PROPRIETA' GENERALI DELLE EQUAZIONI LINEARI

In questa sezione verranno presentate alcune importanti proprietà delle equazioni, dette **LINEARI**,

$$A(u) = f \qquad (*)$$

ove A è un'applicazione lineare da uno spazio vettoriale X ad uno spazio, in genere differente, Y, ed f è un vettore noto in Y. L'esistenza di un vettore u in X che verifichi $(*)$, soprattutto se X è uno spazio di funzioni e A è un operatore differenziale, può essere una questione assai delicata, ma vi sono alcune conseguenze della linearità di A, semplici da verificare e molto utili nella pratica, sulle quali vale la pena di soffermarsi.

Iniziamo col riepilogare quanto già sappiamo sull'equazione

$$A(u) = 0$$

che nel linguaggio tradizionale è anche detta **EQUAZIONE OMOGENEA ASSOCIATA** a $(*)$.

E' stato provato che l'insieme delle soluzioni di tale equazione è un sottospazio di X, che si riduce al solo 0 se e solo se A è iniettiva.

Un'utile proprietà delle equazioni lineari è espressa dal

LEMMA (Principio di sovrapposizione)

Sia $A(u) = f$ un'equazione lineare e siano $u_1, u_2 \in X$ $f_1, f_2 \in Y$ tali che

$$A(u_1) = f_1 \quad e \quad A(u_2) = f_2$$

Allora $A(u_1 + u_2) = f_1 + f_2$ e $A(\lambda u_1) = \lambda f_1$

Dim. Immediato dalla linearità di A.

Il nome, molto usato in Fisica, esprime il fatto che la risposta del sistema alla somma di due forze è la somma (o sovrapposizione) degli effetti delle due forze, applicate singolarmente.

Una conseguenza immediata di tale semplice proprietà è che se $A(u) = f$ e $w \in \operatorname{Ker} A$, ossia se $A(w) = 0$, allora

$$A(u + w) = A(u) + A(w) = f + 0 = f$$

e dunque, da una qualunque soluzione di $(*)$ se ne possono ottenere altre sommandovi gli elementi del nucleo di A. In realtà le cose vanno ancora meglio: così si ottengono TUTTE le soluzioni.

LEMMA (di struttura delle soluzioni delle equazioni lineari)

Sia

$$A(u) = f \qquad (*)$$

un'equazione lineare dotata di soluzioni e sia $\bar{u} \in X$ una di esse, scelta ad arbitrio.

Allora per ogni soluzione v di $(*)$ esiste $w \in \operatorname{Ker} A$ tale che

$$v = \bar{u} + w$$

Dim. Si ha

$$A(v - \bar{u}) = A(v) - A(\bar{u}) = f - f = 0$$

e dunque $w = v - \bar{u} \in \mathrm{Ker}\, A$.

Ad onta dell'estrema semplicità della dimostrazione, tale lemma fornisce un'importante procedura per la determinazione di **TUTTE** le soluzioni di un'equazione lineare. Infatti, si deve "solo"

1) Determinare il nucleo di A o, nel linguaggio più antico, **TUTTE LE SOLUZIONI DELL'EQUAZIONE OMOGENEA ASSOCIATA.**

2) Determinare **UNA SOLUZIONE PARTICOLARE DELLA EQUAZIONE COMPLETA** $A(x) = f$, qualsiasi.

3) **SOMMARE** alla soluzione particolare tutte le soluzioni nel nucleo.

Si pensi al problema delle primitive $u' = f$: in tal caso A è l'operatore individuato dalla derivata, che è lineare per i noti teoremi sulla derivata della somma e del multiplo costante. Inoltre, il nucleo $\{u : u' \equiv 0\}$ è costituito dalle funzioni costanti su ogni sottintervallo del dominio di f, e la soluzione particolare è fornita, in teoria, dal teorema di TORRICELLI (il teorema fondamentale del calcolo integrale),

$$F(x) = \int_a^x f(t)\,dt$$

mentre, in pratica, non si dà molta importanza al fatto che tale primitiva si annulli 0 per $x = a$, e se ne utilizzano altre ricavate direttamente dall'equazione $u' = f$ mediante i procedimenti di integrazione "indefinita" (decomposizione in somme, sostituzione, parti ...).

