

Note di Algebra Lineare

I Modulo

Luca Barbieri Viale

<http://users.unimi.it/barbieri/>

Dipartimento di Matematica
Università degli Studi di Milano

© 2003 Versione L^AT_EX

Avvertenza!

Le Note di Algebra Lineare si dividono in un I ed un II Modulo: raccolgono i **concetti essenziali** utilizzabili per un buon orientamento tra le nozioni impartite nei corsi canonici di base.

Queste pagine non costituiscono un testo compiuto ma solamente una sorta di organigramma di nozioni, definizioni e risultati, algoritmi e teoremi, problematiche e motivazioni.

Per la comprensione (ovvero la digestione) di queste materie prime è indispensabile molto esercizio (ovvero un buon addestramento alla produzione di acidi gastrici) e solamente l'esercizio vi renderà consapevoli (di quel che non avete digerito).

Nelle note a fondo pagina sono raccolti alcuni commenti ulteriori che intendono stimolare il lettore ad aguzzare l'ingegno, precisare e approfondire eventuali manchevolezze dell'autore.

Utilizzerò una *cornice* per indicare alcuni termini che vengono introdotti o utilizzati per la prima volta nel testo e che sono anche le parole chiave del vostro glossario.

Prerequisiti I Modulo

Assumo che il lettore abbia una certa domestichezza con i numeri naturali, interi, razionali e reali, equazioni e disequazioni, elementi di trigonometria, polinomi e radici. Assumo noti i rudimenti della teoria degli insiemi: prodotto cartesiano, corrispondenze, relazioni d'ordine e d'equivalenza, applicazioni (iniettive, surgettive e bigettive), immagini e controimmagini. In particolare, per le notazioni, si faccia riferimento al seguente paradigma.

Paradigma

Sia X un insieme e \mathcal{P} una certa proprietà. Un insieme


$$A \stackrel{\text{def}}{=} \{x \in X \mid x \text{ verifica la proprietà } \mathcal{P}\}$$

esiste sempre.¹ Il simbolo $\stackrel{\text{def}}{=}$ indica che l'insieme considerato verrà in seguito denotato A . Ad esempio, l'insieme $\emptyset \stackrel{\text{def}}{=} \{x \in X \mid x \neq x\}$ esiste e viene detto insieme vuoto. Le lettere \mathbb{Z} , \mathbb{Q} , \mathbb{R} e \mathbb{C} denotano gli insiemi dei numeri interi, razionali, reali e complessi. Osserviamo che

$$\emptyset = \{x \in \mathbb{Q} \mid x^2 = 2\} = \{x \in \mathbb{R} \mid x^2 = -1\}$$

dove ora il simbolo $=$ di uguaglianza significa che gli insiemi hanno gli stessi elementi² e quindi tali uguaglianze descrivono alcune peculiarità degli insiemi numerici considerati.

Un'applicazione tra insiemi verrà denotata $f : A \rightarrow B$ e la sua composizione con $g : B \rightarrow C$ verrà denotata $gf : A \rightarrow C$.

 **Attenzione!** Non abusare delle notazioni sopra esposte, in particolare dell'esistenza d'insiemi stravaganti.³

¹Per l'assioma di specificazione se esiste un insieme ne esiste sempre un altro che si ottiene dal primo specificandone gli elementi mediante una qualunque proprietà.

²Per l'assioma di estensione due insiemi sono uguali se hanno gli stessi elementi. Questo assioma garantisce che l'assioma di specificazione determina un unico insieme.

³Sia X un insieme e sia \mathcal{P} la proprietà di non appartenere a sé stessi. Dunque esiste l'insieme $A \stackrel{\text{def}}{=} \{x \in X \mid x \notin x\}$. Allora si ha: $y \in A \Leftrightarrow y \in X$ e inoltre $y \notin y$. Supponendo che $A \in A$ si ha che $A \in X$ ma pure che $A \notin A$: assurdo! Dunque $A \notin A$. Allora $A \notin X$, infatti: supponendo $A \in X$ e $A \notin A$ si ottiene che $A \in A$: assurdo! Ora l'insieme X

Programma I Modulo

In questa prima parte presentiamo lo studio dei sistemi lineari e del principale algoritmo per la loro soluzione. Introducendo la nozione di matrice, come sinonimo di tabella associata ad un sistema lineare, sviluppiamo una vera algebra, che risulta il *leit-motiv* dei nostri studi (anche nel II Modulo). Lo studio delle matrici associate ai sistemi da un lato e quello delle loro soluzioni dall'altro conducono al concetto di spazio vettoriale. Presenteremo, quindi, i tratti essenziali della teoria degli spazi vettoriali di dimensione finita.

