Lezione del 05.04

Questa lezione si riferisce al Cap.5 “Applicazioni lineari”, Par.5.1 “Definizione di applicazione lineare”, Par.5.2 “Applicazioni lineari e matrici”, Par.5.3 “La composizione di applicazioni lineari”. Si sono considerate solo applicazioni lineari da uno spazio vettoriale $\mathbb{R}^n$ ad uno $\mathbb{R}^m$ (in particolare non si è trattato il Th.5.17). Il tema principale è l'introduzione delle applicazione lineari fra spazi vettoriali come generalizzazioni delle applicazioni $\mathbb{R} \to \mathbb{R}$, $x \mapsto mx$ (m costante in $\mathbb{R}$), la rappresentazione di applicazioni lineari $\mathbb{R}^n \to \mathbb{R}^m$ con matrici $m \times n$ e la rappresentazione della composizione di applicazioni lineari con il prodotto di matrici.

Applicazioni lineari. Si sono ricordate le prime nozioni sulle funzioni fra insiemi: funzione, dominio e codominio di una funzione, immagine di un elemento tramite una funzione o valore di una funzione su un elemento; si è segnalato che oltre al termine “funzione” si usano sinomimi come “applicazione”. Si sono usati anche termini come “variabile” e locuzioni come “la funzione che in entrata prende una variabile ... e in uscita restituisce ...”

Si è considerato il tipo di relazione più semplice che può sussistere fra due variabili reali, descritta da una funzione del tipo

$$\ell : \mathbb{R} \to \mathbb{R}, \quad \ell(x) = mx$$

dove m è una costante in $\mathbb{R}$, che ha per grafico una retta per l'origine con una certa pendenza m. Si è osservato che una tale funzione si comporta bene rispetto alla struttura di spazio vettoriale di $\mathbb{R}$: si comporta bene rispetto alla somma: $\ell(u + v) = m(u+v) = mu+mv = \ell(u)+\ell(v)$ (per ogni u, v in $\mathbb{R}$); si comporta bene rispetto alla moltiplicazione per scalari: $\ell(\alpha u) = m(\alpha u) = \alpha(mu) = \alpha\ell(u)$ (per ogni α, u in $\mathbb{R}$).

Si è data la definizione di applicazione lineare fra due spazi vettoriali

Definizione [Def.5.1.2] Un'applicazione $F : V \to W$ fra due spazi vettoriali V, W si dice applicazione lineare se

(1) $$F(\underline{u} + \underline{v}) = F(\underline{u}) + F(\underline{v}) \qquad \forall \underline{u}, \underline{v} \in V;$$

(2) $$F(\alpha\underline{u}) = \alpha F(\underline{u}) \qquad \forall \alpha \in \mathbb{R},\ \underline{u} \in V.$$

Ci si è limitati alle applicazioni lineari da uno spazio vettoriale $\mathbb{R}^n$ ad uno $\mathbb{R}^m$. Si sono considerati i seguenti esempi e controesempi

(1) ogni funzione $\ell : \mathbb{R} \to \mathbb{R}$, $\ell(x) = mx$ (m costante in $\mathbb{R}$) è lineare;

(1') la funzione $Q : \mathbb{R} \to \mathbb{R}$, $Q(x) = x^2$ non è lineare. Infatti non è vero che $Q(u+v) = (u + v)^2$ sia uguale a $Q(u) + Q(v) = u^2 + v^2$ per ogni $u, v \in \mathbb{R}$.

(1") la funzione $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x + 1$ non è lineare. Infatti ...

