ALGEBRA LINEARE

ISTRUZIONI PER L'USO.

(II)

Placido Longo

Indice

Capitolo 1

Rappresentazione delle applicazioni lineari fra spazi di dimensione finita

Lo scopo di queste note è di approfondire la struttura delle applicazioni lineari fra gli spazi di dimensione finita, avvalendosi del ruolo del tutto particolare rivestito da $\mathbb{R}^n$ nell'ambito di tali spazi: in un senso opportuno, $\mathbb{R}^n$ è *l'unico* spazio di dimensione n.
Per tale ragione verranno inizialmente studiate in dettaglio le applicazioni lineari fra spazi $\mathbb{R}^n$ e verrà provato che la loro forma generale è quella di un prodotto di un oggetto fisso per il vettore variabile nel dominio. Di volta in volta l'oggetto sarà uno scalare, un vettore, o una matrice, e il prodotto sarà di numeri, o di un numero per un vettore, o un prodotto scalare di vettori, o un prodotto di una matrice per un vettore. Verrà poi preso in considerazione il caso generale: fissate delle basi nel dominio e nel codominio, ad ogni applicazione lineare (o bilineare) verrà associata una matrice, e si vedrà come tale caso non differisca sostanzialmente da quello euclideo. Infine verrà esaminato il comportamento di tali matrici associate quando vengano cambiate le basi.
I risultati elementari e apparentemente modesti di queste sezioni permetteranno, a suo tempo, di scrivere l'espressione del differenziale di una funzione di più variabili mediante le derivate parziali, che sono facili da calcolare, e di introdurre il problema della diagonalizzazione fornendo le condizioni sotto le quali una matrice sia diagonalizzabile: questioni davvero importanti, fra molte altre!
Analogamente a quanto già visto nel capitolo su $\mathbb{R}^n$, anche questi resultati hanno ispirato alcuni sviluppi dell'Analisi Funzionale negli spazi di dimensione infinita: la loro estensione a tali ambienti ha richiesto assai più che una semplice riformulazione, ed ha ampliato notevolmente il loro campo d'applicazione. Il teorema di Riesz, ad esempio, che estende agli spazi euclidei astratti – in realtà, a quelli *completi*, i cosiddetti *spazi di Hilbert* – il teorema di struttura delle applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}$, può essere impiegato per dimostrare un teorema di esistenza e unicità per le soluzioni dell'equazione di Laplace, che sono i potenziali (gravitazionale, elettrostatico) prodotti da una distribuzione di cariche.

1.1 Applicazioni lineari fra spazi $\mathbb{R}^n$

Iniziamo con il caso più semplice: le applicazioni fra spazi euclidei. Esamineremo nell'ordine le applicazioni lineari da $\mathbb{R}$ in $\mathbb{R}$, da $\mathbb{R}$ in $\mathbb{R}^n$, da $\mathbb{R}^n$ in $\mathbb{R}$, e da $\mathbb{R}^n$ in $\mathbb{R}^m$. Introdurremo poi la convenzione di Einstein (sì, quello della Relatività), per poter scrivere in forma abbreviata alcune sommatorie.

1.1.1 La struttura delle applicazioni lineari su $\mathbb{R}$

Data un'applicazione lineare $T : \mathbb{R} \to \mathbb{R}$ si può osservare che

$$T(x) \,=\, T(x\,1) \,=$$

(per la linearità di T)

$$=\, x\,T(1)$$

L'elemento $T(1)$ è uno scalare. Posto $a \equiv T(1)$, si osserva subito che la forma generale di una funzione lineare da $\mathbb{R}$ in $\mathbb{R}$ è

$$T(x) \,=\, a\,x$$

In modo pressoché identico si possono studiare le applicazioni da $\mathbb{R}$ in $\mathbb{R}^n$: basta infatti ripetere il ragionamento e osservare che stavolta il valore $T(1)$ è un vettore di $\mathbb{R}^n$. Posto di nuovo $a \equiv T(1)$, sarà ancora vero che

$$T(x) \,=\, a\,x$$

ma stavolta il prodotto non va inteso come un prodotto dei numeri a e x, ma come quello fra il numero x ed il vettore a.

1.1.2 La struttura delle funzioni lineari su $\mathbb{R}^n$

Il prossimo caso trattato è quello delle applicazioni lineari T da $\mathbb{R}^n$ a $\mathbb{R}$. Questa volta, invece di scrivere $x = x\,1$ scriveremo $x \,=\, \sum_{i=1}^{n} x_i e_i$, ove e_i è la base canonica di $\mathbb{R}^n$. Dalla linearità di T si ottiene:

$$T(x) \,=\, T(\sum_{i=1}^{n} x_i e_i) \,=\, \sum_{i=1}^{n} x_i T(e_i)$$

Osserviamo che gli n valori $T(e_i)$ sono numeri, e introduciamo il vettore

$$a \equiv (T(e_1), \cdots, T(e_n))$$

Dalla definizione di prodotto scalare si ottiene (ancora!)

$$T(x) \,=\, a\,x$$

ove però il prodotto va stavolta inteso come il prodotto scalare fra i due vettori a e x.
Ricordando che lo *spazio duale* di uno spazio vettoriale è lo spazio delle funzioni lineari su di esso, ne segue che ogni elemento del duale di uno spazio euclideo è un prodotto scalare per un opportuno vettore.