1 Sistemi Lineari

Un sistema lineare nelle incognite X_1, X_2, \dots, X_n si scrive

$$S : \begin{cases} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n & = b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n & = b_2 \\ \dots & = \dots \\ \dots & = \dots \\ a_{m-11}X_1 + a_{m-12}X_2 + \dots + a_{m-1n}X_n & = b_{m-1} \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n & = b_m \end{cases} \quad (1)$$

dove i numeri a_{ij} si dicono coefficienti mentre i numeri b_i sono i corrispondenti termini noti.

- I numeri naturali $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$ sono detti indici. L'indice i fissato corrisponde alla equazione

$$a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n = b_i$$

e quindi il sistema S risulta avere m -equazioni. L'indice j fissato corrisponde alla incognita X_j .

è completamente arbitrario e a partire da questo insieme abbiamo costruito un insieme A che non appartiene a X . Se X fossero tutti gli insiemi si avrebbe così un insieme A che non appartiene a X : assurdo! Se ne deduce che tutti gli insiemi non sono un insieme. Per evitare tali inconvenienti si può introdurre il concetto di universo nel quale tutte le operazioni della teoria degli insiemi sono ammesse.

- Una soluzione di tale sistema è una n -upla (s_1, s_2, \dots, s_n) di numeri tale che ponendo $X_j = s_j$ in tutte le singole equazioni si ottengono i termini noti b_i corrispondenti. Bisogna dunque specificare dove si vogliono ottenere queste soluzioni (ovvero se gli s_j sono numeri razionali, reali, complessi, etc.) L'insieme delle soluzioni è

$$\text{Sol}(S) \stackrel{\text{def}}{=} \{(s_1, s_2, \dots, s_n) \text{ tale che è soluzione di } S\}.$$

Problema 1.1 *Trovare un metodo, possibilmente un algoritmo, per trovare le soluzioni di un sistema lineare ovvero determinare l'insieme $\text{Sol}(S)$ dato un sistema S .*

⊙ **Attenzione!** Un sistema S con n -incognite ed n -equazioni può avere un'unica soluzione come averne infinite come nessuna!

Se il sistema S non ha soluzioni allora $\text{Sol}(S)$ è vuoto; in questo caso basta che una singola equazione non ammetta soluzioni oppure può accadere che le singole equazioni ammettano soluzioni, ma il sistema sia incompatibile e quindi $\text{Sol}(S)$ risulti vuoto.

Operazioni elementari e sistemi equivalenti

Se consideriamo una singola equazione del sistema

$$a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n = b_i$$

possiamo moltiplicare per un numero non nullo $k \neq 0$ ottenendo

$$ka_{i1}X_1 + ka_{i2}X_2 + \dots + ka_{in}X_n = kb_i$$

senza modificare $\text{Sol}(S)$. Analogamente riordinando le equazioni di S non modifichiamo $\text{Sol}(S)$.

Definizione 1.2 *Diciamo che due sistemi S ed S' sono equivalenti se $\text{Sol}(S) = \text{Sol}(S')$.*

Si possono quindi individuare facilmente le seguenti operazioni su S dette *operazioni elementari* che producono sistemi equivalenti:

- moltiplicare una qualunque equazione di S per un numero non nullo,
- scambiare equazioni di S ,
- sostituire una equazione di S con quella ottenuta sommando ad essa un'altra equazione di S .⁴

Algoritmo di riduzione di Gauss

Utilizzando le operazioni elementari si può sempre semplificare un sistema lineare S , trasformandolo in un sistema S' equivalente e con un maggior numero di coefficienti nulli. I coefficienti e termini noti del sistema S si possono rappresentare sinteticamente mediante una tabella

$$\begin{array}{cccccc}
 a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\
 a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 a_{m-11} & a_{m-12} & a_{m-13} & \dots & a_{m-1n} & b_{m-1} \\
 a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & b_m
 \end{array} \tag{2}$$

La prima riga R_1 corrisponde alla prima equazione nel sistema S in (1), la seconda riga R_2 corrisponde alla seconda equazione, etc. Sarà comodo denotare C_1, C_2, \dots la prima, seconda, \dots colonna della tabella.