(2) La funzione $L : \mathbb{R}^2 \to \mathbb{R}$, $L(x_1, x_2) = 2x_1 + 3x_2$ è lineare. Infatti:

$$L((u_1, u_2) + (v_1, v_2)) = L(u_1 + v_1, u_2 + v_2) = 2(u_1 + v_1) + 3(u_2 + v_2)$$

è uguale a

$$L(u_1, u_2) + L(v_1, v_2) = 2u_1 + 3u_2 + 2v_1 + 3v_2$$

per ogni $(u_1, u_2), (v_1, v_2) \in \mathbb{R}^2$;

$$L(\alpha(u_1, u_2)) = L(\alpha u_1, \alpha u_2) = 2(\alpha u_1) + 3(\alpha u_2)$$

è uguale a

$$\alpha L(u_1, u_2) = \alpha(2u_1 + 3u_2)$$

per ogni $\alpha \in \mathbb{R}$ ed ogni $(u_1, u_2) \in \mathbb{R}^2$.

(2') più in generale, una qualsiasi funzione polinomiale omogenea di primo grado (eventualmente degenere nel polinomio nullo) in due variabili

$$L : \mathbb{R}^2 \to \mathbb{R}, \ L(x_1, x_2) = ax_1 + bx_2,$$

(a, b costanti in $\mathbb{R}$) è un'applicazione lineare;

(3) Si è detto che viene da pensare che una qualsiasi terna di funzioni polinomiali omogenee di primo grado (eventualmente degeneri ...) in due variabili

$$L : \mathbb{R}^2 \to \mathbb{R}^3, \ L(x_1, x_2) = (ax_1 + bx_2, cx_1 + dx_2, ex_1 + fx_2),$$

($a, \dots, f$ costanti in $\mathbb{R}$) sia un'applicazione lineare; si è osservato che, scrivendo i vettori come vettori colonna si è condotti ad osservare che

$$L\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \\ ex_1 + fx_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

cioè L agisce su un vettore 2×1 come la moltiplicazione a sinistra per una matrice 3×2, creando dunque un vettore 3×1; ciascuna riga della matrice consiste dei coefficienti del corrispondente polinomio.

Applicazioni lineari e matrici. Si è considerata una qualsiasi $m-$pla di funzioni polinomiali omogenee di primo grado (eventualmente degeneri ...) in n variabili

$$L : \mathbb{R}^n \to \mathbb{R}^m, \quad L(x_1, \dots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \dots, a_{m1}x_1 + \cdots + a_{mn}x_n)$$

($a_{11}, \dots, a_{mn}$ costanti in $\mathbb{R}$); si è osservato che, scrivendo i vettori come vettori colonna si è condotti ad osservare che

$$L\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

e, usando il formalismo matriciale come

$$L : \mathbb{R}^n \to \mathbb{R}^m, \quad L(\underline{x}) = A\underline{x};$$

cioè L agisce su un vettore $n \times 1$ come la moltiplicazione a sinistra per una matrice A di tipo $m \times n$, creando dunque un vettore $m \times 1$; ciascuna delle m righe della matrice A consiste degli n coefficienti del corrispondente polinomio.

Questa applicazione è lineare, in quanto

$$L(\underline{u} + \underline{v}) = A(\underline{u} + \underline{v}) = A\underline{u} + A\underline{v} = L(\underline{u}) + L(\underline{v});$$

$$L(\alpha\underline{u}) = A(\alpha\underline{u}) = \alpha A\underline{u} = \alpha L(\underline{u});$$

(nei passaggi centrali si sono usate le proprietà del prodotto di matrici rispetto alla somma (proprietà distributiva) e rispetto al prodotto per scalari).

Ci si è chiesti se tutte le applicazioni lineari $\mathbb{R}^n \to \mathbb{R}^m$ sono di questo tipo. A tale scopo, si sono considerati prima alcuni casi.

(1) Sia $L : \mathbb{R} \to \mathbb{R}$ un'applicazione lineare. Allora per ogni $x \in \mathbb{R}$ si ha $L(x) = L(x1) = xL(1)$; cioè L è il polinomio omogeneo di primo grado (eventualmente degenere ...) in x con coefficiente il valore $L(1)$ di L su 1.