A titolo d'esempio, consideriamo anche l'equazione differenziale

$$u' + u = 1 \qquad (**)$$

Consideriamo per prima cosa l'<u>omogenea associata</u>

$$u' + u = 0$$

Un teorema fondamentale della teoria delle equazioni differenziali assicura che la dimensione dello spazio delle soluzioni è uguale all'ordine dell'equazione (ossia al massimo ordine di derivata che appare nell'equazione) e dunque, per trovare tutti gli elementi del nucleo di $(**)$ basta trovarne uno non nullo, che genererà coi suoi multipli tutti gli altri (dimensione 1).

Cercando soluzioni del tipo $u = e^{\lambda t}$ si ottiene

$$u' = \lambda e^{\lambda t} \qquad e \qquad \lambda e^{\lambda t} + e^{\lambda t} = 0$$

da cui, dividendo per $e^{\lambda t}$ (mai nullo), si ottiene $\lambda = -1$.
Un elemento non nullo del nucleo sarà dunque e^{-t} e dunque **<u>OGNI</u>** soluzione di $u' + u = 0$ sarà del tipo

$$w = \alpha e^{-t} \quad \alpha \in \mathbb{R}$$

Cerchiamo ora una soluzione qualsiasi di $(**)$. La teoria ci assicura che devono esistere soluzioni costanti e infatti, sostituendo $u = k$ si ottiene $k = 1$ e dunque $\overline{u} \equiv 1$ è una soluzione particolare di $(**)$. Dal lemma, **<u>OGNI</u>** soluzione di $(**)$ è del tipo $v = 1 + \alpha e^{-t}$, $\alpha \in \mathbb{R}$.

RAPPRESENTAZIONE DELLE FUNZIONI LINEARI FRA SPAZI DI DIMENSIONE FINITA

Sia $A : X \to Y$ lineare, ma siano X e Y entrambi di dimensione finita. Ricordiamo che possedere basi comporta la possibilità di introdurre coordinate su X e Y. Rappresenteremo l'azione di A sui vettori di X per ottenere quelli di Y come il prodotto di una matrice appropriata, **LA MATRICE DI RAPPRESENTAZIONE DI A**, sulle coordinate di $x \in X$ per ottenere quelle di $A(x) \in Y$.

Sia dunque $x_1 \dots x_n$ una base in X, e $y_1 \dots y_m$ una in Y, fissate ad arbitrio. Sia inoltre $u = \sum_1^n u_j x_j$ un vettore generico in X. Si ha

$$A(u) = A\left(\sum_1^n u_j x_j\right) = (\text{linearità}) \sum_1^n u_j A(x_j)$$

e dunque il valore $A(u)$ in Y associato ad u si potrà determinare conoscendo le coordinate di u rispetto alla base $x_1 \dots x_n$ e n vettori "speciali" $A(x_i)$, immagini dei vettori della base. A loro volta gli n vettori $A(x_1), A(x_2), \dots, A(x_n)$ possono essere scritti come combinazioni lineari dei vettori della base $y_1 \dots y_m$ dello spazio "d'arrivo" Y

$$A(x_j) = \sum_{i=1}^m a_{ij} y_i$$

da cui

$$A(u) = \sum_{j=1}^n u_j \sum_{i=1}^m a_{ij} y_j = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} u_j\right) y_i$$

Dunque, se si definisce una matrice $A=(a_{ij})$, avente per **COLONNA** j-esima le coordinate di $\mathcal{A}(x_j)$ rispetto alla base d'arrivo, si è appena visto che le coordinate di $\mathcal{A}(u)$ rispetto a tale base sono date da

$$\sum_{j=1}^{n} a_{ij} u_j = A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

ossia dal prodotto della matrice A per il vettore (in $\mathbb{R}^n$) delle coordinate di u rispetto alla base $x_1 \dots x_n$ in X.
Tale matrice verrà detta **MATRICE DI RAPPRESENTAZIONE (O MATRICE ASSOCIATA A $\mathcal{A}$) RISPETTO ALLE BASI $x_1 \dots x_n$ in X, e $y_1 \dots y_m$ in Y.**