Passo 1

L'algoritmo⁵ di riduzione parte controllando se a_{11} è nullo:

- se $a_{11} = 0$ e inoltre $a_{k1} = 0$ con $k > 1$, ovvero la prima colonna C_1 della tabella (2) è nulla, allora controlla se a_{12} è nullo, etc. operando in seguito sulla nuova tabella ottenuta da (2) scartando la prima colonna;

⁴Se (s_1, s_2, \dots, s_n) è una soluzione di S , quindi una soluzione di ogni equazione E_i per $i = 1, \dots, m$, è anche una soluzione del sistema S' dove l' i -esima equazione E_i viene sostituita con $E_i + E_j$ per i e j fissati; e si ha che $\text{Sol}(S) = \text{Sol}(S')$ se $i \neq j$.

⁵Brevemente un algoritmo è una sequenza di procedure di calcolo che trasformano certi dati input in altri dati output.

- nel caso in cui $a_{11} = 0$ ma esiste $a_{k1} \neq 0$ con $k > 1$ allora opera lo scambio

$$R_1 \leftrightarrow R_k$$

e passa al caso seguente, nel quale $a_{11} \leftrightarrow a_{k1}$ non è nullo:

- nel caso $a_{11} \neq 0$ nella tabella (2) o in quelle ottenute come sopra, opera la sostituzione

$$R_k \rightarrow R_k - \frac{a_{k1}}{a_{11}} R_1$$

per ogni riga R_k con $k > 1$ della tabella (2); questo produce un sistema S' equivalente ad S la cui tabella è

a_{11}	a_{12}	a_{13}	\dots	a_{1n}	b_1
0	a'_{22}	a'_{23}	\dots	a'_{2n}	b'_2
\dots	\dots	\dots	\dots	\dots	\dots
\dots	\dots	\dots	\dots	\dots	\dots
0	a'_{m-12}	a'_{m-13}	\dots	a'_{m-1n}	b'_{m-1}
0	a'_{m2}	a'_{m3}	\dots	a'_{mn}	b'_m

(3)

Passo 2

Applica il Passo 1 alla tabella

a'_{22}	a'_{23}	\dots	a'_{2n}	b'_2
\dots	\dots	\dots	\dots	\dots
\dots	\dots	\dots	\dots	\dots
a'_{m-12}	a'_{m-13}	\dots	a'_{m-1n}	b'_{m-1}
a'_{m2}	a'_{m3}	\dots	a'_{mn}	b'_m

(4)

ottenuta dalla (3) estraendo la prima riga R_1 e la prima colonna e quindi si avrebbe che a'_{22} diventa a_{11} del Passo 1. Chiaramente molti altri coefficienti a'_{k2}, a'_{k3}, \dots possono risultare nulli.

Iterazione

La tabella (4) è più piccola della tabella (2) da cui siamo partiti. Ripetendo il Passo 1 a tabelle sempre più piccole il processo termina.

Fine

L'algoritmo termina producendo una tabella del tipo seguente

$$\begin{array}{cccccc}
 \star & \star & \star & \dots & \star & \star \\
 0 & 0 & \star & \dots & \star & \star \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 0 & 0 & 0 & \dots & \star & \star \\
 0 & 0 & 0 & \dots & 0 & 0
 \end{array} \tag{5}$$

dove \star sono numeri possibilmente⁶ non nulli che risultano dalle varie operazioni.

Il primo termine non nullo, da sinistra, in ogni riga di questa tabella, si dice pivot e corrisponde ad un coefficiente non nullo di una incognita del sistema che risulta alla fine dell'algoritmo. Ad esempio, se fosse risultato che $a_{11} \neq 0$ in (2) si avrebbe $a_{11} = \star$ nella (5). Nella prima equazione quindi si potrebbe facilmente esprimere X_1 in funzione delle rimanenti incognite. Un sistema con una tabella come in (5) si risolve subito, come vedremo in seguito.

Definizione 1.3 *Un sistema si dice ridotto se il primo coefficiente non nullo di ogni equazione è più a sinistra del primo coefficiente non nullo dell'equazione successiva.*

Soluzioni di un sistema ridotto

Supponiamo di partire da un sistema S come in (1) con n -incognite ed m -equazioni. Applicando l'algoritmo di riduzione si ottiene un sistema ridotto equivalente ad S . In tale sistema ridotto le equazioni significative o non nulle sono quelle che determinano le soluzioni. Quindi basta studiare le soluzioni di queste ultime per descrivere $\text{Sol}(S)$ e rispondere al Problema 1.1. In un sistema ridotto si possono verificare due casi.