(2) Sia $L : \mathbb{R}^2 \to \mathbb{R}$ un'applicazione lineare. Allora per ogni $(x_1, x_2) \in \mathbb{R}^2$ si ha

$$L(x_1, x_2) = L(x_1(1,0) + x_2(0,1)) = x_1L(1,0) + x_2L(0,1);$$

cioè L è il polinomio omogeneo di primo grado (eventualmente degenere ...) in x_1, x_2 con coefficienti rispettivi i valori $L(1,0)$ ed $L(0,1)$ di L sui vettori $(1,0)$ e $(0,1)$ della base canonica di $\mathbb{R}^2$.

In generale. Sia $L : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora per ogni $(x_1, \ldots, x_n) \in \mathbb{R}^n$ si ha

$$L(x_1, \ldots, x_n) = L(x_1\underline{e}_1 + \cdots + x_n\underline{e}_n) = x_1L(\underline{e}_1) + \cdots + x_nL(\underline{e}_n) = (*);$$

posto

$$L(\underline{e}_1) = (a_{11}, \ldots, a_{m1}), \ldots, L(\underline{e}_n) = (a_{1n}, \ldots, a_{mn}),$$

si ha

$$\begin{aligned}(*) &= x_1(a_{11}, \ldots, a_{m1}) + \cdots + x_n(a_{1n}, \ldots, a_{mn}) \\ &= (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n);\end{aligned}$$

cioè L è una $m-$pla di polinomi omogenei di primo grado (eventualmente ...) in $x_1, \ldots, x_n$; la $m-$pla dei coefficienti di x_i nei vari polinomi è il valore $L(\underline{e}_i)$ di L sull'$i-$mo vettore $\underline{e}_i$ della base canonica di $\mathbb{R}^n$ (per ogni $i = 1, \ldots, n$). Riassumendo, si è provato il

Teorema[Th.5.2.2] Per ogni applicazione $L : \mathbb{R}^n \to \mathbb{R}^m$ le seguenti condizioni son equivalenti

(1) L è lineare;

(2) Esistono $m \times n$ costanti $a_{11}, \ldots, a_{mn}$ in $\mathbb{R}$ tali che

$L(x_1, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n) \qquad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n;$

(3) Esiste una matrice A $m \times n$ tale che, identificati i vettori con vettori colonna,

$$L(\underline{x}) = A\underline{x} \qquad \forall \underline{x} \in \mathbb{R}^n.$$

Inoltre: nelle varie righe di A compaiono i coefficienti di $x_1, \ldots, x_n$ nei vari polinomi; le varie colonne di A son le immagini dei vari vettori della base canonica $\underline{e}_1, \ldots, \underline{e}_n$ di $\mathbb{R}^n$.

Per ogni matrice A $m \times n$ si è denotata con L_A la corrispondente applicazione lineare

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \; L_A(\underline{x}) = A\underline{x}.$$

Composizione di applicazioni lineari e prodotto di matrici. Si è ricordata la nozione di composizione di funzioni fra insiemi, nella forma seguente. Data una prima funzione $f : A \to B$ ed una seconda funzione $g : B \to C$ (A, B, C insiemi) tali che il codominio di f coincida col dominio di g, si dice funzione composta "g dopo f" denotata $g \circ f$ la funzione

$$g \circ f : A \to C, \quad (g \circ f)(x) = g(f(a), \quad \forall a \in A.$$

Se il codominio di f non coincide col dominio di g allora $g \circ f$ non viene definita. Si sono ricordate la proprietà di di associatività della composizione di funzioni, e le proprietà che caratterizzano le funzioni identità. Per ogni insieme A si definisce la funzione id_A identità su A come la funzione da A in se stesso che manda ogni elemento in se stesso

$$\mathrm{id}_A : A \to A, \quad \mathrm{id}_A(x) = x, \ \forall x \in A.$$

Le funzioni identità soon caratterizzate dalle proprietà

$$\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A, \quad \forall f : A \to B.$$

Si è verificato che la composizione di due applicazioni lineari è un'applicazione lineare:

Proposizione[Prop.5.3.3] Se due applicazioni fra spazi vettoriali $F : U \to V$ e $G : V \to W$ sono lineari, allora anche l'applicazione composta $G \circ F : V \to W$ è lineare.