A titolo d'esempio consideriamo $\mathcal{A}(u)=u'$, definita sullo spazio dei polinomi di grado massimo 2 a valori in quello dei polinomi di grado massimo 1. Fissiamo le basi $x_1=1$ $x_2=t$ $x_3=t^2$ nello spazio X di partenza e $y_1=1$ e $y_2=t$ in quello d'arrivo. Allora

$$\mathcal{A}(x_1)=\mathcal{A}(1)=0=0\cdot 1+0\cdot t$$

$$\mathcal{A}(x_2)=\mathcal{A}(t)=1=1\cdot 1+0\cdot t$$

$$\mathcal{A}(x_3)=\mathcal{A}(t^2)=2t=0\cdot 1+2\cdot t$$

La matrice associata avrà le <u>colonne</u> costituite dalle coordinate delle immagini dei vettori della base di partenza prescelta, rispetto alla base d'arrivo prescelta, e cioè

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Notiamo che la matrice ha tante righe quante sono le dimensioni d'arrivo e tante colonne quante sono quelle di partenza.

La matrice associata è ovviamente dipendente dalla scelta delle basi. Supponiamo, nell'esempio precedente, di scegliere come base d'arrivo, invece di $\{1, t\}$, $\{2, 3t-1\}$, che sono indipendenti in quanto non sono uno multiplo dell'altro.
Si ha allora

$$\mathcal{A}(x_1) = \mathcal{A}(1) = 0 = 0 \cdot 2 + 0(3t-1)$$

$$\mathcal{A}(x_2) = \mathcal{A}(t) = 1 = \frac{1}{2} \cdot 2 + 0(3t-1)$$

$$\mathcal{A}(x_3) = \mathcal{A}(t^2) = 2t = \frac{1}{3} \cdot 2 + \frac{2}{3}(3t-1)$$

e dunque la matrice associata alla derivata rispetto alle basi $\{1, t, t^2\}$ in X e $\{2, 3t-1\}$ in Y è

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

Esercizio: Determinare la matrice di rappresentazione di $\mathcal{A}(u) = u'$ dallo spazio $\langle \sin t, \cos t \rangle$ in sé, rispetto alle basi $\{\sin t, \cos t\}$ $\{\sin t, \cos t\}$

Riassumendo:

LEMMA (di rappresentazione delle applicazioni lineari)

Sia $\mathcal{A}: X \to Y$ e siano $x_1 \dots x_n$ una base in X, e $y_1 \dots y_m$ una in Y.

Allora, definita una matrice $a_{ij} \in \mathbb{R}^{m\times n}$, avente per COLONNA j-esima le coordinate di $\mathcal{A}(x_j)$ rispetto a $y_1 \dots y_m$ si ha, per ogni $u \in X$

$$\mathcal{A}(u) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} u_j y_i$$

ove $u_1 \dots u_n$ sono le coordinate di u rispetto a $x_1 \dots x_n$.

Un interessante caso particolare riguarda le applicazioni da $\mathbb{R}^n$ ad $\mathbb{R}^m$ definite dal prodotto per una matrice, che vedremo presto essere, in realtà, il caso generale.

Sia $A \in \mathbb{R}^{m\times n}$ e se

$$\mathcal{A}(u) = Au \quad \forall u \in \mathbb{R}^n.$$

Calcoliamo la matrice di rappresentazione di $\mathcal{A}$ rispetto alle basi canoniche in $\mathbb{R}^n$ ed $\mathbb{R}^m$. Si ha subito che $\mathcal{A}(e_i) = Ae_i =$ (colonna i-esima di A) e le coordinate di tale vettore rispetto alla base d'arrivo (la base canonica) coincidono con le componenti di tale vettore colonna. Dunque la matrice di rappresentazione ha le stesse colonne di A e dunque coincide con A. Ovviamente tale particolarità è legata strettamente alla scelta felice delle basi di partenza e d'arrivo.

FORMA GENERALE DELLE APPLICAZIONI LINEARI FRA SPAZI EUCLIDEI $\mathbb{R}^n$.

Esaminiamo brevemente quattro casi particolari del risultato precedente, che per la loro semplicità ed importanza meritano di essere trattati direttamente.

I) Applicazioni lineari da $\mathbb{R}$ in $\mathbb{R}$.