1.1.3 La struttura generale delle applicazioni lineari fra spazi euclidei

Trattiamo ora il caso generale: quello di un'applicazione lineare fra spazi euclidei. Sia dunque $T : \mathbb{R}^n \to \mathbb{R}^m \;\; n, m > 1$, lineare. Ripetendo il ragionamento della sezione precedente, si ottiene

$$T(x) \,=\, \sum_{i=1}^{n} x_i T(e_i)$$

Gli n valori $T(e_i)$ sono ora vettori in $\mathbb{R}^m$. Introducendo la matrice

$$a \qquad \equiv \qquad (T(e_1), \cdots, T(e_n))$$

avente come prima colonna il vettore $T(e_1)$, come seconda $T(e_2)$ e così via fino all'n-esima colonna che è $T(e_n)$ (che risulta una matrice ad m righe ed n colonne) si osserva subito che vale

$$a\,x \,=\, \sum_{i=1}^{n} x_i T(e_i)$$

Si può dunque concludere che:

Teorema 1 *Ogni applicazione lineare fra $\mathbb{R}^n$ ed $\mathbb{R}^m$, con $n, m > 1$, si rappresenta come un prodotto di una matrice $m \times n$ per il vettore variabile nel dominio.*

In simboli,

Formula 1

$$\forall T : \mathbb{R}^n \to \mathbb{R}^m \text{ lineare, } n, m > 1, \qquad \exists a \in \mathbb{R}^{m \times n} \qquad T(x) \;=\; a\,x$$

Tale prodotto assume però le forme speciali esaminate nelle precedenti sezioni, se n o m (o entrambe) valgono 1. L'utilità di poter considerare i casi delle sezioni precedenti come casi particolari di questa formula generale ci porta diritti alla prossima sezione, che tratta (solo in apparenza!) di una questione di notazione.

1.1.4 La convenzione di Einstein

Nella formulazione della teoria della relatività generale vennero adoperati sistematicamente da Einstein i metodi del "Calcolo differenziale assoluto" di Ricci-Curbastro, come allora si chiamava il calcolo tensoriale sulle varietà riemaniane, e dovette esserne rimasto impressionato (in tutti i sensi!) perché da un canto ne disse e ne scrisse ogni bene in ogni occasione per la sua efficacia, ma dall'altro si avvalse largamente della collaborazione della moglie Mileva Marić, brava e più paziente di lui, per eseguire i terrificanti calcoli richiesti dalla sua teoria. Un risultato utile per tutti noi fu la convenzione, a lui intitolata, di sottintendere tutte le sommatorie relative ad indici ripetuti, il che riduce notevolmente l'intensità delle raffiche di sommatorie, così abbondanti nel calcolo tensoriale. Più precisamente, la *convenzione di Einstein* consiste nel sottintendere la somma ogni qual volta appaiano indici ripetuti in alto e in basso, dei quali non sappiamo ancora nulla e che dovremo cominciare ad usare presto! La adopereremo con i nostri tensori di tipo molto semplice: scalari, vettori riga, vettori colonna (ben distinti dai vettori riga), e matrici. In tale ordine d'idee, tutti gli oggetti saranno pensati come matrici. In particolare: gli scalari appartengono a $\mathbb{R}^{1\times 1}$; i vettori riga a $\mathbb{R}^{1\times n}$; i vettori colonna a $\mathbb{R}^{n\times 1}$. Dopo aver trasformato tutto in matrici, è semplice introdurre la notazione con gli indici in alto e in basso, e la *convenzione di Einstein* che ne consegue: basta scrivere gli indici di *riga* in alto e quelli di *colonna* in basso. Dunque:

Definizione 1 *Il simbolo a^i_j denoterà l'elemento di riga i e colonna j della matrice a.*

Definizione 2 (Convenzione di Einstein) *In ogni espressione di matrici (e in particolare di scalari, vettori riga o vettori colonna), ovunque appaiano indici ripetuti in alto e in basso deve considerarsi sottintesa una somma, al variare dell'indice ripetuto nell'insieme dei suoi valori.*

Il campo di variazione degli indici (cioè la dimensione degli spazi, almeno nei casi di uso più frequente) è noto *a priori*, e dunque non ci si rimette nulla a non indicarlo ossessivamente molte volte in ogni pagina. Un fenomeno analogo accade in molti linguaggi di programmazione, dove il *tipo* degli *array* viene dichiarato una volta per tutte al momento di allocare la memoria e non cambia più. Altra analogia si ravvisa nel sistematico *overloading* del prodotto (o di altri simboli), che vuol dire cose diverse, o addirittura non è definito, a seconda del *tipo* delle variabili fra le quali opera.

Per quanto scomodo (e noioso), dobbiamo riesaminare quanto già fatto per tenere nel debito conto gli effetti della notazione tensoriale e della convenzione di Einstein.
Per cominciare, denoteremo con a_i l'elemento i-esimo del vettore riga $a = (a_1, a_2, ..., a_n)$ (l'indice indica le colonne!), mentre a^i sarà l'elemento i-esimo del vettore colonna $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ (varia l'indice di riga).

Consideriamo ora il prodotto scalare. Si può facilmente sottintendere la somma contenuta nella definizione scrivendo $a_i\, b^i$ (o anche $b_i a^i$), ma lo scotto da pagare è la necessità di definire il prodotto solo fra un vettore riga e uno colonna. Niente di scandaloso, perché non si può definire il prodotto di matrici con risultato scalare (e quindi in $\mathbb{R}^{1\times 1}$) fra due matrici in $\mathbb{R}^{n\times 1}$, ma occorre che siano la prima di tipo $1 \times n$ e la seconda di tipo $n \times 1$, e quindi la prima un vettore riga e la seconda un vettore colonna. Se si vuole definire definire il vecchio prodotto scalare fra vettori colonna basta moltiplicare il trasposto del primo per il secondo, ma la convenzione di Einstein funziona in modo "pulito" solo fra un vettore riga e uno colonna. Riesaminiamo brevemente, sotto questa luce, il teorema di struttura delle funzioni lineari su $\mathbb{R}^n$. Si ha:

