora come sono fatte le funzioni lineari

$$f : V \to W,$$

dove V e' uno spazio vettoriale di dimensione n e W e' uno spazio vettoriale di dimensione m. Siano

$$\mathfrak{V} = \{v_1, \ldots, v_n\} \text{ una base di } V$$

$$\mathfrak{W} = \{w_1, \ldots, w_m\} \text{ una base di } W.$$

Sia $x \in V$. Poniamo

$$x = v_1x_1 + \cdots + v_nx_n, \text{ cioe' } \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x]_{\mathfrak{v}}\,.$$

Per la linearita' di f si ha

$$f(x) = f(v_1x_1 + \cdots + v_nx_n) = f(v_1)x_1 + \cdots + f(v_n)x_n.$$

Applicando ad ambo i membri la funzione coordinata rispetto a $\mathfrak{W}$ si ha

$$\begin{aligned}[f(x)]_{\mathfrak{w}} &= [f(v_1)x_1 + \cdots + f(v_n)x_n]_{\mathfrak{w}} \\ &= [f(v_1)]_{\mathfrak{w}}\, x_1 + \cdots + [f(v_n)]_{\mathfrak{w}}\, x_n.\end{aligned}$$

Ora

$$\begin{aligned}[f(v_1)]_{\mathfrak{w}}\, x_1 + \cdots + [f(v_n)]_{\mathfrak{w}}\, x_n &= \left[[f(v_1)]_{\mathfrak{w}}\,,\ \ldots,\ [f(v_1)]_{\mathfrak{w}}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \left[[f(v_1)]_{\mathfrak{w}}\,,\ \ldots,\ [f(v_1)]_{\mathfrak{w}}\right] [x]_{\mathfrak{v}}\,.\end{aligned}$$

In definitiva si ha

$$[f(x)]_{\mathfrak{w}} = M^{\mathfrak{v}}_{\mathfrak{w}}(f)\ [x]_{\mathfrak{v}}\,, \quad x \in V,$$

dove

$$M^{\mathfrak{v}}_{\mathfrak{w}}(f) = \left[[f(v_1)]_{\mathfrak{w}}\,,\ \ldots,\ [f(v_1)]_{\mathfrak{w}}\right].$$

Diciamo che $M^{\mathfrak{v}}_{\mathfrak{w}}(f)$ e' la matrice che rappresenta la funzione lineare $f : V \to W$ rispetto alle basi $\mathfrak{V}$ di V e $\mathfrak{W}$ di W.

9. **Esempio** Consideriamo di nuovo la funzione lineare $f = \mathrm{pr}_l : \mathfrak{P}_O \to \mathfrak{P}_O$ proiezione ortogonale su una retta l passante per O, gia' considerata all'inizio del punto 4. Sia $v_1 \neq 0$ un vettore non nullo sulla retta l, e sia $v_2 \neq 0$ un vettore non nullo ortogonale ad l; questi due vettori formano una base $\mathfrak{V} = \{v_1, v_2\}$ di $\mathfrak{P}_O$.

   Per quanto appena visto si ha

   $$[f(x)]_{\mathfrak{v}} = M^{\mathfrak{v}}_{\mathfrak{v}}(f)\ [x]_{\mathfrak{v}}\,, \quad x \in \mathfrak{P}_O,$$

   dove

   $$M^{\mathfrak{v}}_{\mathfrak{v}}(f) = \left[[f(v_1)]_{\mathfrak{v}}\,,\ [f(v_2)]_{\mathfrak{v}}\right].$$

   Ora

   $$f(v_1) = v_1 = v_1\, 1 + v_2\, 0,\ \text{cioe'}\ [f(v_1)]_{\mathfrak{v}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

   $$f(v_2) = 0\ \ = v_1\, 0 + v_2\, 0,\ \text{cioe'}\ [f(v_2)]_{\mathfrak{v}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   $$M^{\mathfrak{v}}_{\mathfrak{v}}(f) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$