Si ha, per ogni $x \in \mathbb{R}$, $A(x) = A(x \cdot 1) = x\,A(1)$ e, posto $a = A(1)$ si ottiene la forma generale delle funzioni lineari su $\mathbb{R}$

$$\boxed{A(x) = ax \quad \forall x \in \mathbb{R}}$$

II) Applicazioni lineari da $\mathbb{R}$ in $\mathbb{R}^n$

Esattamente come prima, si ottiene $A(x) = x\,A(1)$ e, posto $a = A(1) \in \mathbb{R}^n$, si ha ancora che ogni applicazione lineare da $\mathbb{R}$ in $\mathbb{R}^n$ assume la forma

$$\boxed{A(x) = ax}$$

ma stavolta x è scalare e a è un vettore di $\mathbb{R}^n$.

III) Applicazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}$

Sia $x \in \mathbb{R}^n$ e sia $e_1, \dots, e_n$ la base canonica in $\mathbb{R}^n$. Da $x = \sum_1^n x_i e_i$ si ha subito $\mathcal{A}(x) = \mathcal{A}\left(\sum_1^n x_i e_i\right) = \sum_{i=1}^n x_i \mathcal{A}(e_i)$.
Definiamo ora un vettore $a \in \mathbb{R}^n$ ponendo

$$a = \left(\mathcal{A}(e_1), \mathcal{A}(e_2), \dots, \mathcal{A}(e_n)\right)$$

da cui segue subito

$$\boxed{\mathcal{A}(x) = ax}$$

ove sia x sia a sono vettori, e il prodotto indicato è il prodotto scalare in $\mathbb{R}^n$.

IV) Applicazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}^m$.

Come prima, si ottiene $\mathcal{A}(x) = \sum_{i=1}^n x_i \mathcal{A}(e_i)$ e, definita una matrice

$$a = \left(\mathcal{A}(e_1), \mathcal{A}(e_2), \dots, \mathcal{A}(e_n)\right)$$

le colonne della quale sono i vettori di $\mathbb{R}^m$ immagini di $e_1 \dots e_n$ mediante $\mathcal{A}$, e ricordando che

$$(A_1 A_2 \dots A_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_1^n x_i A_i$$

si ottiene ancora

$$\boxed{A(x) = ax}$$

ove però $x \in \mathbb{R}^n$, $a \in \mathbb{R}^{m \times n}$, e il prodotto indicato è quello della matrice a per il vettore (colonna) x.

Questo appena trovato dimostra che i vari prodotti non sono solo esempi di applicazioni lineari: sono anche le uniche applicazioni lineari fra spazi $\mathbb{R}^n$. Ogni applicazione lineare da $\mathbb{R}^n$ in $\mathbb{R}^m$ è del tipo $A(x) = ax$ ove il prodotto e i suoi termini di volta in volta, vanno intesi in modo opportuno.

Queste poche osservazioni mostrano chiaramente come vettori e matrici costituiscano il substrato profondo sul quale viene edificata la teoria delle applicazioni lineari.

La situazione cambia notevolmente negli spazi di dimensione infinita, oggetto principale di studio dell'Analisi Funzionale.

FUNZIONALI LINEARI E SPAZIO DUALE

Sia X uno spazio vettoriale di dimensione finita, su $\mathbb{R}$ o $\mathbb{C}$. Denotiamo con $\mathbb{K}$ l'insieme degli scalari, qualunque esso sia, e definiamo

$$X' = \{ f : X \to \mathbb{K} \ , \ f \text{ lineare} \}$$

Notiamo fra parentesi che, nel caso X sia di dimensione infinita, la teoria è notevolmente diversa, ed f dovrà essere anche continua.

Osserviamo che se $f, g \in X'$, e $\lambda \in \mathbb{K}$, si possono definire

$$(f+g)(x) = f(x) + g(x) \qquad (\lambda f)(x) = \lambda f(x)$$

Dalle proprietà di campo di $\mathbb{K}$ (commutatività e associatività di somma e prodotto di reali o complessi, distributività del prodotto rispetto alla somma, esistenza di 0 e unità) segue immediatamente che X' è uno spazio vettoriale rispetto alla somma e il multiplo scalare appena introdotto, lo zero del quale è la funzione $0(x) \equiv 0 \ \forall x \in X$ e l'opposto di $f \in X'$ è $(-f)(x) = -f(x) \ \forall x \in X$. Tale spazio vettoriale su $\mathbb{K}$ si chiamerà **SPAZIO DUALE DI X**. I suoi elementi, le funzioni lineari su X, verranno dette **FUNZIONALI (LINEARI) SU X.**