$$f(x) \,=\, x^i f(e_i) \,=\, x^i a_i$$

dove a è il vettore *riga* ottenuto dalle immagini dei vettori della base di partenza. Una domanda scaturisce immediatamente: "Forse i vettori riga rappresentano le funzioni su $\mathbb{R}^n$, cioè gli elementi del suo spazio duale"? Risposta affermativa!
Consideriamo l'applicazione che associa ad ogni $f : \mathbb{R}^n \to \mathbb{R}$ il vettore riga a che la rappresenta. Dalla linearità del prodotto scalare e della f segue subito che tale applicazione è lineare. È anche iniettiva, perché da

$$x^i a_i \,=\, x^i b_i \qquad \Rightarrow \qquad x^i(a_i - b_i) = 0 \qquad \forall x \in \mathbb{R}^n$$

da cui, ponendo $x = e_i, i = 1..n$ segue subito $a_i - b_i = 0 \quad \forall i = 1..n$. La suriettività è immediata: dato qualsiasi vettore riga a la funzione $x^i a_i$ ha come immagine esattamente quel vettore. L'applicazione che associa ad un elemento del duale il vettore riga che lo rappresenta, essendo iniettiva, suriettiva e lineare, è un isomorfismo di spazi vettoriali non dissimile dall'isomorfismo canonico, introdotto più avanti. Ha senso chiamarlo *isomorfismo canonico duale*, per via del legame che stabilisce fra i vettori riga e il duale, ossia l'insieme delle funzioni lineari sullo spazio $\mathbb{R}^n$.

Le matrici presentano meno problemi. Infatti, data $A = (a^i_j)$, l'espressione

$$a^i_j\, x^j \,=\, \sum_{j=1}^{n} a^i_j\, x^j$$

rappresenta, al variare di i, un vettore colonna, combinazione lineare delle colonne di A secondo i coefficienti x^j e dunque è il vettore $A\,x$.
C'è un certo numero di "bizzarrie" conseguenti alla rappresentazione di tutti gli oggetti con matrici, e dei loro prodotti con il prodotto di matrici, che vale la pena di esaminare in dettaglio. Ad esempio, il prodotto dello scalare α, che è una matrice 1×1 per il vettore u va *di necessità* scritto nella forma $\alpha u = \alpha u_i$ se $u \in \mathbb{R}^{1\times n}$ e cioè è un vettore riga, mentre va scritto come $u\alpha$ se $u \in \mathbb{R}^{n\times 1}$ è un vettore colonna. Del prodotto scalare dei vettori u, v abbiamo già detto: va obbligatoriamente scritto in una delle forme $u_i v^i$ o $u^i v_i$: per definirlo su una coppia di vettori colonna occorre trasporre il primo vettore per trasformarlo in una riga, cosa che faremo di frequente. Queste ed altre stranezze hanno fatto sì che la convenzione di Einstein non abbia soppiantato del tutto il modo solito di scrivere gli oggetti dell'algebra. Ad ogni modo, la utilizzeremo ovunque sia vantaggioso e qualunque sia il senso della somma e del prodotto (scalari, vettori, matrici): ad esempio, la somma (sottintesa) $x^i e_i$, prima utilizzata, vale in realtà $\sum_1^n x^i e_i$, che è la combinazione lineare dei vettori e_i della base canonica con i coefficienti x^i. Ciascuno dei vettori della base canonica è una colonna di numeri, e il sistema e_i è in realtà pensato come una riga di vettori colonna che moltiplica a sinistra, con la regola di Einstein, la colonna x^i.
In questo ordine di idee, quando intendiamo usare la convenzione di Einstein rappresenteremo i vettori di $\mathbb{R}^n$ come vettori colonna (da $x = x^i e_i$) mentre rappresenteremo le funzioni lineari su $\mathbb{R}^n$ con i vettori riga (da $f(x) \,=\, x^i f(e_i) \,=\, x^i a_i$).
Per via del loro comportamento nel cambio di base, che studieremo tra breve, i primi si chiamano anche *vettori controvarianti* e i secondi *vettori covarianti*.

Come si diceva più su, il trionfo della convenzione di Einstein è l'analisi tensoriale, strumento principe per lo studio della curvatura, che non esulerebbe affatto dall'ambito di queste note (al contrario!), ma che purtroppo eccederebbe largamente le limitazioni di estensione per un corso introduttivo di algebra anche se ci limitassimo ai soli aspetti strettamente algebrici. Un'esposizione (a giudizio dell'autore) molto chiara dell'algebra dei tensori, almeno per chi sappia l'inglese, è contenuta nel secondo capitolo del libro di R.Bishop e S. Goldberg "Tensor Analysis on Manifolds" ed. Dover.
Una curiosità: il nome *tensori* trae la sua origine dal fatto che l'esempio più pregnante proveniente dalla fisica per un tale oggetto è il *tensore degli sforzi*, una delle chiavi di volta della meccanica dei sistemi continui (materiali elastici o fluidi). L'idea consiste nel considerare una piccola superficie piana e "misurare" la forza attraverso quella superficie: una breve riflessione convincerà il lettore che tale forza non dipende solo dallo stato di tensione del materiale, ma anche dell'orientamento della superficie rispetto alla "direzione di provenienza" dello sforzo; così mentre un campo di forze associa ad ogni punto un vettore (la forza), un "campo di sforzi" associa ad ogni punto e ad ogni direzione (la normale alla piccola superficie) una forza. È una cosa diversa, e per approfondimenti i riferimenti d'obbligo sono i corsi di Scienza delle Costruzioni o quelli di Dinamica dei Fluidi nelle Facoltà d'Ingegneria, con le relative bibliografie. Un bel libro, non facile, che è stato consigliato all'autore da un amico di Aerospaziale, è "Introduction to Fluid Dynamics" di G.K. Batchelor ed. Cambridge University Press. Sarà chiaro a tutti, a questo punto, che è meglio imparare alla svelta l'inglese!!