Caso 1

L'ultima equazione significativa ha tutti i coefficienti nulli ma il termine noto è non nullo. Ovvero si ottiene una tabella

⁶In effetti la tabella (5) non contempla tutte i possibili risultati della riduzione; ad esempio, si può verificare che la \star del coefficiente nella penultima equazione sia nulla.

★	★	★	...	★	★
0	0	★	...	★	★
...
...
0	0	0	...	0	★
0	0	0	...	0	0

(6)

dalla quale risulta una equazione $0 = \star$ con \star un numero diverso da zero. In questo caso chiaramente il sistema **non** ha soluzioni.

Caso 2

Si hanno p -equazioni significative, ovvero p -righe non nulle nella tabella (5), nelle quali si trova almeno un coefficiente non nullo. Ovvero, nella tabella (5), si possono anche trovare diverse righe nulle ma non si ha il caso descritto nella tabella (6). Allora si hanno $n-p$ incognite libere. In effetti tali incognite sono parametri che descrivono al loro variare tutte le soluzioni del sistema. Si dice dunque che il sistema ha ∞^{n-p} soluzioni⁷.

Esempio

Ad esempio, la tabella potrebbe essere

★	★	★	★	★	★
0	0	★	★	★	★
0	0	0	★	★	★
0	0	0	0	0	0
0	0	0	0	0	0

In questo caso abbiamo 5 incognite e 3 equazioni significative ed il sistema ha ∞^2 soluzioni.

In generale, se abbiamo operato la riduzione ed m era il numero di equazioni si avrà che $p \leq m$. In particolare se $m = n$, ovvero abbiamo tante incognite quante equazioni, e si ha anche $p = n$ allora abbiamo un'unica soluzione del sistema.

⁷Questo è un tipico abuso di notazione.

Definizione 1.4 Diciamo caratteristica di un sistema ridotto il numero delle sue equazioni significative.

Si noti che per rispondere veramente al Problema 1.1 però bisogna adesso rispondere al seguente.

Problema 1.5 *Mostrare che un qualunque sistema ridotto S' equivalente ad S ha sempre la stessa caratteristica.*

Siccome la caratteristica è una proprietà della tabella dobbiamo ora affrontare seriamente il concetto di tabella.

2 Matrici

Una **matrice** è appunto una tabella che viene indicata sinteticamente (a_{ij}) con i e j indici di riga e colonna rispettivamente. I singoli numeri (reali, interi, etc.) a_{ij} sono anche dette **entrate** della matrice ed il **formato** è il numero di righe \times il numero di colonne⁸. Adotteremo la seguente notazione

$$M_{m,n}(\mathbb{R}) \stackrel{\text{def}}{=} \{\text{matrici } m \times n \text{ ad entrate in } \mathbb{R}\}$$

In una matrice $A = (a_{ij}) \in M_{m,n}(\mathbb{R})$ con $m \leq n$ i termini a_{ii} per $i = 1, \dots, m$ vengono detti **diagonale** principale. Una **matrice diagonale** è una matrice **quadrata** ovvero con $m = n$ tale che tutti gli elementi a_{ij} con $i \neq j$ sono nulli.

In corrispondenza del nostro sistema S in (1) abbiamo una matrice $A \stackrel{\text{def}}{=} (a_{ij})$ di formato $m \times n$ con entrate i coefficienti a_{ij} del sistema. Analogamente i termini noti $b \stackrel{\text{def}}{=} (b_i)$ sono una matrice colonna $m \times 1$ e le incognite $X \stackrel{\text{def}}{=} (X_j)$ sono una matrice colonna $n \times 1$. La matrice visualizzata nella tabella (2) si dice matrice **completa** del sistema S ed il suo formato è $m \times (n + 1)$.

Per rappresentare il sistema S mediante le matrici come segue

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_{n-1} \\ X_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_{m-1} \\ b_m \end{pmatrix}$$

ovvero molto più sinteticamente

$$AX = b \tag{7}$$

pensiamo le equazioni come risultato del prodotto **righe \times colonne**, riscrivendo la singola i -esima equazione

$$a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n = b_i$$

⁸Più precisamente, una matrice $m \times n$ di numeri reali è un'applicazione dall'insieme $\{1, \dots, m\} \times \{1, \dots, n\}$ a valori in \mathbb{R} .

come segue

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_{n-1} \\ X_n \end{pmatrix} = b_i$$

Questa operazione ha senso poichè il numero di colonne della matrice A coincide con il numero di righe della matrice colonna X .