Nelle poche note che seguono supporremo fissata una base di X, $e_1 \dots e_n$ (che non esisterebbe se X fosse di dimensione infinita), e costruiremo una base di X', ad essa associata: **LA BASE DUALE DELLA BASE DATA**, $e'_1 \dots e'_n$

Sia dunque $e_1 \dots e_n$ una base fissata di X e siano $x_1 \dots x_n$ le coordinate del vettore x rispetto a tale base.
Come già visto nei teoremi di rappresentazione, si ha, per ogni $f \in X'$

$$f(x) = \sum_1^n x_i f(e_i) \quad \forall x \in X \qquad (*)$$

Si possono ora definire n vettori in X', ossia n funzionali $e_i' : X \to \mathbb{K}$, ponendo $e_i'(x) = x_i$: e_i' è il funzionale che, ad ogni vettore x, associa la sua coordinata i-esima rispetto a $e_1 \dots e_n$.
Proviamo che i vettori $e_i' \in X'$ formano una base di X'.
Dalla precedente $(*)$ si ha subito

$$X' = \langle e_1', e_2', \dots, e_n' \rangle$$

ove i coefficienti della combinazione che genera f sono gli scalari $f(e_i)$, dipendenti dalla scelta della base $e_1 \dots e_n$.
Per provare l'indipendenza di $e_1', e_2', \dots, e_n'$ proviamo prima che

$$e_i'(e_j) = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases} \qquad (**)$$

Infatti, $e_i'(e_j)$ è l'i-esima coordinata di e_j rispetto a $e_1 \dots e_n$,
Poiché $e_j = \sum_1^n \alpha_i e_i$ ove $\alpha_i = 0$ per $i \neq j$ e $\alpha_j = 1$, e le coordinate α_i sono uniche (perché $e_1 \dots e_n$ è una base) segue subito $(**)$.
Sia ora

$$\sum_1^n \alpha_i e_i'(x) = 0(x) \qquad \left(\begin{smallmatrix} * * \\ * \end{smallmatrix}\right)$$

e proviamo che $\alpha_i = 0$, $i = 1 \dots n$.
Poiché lo zero in X' è il funzionale identicamente nullo, calcolando ambo i membri di $\left(\begin{smallmatrix} * * \\ * \end{smallmatrix}\right)$ per $x = e_j$, segue

$$\sum_1^n \alpha_i e'_i(e_j) = 0(e_j) = 0$$

Poiché per (**) il primo membro vale α_j, segue $\alpha_j = 0$. Al variare di j fra 1 ed n segue $\alpha_j = 0$ $\forall j = 1 \dots n$ e dunque $e'_1 \dots e'_n$ sono indipendenti. Ne segue infine il

TEOREMA (della base duale):

Sia X di dimensione n e sia $e_1 \dots e_n$ una base qualunque di X. Allora, i funzionali $e'_i(x)$, che associano al vettore x le sue componenti i-esime, formano una base di X', che è dunque uno spazio vettoriale di dimensione n.

Lo spazio duale è di importanza vitale nella teoria dei tensori e nella teoria avanzata delle forme differenziali, per esempio. Nella teoria classica delle forme differenziali, la base duale della base canonica ha una notazione diversa da quella qui adottata: invece di scrivere $e'_1 \dots e'_n$ per le funzioni lineari che associano ad un generico vettore w rispettivamente le sue componenti $w_1 \dots w_n$ si usa la notazione $dx_1\ dx_2 \dots dx_n$. Tale notazione fa riferimento al fatto che le applicazioni

$$x_i(w) = w_i$$

essendo lineari, coincidono coi loro differenziali in ogni punto, sicché

$$dx_i(x_0, w) = w_i \qquad \forall x_0, w$$

e dunque, in luogo di scrivere $\sum_{i=1}^n a_i(x) w_i$ si scrive $\sum_1^n a_i(x) dx_i$.

Tale notazione è la più comune in Fisica e nelle applicazioni.