1.2 Applicazioni lineari e matrici

Lo stretto legame fra applicazioni lineari in $\mathbb{R}^n$ e matrici non è conseguenza della particolare struttura dello spazio euclideo, ma del fatto che esso è uno spazio di dimensione finita, e dunque il comportamento dell'applicazione lineare è completamente determinato da quello su una base. Dopo aver provato che tutti gli spazi di uguale dimensione sono isomorfi, e cioè sostanzialmente uguali, ci si potrà allora aspettare un analogo comportamento anche per le applicazioni fra spazi qualunque, purché di dimensione finita. Nelle sezioni che seguono verrà esaminata questa questione, e verranno estesi in modo opportuno alle applicazioni lineari fra generici spazi di dimensione finita i resultati di rappresentazione già ottenuti per quelle fra spazi euclidei.

1.2.1 L'isomorfismo canonico associato ad una base

Supponiamo che X sia uno spazio vettoriale di dimensione finita e che dunque esistano un numero finito di vettori indipendenti $u_1, u_2, ..., u_n$ tali che

$$X = \langle u_1, u_2, ..., u_n \rangle$$

Per ogni x in X esisteranno dei numeri $x_1, x_2, ..., x_n$ tali che (usando la convenzione di Einstein dopo aver posto i coefficienti in colonna) $x = x^i u_i$. La colonna definita dai coefficienti x^i della combinazione lineare che genera x è un vettore in $\mathbb{R}^n$. Ogni vettore ha un unico sistema di coefficienti associato: infatti da $x = x^i u_i = y^i u_i$ segue subito $(x^i - y^i)u_i = 0$ e, dall'indipendenza lineare, $x^i - y^i = 0$ per ogni i, e dunque i due sistemi di coefficienti coincidono. Quanto osservato giustifica la seguente definizione:

Definizione 3 *Ad ogni base $u_1, u_2, ..., u_n$ in uno spazio vettoriale X di dimensione finita n è associata una funzione $I(x)$, detta* isomorfismo canonico*, definita su X, a valori in $\mathbb{R}^n$, che associa ad ogni x in X il vettore colonna delle sue componenti (o* coordinate*) rispetto alla base fissata.*

In simboli:

Formula 2

$$\forall x \in X, \qquad x = x^i u_i \qquad \Rightarrow \qquad I(x) \qquad \equiv \qquad \begin{pmatrix} \vdots \\ x^i \\ \vdots \end{pmatrix}$$

È stato già osservato che i è definita per ogni vettore dello spazio, e che è iniettiva, per effetto dell'indipendenza dei vettori della base. È anche immediato, dalla distributività del prodotto, provare che è lineare. Osserviamo anche che I è suriettiva: infatti, ogni sistema α^i è immagine del vettore $\alpha^i u_i$. Poiché è iniettiva e suriettiva, è invertibile e l'inversa sarà lineare: è quindi un isomorfismo.
In sostanza, l'isomorfismo canonico associa ad ogni vettore la colonna dei coefficienti da attribuire alla base fissata per ottenerlo, mentre il suo inverso associa ad ogni vettore di $u \in \mathbb{R}^n$ il vettore di X ottenuto considerando la combinazione lineare dei vettori della base che si ottiene utilizzando come coefficienti le componenti di u.
Il motivo del nome "importante" (le radici greche della parola isomorfismo indicano che i due spazi hanno la stessa forma... lo stesso aspetto, insomma) indica che l'isomorfismo, oltre ad essere una corrispondenza biunivoca fra lo spazio originale X e lo spazio euclideo $\mathbb{R}^n$ di uguale dimensione, ne conserva le operazioni: invece di sommare due vettori di X direttamente, si può ottenere lo stesso resultato considerando i loro corrispondenti vettori di $\mathbb{R}^n$, sommandoli in $\mathbb{R}^n$ (e cioè sommandone le coordinate), e prendendo infine il corrispondente in X della somma così ottenuta mediante l'inverso dell'isomorfismo canonico. Analogamente si può ragionare per il prodotto per uno scalare.
È questo il senso nel quale si diceva nell'introduzione che tutti gli spazi di uguale dimensione n sono "uguali" ad $\mathbb{R}^n$: non sono proprio uguali... sono isomorfi; hanno punti corrispondenti e operazioni corrispondenti. Osserviamo esplicitamente che la corrispondenza introdotta (l'isomorfismo canonico) dipende dalla base prescelta: ciò si rivelerà di grande importanza.

Un'altra considerazione di tipo "geometrico" può aiutare a visualizzare meglio la questione: fissare una base equivale a fissare un sistema di coordinate per ogni punto dello spazio, rappresentate dai coefficienti della combinazione lineare che lo genera, che sono definiti in modo unico. L'isomorfismo canonico associa ad ogni punto dello spazio le sue coordinate, e il suo inverso associa ad ogni sistema di n numeri il vettore generato dalla combinazione lineare degli elementi di *quella* base con *quei* coefficienti, ossia il vettore che ha *quei* numeri come coordinate. Applicare ad un vettore l'isomorfismo canonico equivale a prenderne le coordinate, mentre applicare ad un sistema di numeri il suo inverso equivale a considerare il punto con quelle coordinate. Se si sceglie come spazio $\mathbb{R}^n$ e come base la base canonica, l'isomorfismo canonico è semplicemente l'identità: associa ad ogni vettore, definito da n numeri, il vettore di $\mathbb{R}^n$ formato dai coefficienti da attribuire alla base canonica per ottenere quel vettore, che coincidono ovviamente con i numeri che definiscono il vettore.
Per meglio comprendere il ruolo dell'isomorfismo canonico, occorre dunque allontanarsi da $\mathbb{R}^n$. Consideriamo quindi uno spazio di dimensione finita diverso.