Infatti: per ogni $\underline{u}, \underline{v} \in V$ si ha

$$(G \circ F)(\underline{u} + \underline{v}) = G(F(\underline{u} + \underline{v})) = G(F(\underline{u}) + F(\underline{v})) = \\ G(F(\underline{u})) + G(F(\underline{v})) = (G \circ F)(\underline{u}) + (G \circ F)(\underline{v});$$

per ogni $\alpha \in \mathbb{R}$ e $\underline{u} \in V$ si ha

$$(G \circ F)(\alpha\underline{u}) = G(F(\alpha\underline{u}) = G(\alpha F(\underline{u}) = \alpha G(F(\underline{u}) = \alpha(G \circ F)(\underline{u}).$$

Si è provato che nella corrispondenza fra matrici e applicazioni lineari, al prodotto di matrici corrisponde la composizione di applicazioni lineari:

Proposizione[Cor.5.3.4] Comunque siano date due applicazioni lineari L_B ed L_A associate a due matrici B e A, esiste l'applicazione composta $L_B \circ L_A$ se e solo se esiste la matrice prodotto BA e in tal caso l'applicazione lineare composta $L_B \circ L_A$ è l'applicazione lineare associata alla matrice prodotto BA :

$$L_B \circ L_A = L_{BA}.$$

Infatti, se B ha tipo $q \times p$ e A ha tipo $m \times n$ allora $L_B : \mathbb{R}^p \to \mathbb{R}^q$ e $L_A : \mathbb{R}^n \to \mathbb{R}^m$ e sia il prodotto BA che l'applicazione $L_B \circ L_A$ esistono se e solo se $p = m$. In questo caso, per ogni $\underline{x} \in \mathbb{R}^n$ si ha

$$(L_B \circ L_A)(\underline{x}) = L_B(L_A(\underline{x}) = B(A\underline{x}) = (BA)\underline{x} = L_{BA}(\underline{x})$$

(nel penultimo passaggio si è usata l'associatività del prodotto di matrici).

Applicazione La corrispondenza fra applicazioni lineari e matrici si può usare per semplificare il calcolo della composizione di due applicazioni lineari. Ad esempio, per

$$F : \mathbb{R}^3 \to \mathbb{R}^2, \ F(x, y, z) = (x - 2y + 3z, 4y - 5z);$$
$$G : \mathbb{R}^2 \to \mathbb{R}^4, \ G(x, y) = (2x - 3y, 3x - 2y, 4x, 5y),$$

la funzione composta $G \circ F : \mathbb{R}^3 \to \mathbb{R}^4$ è data da

$$\begin{aligned}&(G \circ F)(x, y, z) = G(F(x, y, z)) = G(x - 2y + 3z, 4y - 5z)\\ &= (2(x-2y+3z) - 3(4y-5z), 3(x-2y+3z) - 2(4y-5z), 4(x-2y+3z), 5(4y-5z))\\ &\qquad = \cdots\end{aligned}$$

D'altro canto,

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 3 & -2 \\ 4 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

dunque

$$\begin{aligned}(G \circ F)\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 2 & -3 \\ 3 & -2 \\ 4 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\\ &= \begin{pmatrix} 2 & -10 & 21 \\ 3 & -14 & 19 \\ 4 & -8 & 12 \\ 0 & 20 & -25 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\end{aligned}$$

così

$$(G \circ F)(x, y, z) = (2x - 10y + 21z, 3x - 14y + 19z, 4x - 8y + 12z, 20y - 25z).$$