Un esempio "pratico" di isomorfismo canonico

Sia X lo spazio dei polinomi di grado non maggiore di 1, che è di dimensione 2 ed è generato (per esempio) dalla base $u = 1$ e $v = t + 2$. Rispetto a questa base, il vettore di $\mathbb{R}^2$ associato dall'isomorfismo canonico al polinomio $t \in X$ (che è uguale a $v - 2u$) sarà

$$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

mentre il polinomio che l'inverso dell'isomorfismo canonico associa alla colonna

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

sarà $3u - v$ e cioè $3 - t - 2 = 1 - t$

1.2.2 La matrice associata ad un'applicazione lineare fra spazi di dimensione finita

Il passo successivo da compiere è di considerare un'applicazione lineare fra spazi di dimensione finita. I discorsi della sezione precedente inducono a sperare che ci debba essere qualche relazione fra esse e le applicazioni lineari fra spazi euclidei, rivelatesi in precedenza come prodotti di matrici per vettori, in quanto il concetto di linearità richiede per la sua formulazione solo somme e prodotti per scalari, che sono "conservati" dall' isomorfismo canonico. Infatti è così!

Siano X, Y due spazi vettoriali di dimensione n ed m rispettivamente, e sia $T : X \to Y$ un'applicazione lineare da X a Y. Siano $u_1, u_2, ..., u_n$ e $v_1, v_2, ..., v_m$ due basi rispettivamente di X e Y. Siano infine I e J gli isomorfismi canonici associati in X e Y alle basi prima introdotte, e x^i e y^j le coordinate dei generici punti $x \in X$ e $y \in Y$ ($x^i = I(x)$ e $y^i = J(y)$). Si ha allora, ragionando come per $\mathbb{R}^n$,

$$T(x) = T(x^i u_i) = x^i T(u_i)$$

I valori $T(u_i)$ sono elementi di Y, e quindi possono scriversi come combinazione lineare degli elementi della base $v_1, v_2, ..., v_m$. Definiamo ora una matrice $A = (A_1, A_2, ..., A_n)$ la cui *colonna* i-esima A_i è formata dalle m coordinate di $T(u_i)$ rispetto alla base di Y, e cioè il valore di J sul vettore $T(u_i)$. In simboli:

$$(A_1, A_2, ..., A_n) \qquad \equiv \qquad (\, J(T(u_1)), J(T(u_2)), \cdots, J(T(u_n))\,)$$

La matrice A è di tipo $m \times n$, e sia a^h_k il suo termine generico. Dalla sua definizione, risulta

$$T(u_i) = a^j_i v_j$$

e di conseguenza, sostituendo nella prima uguaglianza,

$$T(x) = x^i a^j_i v_j$$

Dunque, la coordinata j da attribuire a v_j nella combinazione lineare per ottenere $T(x)$ è

$$x^i a^j_i = \sum_{i=1}^{n} x^i a^j_i$$

Dunque, invece di ottenere $T(x)$ direttamente, si può

- Passare da x a $I(x)$ i cui termini sono x^i
- Calcolare $x^i a^j_i$ (ricordare la convenzione di Einstein sull'indice i), il che equivale a calcolare il prodotto della matrice a^i_j per il vettore $I(x)$ di $\mathbb{R}^n$.
- Utilizzare il vettore di $\mathbb{R}^m$ ottenuto dal prodotto precedente come vettore delle coordinate di $T(x)$ rispetto alla base di Y, dalle quali ottenere infine $T(x)$ applicando J^{-1}.

$$T(x) = J^{-1}(A\, I(x))$$

Se si pone $\tilde{A}(x) = A\,x$, essa risulta un'applicazione lineare definita su $\mathbb{R}^n$, a valori in $\mathbb{R}^m$, che viene detta *applicazione lineare associata ad A.* Essa verifica

$$T = I \circ \tilde{A} \circ J^{-1}$$

ove con il simbolo $\circ$ si è denotato il prodotto di composizione (con l'ordinaria convenzione di applicare i fattori alla variabile nell'ordine in cui appaiono. Ne segue subito il

Teorema 2 *Per ogni applicazione lineare T fra spazi di dimensioni finite, ed ogni scelta di basi per essi, esiste un'unica matrice A, detta* matrice associata all'applicazione lineare e alle basi scelte, *che verifica*

$$T \,=\, I \circ \tilde{A} \circ J^{-1}$$

ove I e J sono rispettivamente gli isomorfismi canonici relativi alle basi scelte nel dominio e nel codominio, e $\tilde{A}$ è l'applicazione lineare fra $\mathbb{R}^n$ ed $\mathbb{R}^m$ definita dalla matrice A .

A dire il vero, non abbiamo verificato l'unicità: è un semplice esercizio.
In altri termini, più "ruspanti"

Teorema 3 *Data un'applicazione $T : X \to Y$, con X, Y di dimensioni finite n ed m, e due basi prefissate di X e di Y, esiste una matrice $A = (a^i_j)$ che moltiplica a sinistra il vettore in $\mathbb{R}^n$ delle coordinate in X di ogni punto x, per ottenere il vettore $a^i_j x^j \;\in \mathbb{R}^m$ delle coordinate in Y del punto $T(x)$; le sue colonne sono formate dalle coordinate, rispetto alla base del codominio, delle immagini dei vettori della base del dominio.*

e infine

Definizione 4 *Data un'applicazione $T : X \to Y$, con X, Y di dimensioni finite n ed m, e due basi prefissate di X e di Y, si definisce* matrice associata a T e alle basi *la matrice le cui n colonne sono formate dalle m coordinate, rispetto alla base scelta in Y, delle immagini dei vettori della base scelta in X, presi nell'ordine.*

Esaminiamo tutto meglio attraverso un esempio.

Calcolo della matrice associata: un esempio

Sia X lo spazio dei polinomi di grado minore o eguale a 2, e Y lo spazio delle combinazioni lineari di $\sin t$ e $\cos t$. Una base ragionevole di X è formata da $1, t, t^2$, mentre per prendere come base di Y il sistema di generatori che appare nella definizione occorre provarne l'indipendenza. Una via molto rapida è la seguente. Se si ha $\lambda \sin t + \mu \cos t \;\equiv\; 0$, derivando ambo i membri si ottiene anche $\lambda \cos t - \mu \sin t \;\equiv\; 0$. Il sistema lineare omogeneo nelle incognite λ e μ (ha determinante $\sin^2 + \cos^2 t \equiv 1$ per cui) ha l'unica soluzione $\lambda = \mu = 0$, da cui l'indipendenza.

Consideriamo ora l'applicazione che ad ogni polinomio p associa

$$T(p) = \sin t \int_0^1 p(s)ds \,+\, \cos t \int_1^2 p(s)ds$$

È immediato, dalla linearità dell'integrale, verificare che T è lineare da X in Y.
Per calcolare la matrice associata ad essa e alle basi scelte più su, consideriamo le immagini dei vettori della base in X. Si ha

$$\int_0^1 1ds = 1 \qquad \int_0^1 sds = \frac{1}{2} \qquad \int_0^1 s^2ds = \frac{1}{3}$$

mentre

$$\int_1^2 1ds = 1 \qquad \int_1^2 sds = \frac{3}{2} \qquad \int_1^2 s^2ds = \frac{7}{3}$$

da cui infine

$$T(1) \,=\, \sin t + \cos t \qquad\qquad T(t) \,=\, \frac{1}{2}\sin t + \frac{3}{2}\cos t \qquad\qquad T(t^2) \,=\, \frac{1}{3}\sin t + \frac{7}{3}\cos t$$

Dunque, $T(1), T(t), T(t^2)$ hanno rispettivamente coordinate $(1,1), (\frac{1}{2}, \frac{3}{2}), (\frac{1}{3}, \frac{7}{3})$ e la matrice richiesta è

$$A \,=\, \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1 & 3/2 & 7/3 \end{pmatrix}$$

Il vantaggio che salta immediatamente agli occhi è che i calcoli macchinosi, nel nostro caso rappresentati dagli integrali, vengono eseguiti una volta per tutte sui soli vettori della base di partenza: da quel momento in poi, sono necessari solo prodotti e somme di numeri, e cioè il prodotto della matrice associata per le coordinate del punto in esame, e una combinazione lineare di vettori per riconvertire le coordinate e determinarne l'immagine nello spazio d'arrivo.
Noti i valori su una base, la linearità permette ovviamente di calcolare direttamente, in pratica, l'immagine di un punto:

$$T(t+1) = T(t) + T(1) = \frac{1}{2}\sin t + \frac{3}{2}\cos t + \sin t + \cos t \;=\; \frac{3}{2}\sin t + \frac{5}{2}\cos t$$

Il calcolo, peraltro, corrisponde esattamente a quello eseguito per definire la matrice associata.

In conclusione, il concetto di matrice associata permette in qualche modo di estendere a qualsiasi applicazione lineare fra spazi di dimensione finita il teorema di struttura già provato per le applicazioni fra spazi euclidei: sono gli isomorfismi canonici associati alle basi, introducendo sistemi di coordinate nel dominio e nel codominio, a trasformare in modo isomorfo l'applicazione iniziale in una fra spazi euclidei, con la sua particolare struttura di prodotto di una matrice (la matrice associata) per un vettore (delle coordinate del punto del dominio).

1.2.3 La matrice associata ad un'applicazione bilineare fra spazi di dimensione finita

Un altro caso interessante è rappresentato dalle applicazioni bilineari, che si prestano anch'esse ad essere rappresentate mediante matrici.

Iniziamo con una definizione.

Definizione 5 *Dato uno spazio vettoriale reale, una funzione (o anche forma, o applicazione)* bilineare *è una funzione $\alpha : X \times X \to \mathbb{R}$ lineare rispetto ad ogni variabile, ossia tale che $u \to \alpha(u, v)$ sia lineare per ogni v fissata, e $v \to \alpha(u, v)$ sia lineare per ogni u fissata.*

La definizione appena presentata si può estendere sia considerando spazi vettoriali complessi e funzioni a valori complessi, sia considerando *funzioni multilineari* su X, ossia funzioni definite su X^n, a valori in $\mathbb{R}$ (o $\mathbb{C}$) e lineari rispetto ad ogni variabile non appena si fissino le altre. Uno degli esempi più importanti di tali funzioni è il determinante, pensato come funzione delle colonne. Un'ulteriore estensione è quella di considerare come dominio un prodotto di n spazi, anche distinti fra loro, invece di imporre che debbano tutti coincidere (con X).
La rappresentazione di funzioni bilineari non richiede idee nuove: basta fissare una base e sviluppare tutti i vettori secondo la base. Infatti, detta $u_1, u_2, ..., u_n$ la base prescelta di X, si ha:

$$\alpha(x, y) \;=\; \alpha(x^i u_i \,,\, y^j u_j) \;=$$

(dalla linearità rispetto alla prima variabile, fissata la seconda)

$$=\; x^i \alpha(u_i \,,\, y^j u_j) \;=$$

(dalla linarità rispetto alla seconde variabili, fissate le prime)

$$=\; x^i y^j \alpha(u_i, u_j)$$

Una volta di più abbiamo ottenuto un'espressione in termini di coordinate: il valore della forma su qualunque coppia di vettori è determinato solo dalle loro coordinate e dai suoi valori sulle coppie di vettori della base, la cui scelta, ovviamente, influenza anche i valori delle coordinate associati ad x e ad y. Occorre resistere alla tentazione di definire di getto la matrice di rappresentazione come quella formata con i numeri $\alpha(u_i, u_j)$, almeno se si vuole continuare ad utilizzare la convenzione di Einstein, perché i sistemi per i quali andrebbe moltiplicata hanno gli indici entrambi in alto,

e le matrici ordinarie hanno righe e colonne e quindi indici in alto e in basso. Alcune soluzioni possibili sono: eliminare definitivamente gli indici in alto e in basso e continuare a mantenere la convenzione sulla somma sugli indici ripetuti sottintesa, ed è il punto di vista di Lev Landau, un altro grande fisico; oppure non ostinarsi a scrivere sempre i vettori in $\mathbb{R}^n$ come se esistessero solo le colonne, e pensare che le matrici rappresentino solo le forme bilineari definite su un vettore riga e uno colonna, e questo è il punto di vista dell'algebra tensoriale moderna, per i dettagli del quale rinviamo il lettore al citato libro di Bishop e Goldberg; oppure ancora "sottintendere", o indicare esplicitamente, come noi faremo, la trasposizione del primo argomento. Vediamo i dettagli.
Torniamo all'ultima espressione e definiamo sì una matrice $A = (a^i_j)$ ponendo

$$a^i_j \;=\; a(u_i, u_j)$$

e osservando che

$$a(x, y) \;=\; x'_i y^j a^i_j \;=\; x' A y$$

dove la trasposizione si incarica di trasformare il primo vettore x da colonna a riga. La matrice A così definita è la *matrice associata* all'applicazione bilineare.
Dunque

Definizione 6 *Data un'applicazione bilineare α definita sulle coppie x e y di vettori di X, a valori reali, e una base $u_i, i = 1..n$ in X, si definisce* matrice associata *all'applicazione bilineare la matrice $a^i_j \;=\; \alpha(u_i, u_j)$. Essa verifica $\alpha(x, y) = x'_i\, a^i_j\, y^j$, ove x'_i è il vettore (riga) trasposto del vettore delle coordinate di x rispetto alla base di X, e y^j è il vettore (colonna) delle coordinate di y.*

Si osservi che tale matrice è biunivocamente determinata, perché se due forme bilineari hanno la stessa matrice associata allora coincidono sulle coppie di vettori della base e dunque, per l'espressione generale delle forme bilineari, coincidono per ogni coppia di vettori.

Tale concetto può essere generalizzato senza fatica al caso in cui a non sia definita su $X \times X$, ma su coppie di vettori appartenenti a spazi differenti. Si procede quasi esattamente allo stesso modo già visto: prima, si calcolano le coordinate dei vettori x e y, ciascuno rispetto alla base prescelta nel proprio spazio di appartenenza, e poi tutto è esattamente uguale.

Osserviamo inoltre che se l'applicazione bilineare più su studiata, invece che essere definita su $X \times X$, fosse stata definita su $X^* \times X$, dove con X^* si è denotato lo spazio delle funzioni lineari su X, sarebbe stato tutto più semplice, in quanto è già stato osservato che gli elementi di X^* sono rappresentati da vettori riga, e quindi non sarebbe neppure necessario di prendere il trasposto del primo vettore.
In algebra tensoriale, si trae profitto da questa circostanza definendo un *tensore di tipo (r, s) su X* come un'applicazione multilineare a valori reali (o complessi, al bisogno), definita sul prodotto cartesiano di r "copie" di X^* per s "copie" di X, e cioè, applicati gli isomorfismi canonici in ogni spazio, una funzione multilineare reale definita su r vettori riga e s vettori colonna. Con questa definizione, le matrici rappresentano direttamente i tensori di tipo $(1, 1)$ (vedi Bishop e Goldberg). Aggiungendo qualche trasposto qua e là, possono essere rappresentati anche i tensori di tipo $(2, 0)$ e $(0, 2)$.

1.3 Cambi di base

1.3.1 La matrice associata ad un cambio di base

Sia X uno spazio di dimensione finita e si considerino due basi $u_1, u_2, ..., u_n$ e $v_1, v_2, ..., v_n$. L'obiettivo di questa sezione è di determinare come cambino le coordinate di un dato punto al passaggio da una base ad un'altra.

Iniziamo con l'osservare che ogni vettore della base "nuova" è combinazione dei vettori di quella vecchia e viceversa. Definiamo allora una matrice M, la *matrice di trasformazione*, ponendo

$$v_j \,=\, m_j^i u_i$$

e cioè la matrice la cui *colonna j-esima* è formata dai coefficienti da attribuire ai vettori della base 'vecchia" per ottenere v_j (già visto qualcosa di simile, vero?)
Posto anche

$$u_j \,=\, n_j^i v_i$$

e sostituendo al posto di v_i l'espressione ottenuta in precedenza, risulta

$$v_j \,=\, m_j^h\, n_h^k v_k$$

da cui segue subito che le due matrici sono una inversa dell'altra, in quanto il loro prodotto è la matrice identica.

Per stabilire le formule di trasformazione delle coordinate, fissiamo un vettore x e siano rispettivamente x^i e y^i le sue coordinate rispetto alle due basi precedenti. Si ha:

$$x \,=\, x^i u_i \,=\, y^i v_i$$

da cui, utilizzando la definizione di matrice di trasformazione, si ottiene

$$x^i u_i \,=\, y^i m_i^j u_j$$

e poiché i vettori u_i sono indipendenti, i coefficienti di ogni vettore della base debbono coincidere, e dunque

$$x^j \,=\, y^i m_i^j$$

Si osserva subito che, mentre applicando la matrice di trasformazione si muta la base "vecchia" in quella "nuova", per le coordinate dei vettori accade l'opposto: le coordinate "nuove" vengono mutate in quelle vecchie. Per questa ragione, i vettori colonna vengono detti anche vettori *controvarianti*. Vedremo in breve come cambino le cose per i vettori riga.

1.3.2 Trasformazione della matrice associata ad un'applicazione lineare

Iniziamo con il caso più semplice: quello delle applicazioni a valori in $\mathbb{R}$, cioè gli elementi del duale. Sia dunque F una funzione lineare su X, e osserviamo che

$$F(x) \,=\, x^i F(u_i) \,=\, y^i F(v_i)$$

detti ora, come in precedenza, a e b i due vettori *riga* definiti da $a = (F(u_i))$ e $b = (F(v_i))$, segue

$$x^i a_i \,=\, y^i b_i$$

da cui, introducendo la formula di trasformazione delle coordinate, si ottiene

$$y^i m_i^j a_j \,=\, y^i b_i \qquad \Rightarrow \qquad y^i(m_i^j a_j - b_i) \,=\, 0$$

e poiché y è arbitrario, scelto $y_i = m_i^j a_j - b_i$ e ricordando che la norma è definita positiva, segue che $y = 0$ e dunque

$$b_i = m_i^j a_j$$

Come si sarà notato, i vettori riga che rappresentano gli elementi del duale di X non sono controvarianti come i vettori colonna, ma *covarianti*: la matrice di trasformazione del cambio di base muta il vettore "vecchio" a in quello "nuovo" b, come accade per le basi stesse.

Il passo successivo sarà di studiare gli effetti del cambio di base sulla matrice associata ad un'applicazione lineare.
Siano X e Y due spazi vettoriali e F un'applicazione lineare fra X e Y. Siano inoltre $u_1, u_2, ..., u_n$, $\tilde{u}_1, \tilde{u}_2, ..., \tilde{u}_n$, $v_1, v_2, ..., v_n$ e $\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_n$ rispettivamente due basi in X e due in Y. Siano infine m ed n rispettivamente le matrici di trasformazione relative ai due cambi di base in X e in Y, e a e b le due matrici che rappresentano l'applicazione F rispetto alle due coppie di basi (una di partenza e una d'arrivo). Si avrà:

$$F(x) \;=\; x^i a_i^j v_j \;=\; \xi^i b_i^j \tilde{v}_j$$

e ricordando che

$$x^i = m_j^i \xi^j \qquad\qquad \tilde{v}_i = n_i^j v_j \;\Rightarrow\; v_i \;=\; (n^{-1})_i^j \tilde{v}_j$$

si ha anche

$$\xi^i b_i^h \tilde{v}_h \;=\; x^i a_i^j v_j \;=\; \xi^k m_k^i a_i^j (n^{-1})_j^h \tilde{v}_h$$

da cui, per l'indipendenza di $\tilde{v}_h$, segue

$$\xi^i\, b_i^h \;=\; \xi^k m_k^i a_i^j (n^{-1})_j^h$$

e dunque, la matrice che rappresenta F rispetto alle due "nuove" basi, di partenza (rispetto alla quale è relativo il vettore delle coordinate ξ) e d'arrivo (costituita dai $\tilde{v}_j$), vale

$$b_k^h \;=\; m_k^i a_i^j (n^{-1})_j^h$$

e cioè, tenendo conto dell'ordine con cui avvengono i prodotti (su ξ operano prima M, poi A e infine N^{-1})

$$\tilde{A} = N^{-1} A M$$

Una questione rilevantissima nella quale vengono impiegate sistematicamente, sin dalla posizione stessa del problema, questi concetti, è la diagonalizzabilità. In tale caso, gli spazi X e Y coincidono e F, applicazione lineare da X in sè, è ciò che viene detto un endomorfismo. Inoltre, le basi di partenza e d'arrivo coincidono e dunque, dopo il cambio di base, la matrice di rappresentazione di F diventa

$$M^{-1} A M$$

ove M è la matrice di trasformazione associata al cambio di base.

1.3.3 Trasformazione della matrice associata ad un'applicazione bilineare

Concludiamo con l'analoga indagine per le funzioni bilineari, che consentirà di applicare il teorema di Sylvester allo studio del segno degli autovalori.
Ci limiteremo per semplicità a considerare il caso delle funzioni bilineari sulle coppie di vettori di X. Con gli stessi simboli della sezione precedente, si ha:

$$x \;=\; x^i u_i \;=\; \xi^i \tilde{u}_i, \qquad\qquad y \;=\; y^i u_i \;=\; \eta^i \tilde{u}_i$$

e anche

$$a(x, y) \;=\; x'_i y^j a_j^i \;=\; \xi'_i \eta^j \tilde{a}_j^i$$

Per stabilire il legame fra a e $\tilde{a}$, introduciamo la matrice di trasformazione m_j^i tale che

$$\tilde{u}_i \;=\; m_i^j u_j$$

e, ricordando che $(AB)' = B'A'$, si ha

$$x^i \;=\; m_j^i \xi^j \qquad \Rightarrow \qquad x'_i \;=\; \xi'_j m'^j_i$$

oltre che

$$y^i = m^i_k \eta^k$$

da cui, sostituendo nella prima delle due espressioni di a, ne segue

$$a(x, y) = \xi'_h m'^h_i \eta^k m^j_k a^i_j = \xi'_h \eta^k m'^h_i a^i_j m^j_k$$

da cui, infine, per l'unicità della matrice di rappresentazione,

$$\tilde{A} = M' A M$$

Nel caso in cui M sia una matrice *unitaria* (altrove detta anche *ortogonale*), e cioè le cui colonne formano una base ortonormale (come nelle rotazioni, ad esempio), si ha $M' = M^{-1}$, e dunque tanto le matrici di rappresentazione delle applicazioni lineari quanto quelle delle forme bilineari mutano allo stesso modo, cosa che può tornare utile nello studio del segno degli autovalori mediante il teorema di Sylvester.