Somma e prodotto di matrici

Possiamo utilizzare questa operazione di prodotto righe per colonne nel modo seguente.

Definizione 2.1 Sia $A = (a_{ij})$ matrice $m \times n$ e sia $B = (b_{ij})$ matrice $n \times l$. Il prodotto di matrici AB è la matrice $m \times l$ che al posto ij ha il risultato

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \dots \\ b_{n-1j} \\ b_{nj} \end{pmatrix} = (a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj})$$

dal prodotto righe per colonne della i -esima riga di A per la j -esima colonna di B .

⊙ **Attenzione!** Il prodotto di matrici non è commutativo! Ad esempio, è facile trovare matrici quadrate ($m = n = l$) tali che $AB \neq BA$.

Prodotto per uno scalare

Data una matrice $A \in M_{m,n}(\mathbb{R})$ si ha inoltre un'operazione naturale che associa alla matrice $A = (a_{ij})$ ed un numero $\lambda \in \mathbb{R}$ la matrice

$$\lambda A \stackrel{\text{def}}{=} (\lambda a_{ij})$$

Questa operazione risulta anche dal prodotto della matrice $m \times m$ seguente, dove λ compare sulla diagonale

$$\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

moltiplicando a sinistra della matrice A . In particolare, si ha che la matrice quadrata

$$I \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

detta matrice identica è tale che $IA = AI = A$ per una qualunque matrice quadrata A , dello stesso formato di I .

Ovviamente abbiamo anche una operazione naturale di somma.

Definizione 2.2 Siano $A = (a_{ij})$ e $B = (b_{ij})$ matrici $m \times n$. La somma di matrici $A + B$ è la matrice $m \times n$ seguente

$$A + B \stackrel{\text{def}}{=} (a_{ij} + b_{ij})$$

L'elemento neutro per la somma è la matrice nulla $m \times n$

$$0 \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Tale somma è ovviamente associativa $A+(B+C) = (A+B)+C$, commutativa $A + B = B + A$ ed esiste sempre l'opposto ovvero la matrice $-A$ tale che $A + (-A) = 0$.

Proprietà

Queste operazioni godono delle seguenti proprietà⁹:

- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$
- $(A + B)C = AC + BC$

dove si sottintende che le matrici A, B e C abbiano il formato opportuno affinché tali operazioni siano definite. Si noti che, in particolare, si hanno corrispondenti proprietà del prodotto per uno scalare: $\lambda(\mu A) = (\lambda\mu)A$, $\lambda(A+B) = \lambda A + \lambda B$ e $(\lambda + \mu)A = \lambda A + \mu A$.

Matrici invertibili

Sia $A \in M_{n,n}(\mathbb{R})$ una matrice quadrata non nulla. Si può facilmente vedere che **non** è sempre garantita l'esistenza dell'inversa moltiplicativa¹⁰ ovvero di una matrice quadrata A^{-1} tale che $A^{-1}A = AA^{-1} = I$. Diremo che una matrice A è invertibile se esiste tale *matrice inversa* A^{-1} .

Consideriamo nuovamente il nostro sistema (7) $AX = b$ e supponiamo che A sia invertibile. Allora, moltiplicando a sinistra abbiamo che

$$A^{-1}AX = A^{-1}b \quad \text{e quindi si ha che} \quad X = A^{-1}b$$

ovvero esiste un'unica soluzione! Ovvero tutte le equazioni sono significative ed il sistema che otteniamo mediante l'algoritmo di riduzione ha caratteristica massima. Deve quindi esserci un qualche legame tra l'algoritmo di riduzione ed il calcolo dell'inversa.

Osserviamo infine che se $A, B \in M_{n,n}(\mathbb{R})$ sono matrici invertibili anche il prodotto AB è invertibile e si ha

$$(AB)^{-1} = B^{-1}A^{-1} \tag{8}$$

Infatti si ha che

$$(B^{-1}A^{-1})(AB) = ((B^{-1}A^{-1})A)B = (B^{-1}(A^{-1}A))B = (B^{-1}I)B = I$$

e analogamente per il prodotto a destra.

⁹La verifica di queste proprietà non è difficile anche se abbastanza laboriosa.

¹⁰Se esiste è unica.

Matrici elementari

Siamo ora interessati ad individuare delle matrici che corrispondano alle operazioni elementari sul sistema (7) attraverso moltiplicazione sulla sinistra. Diremo che le seguenti matrici $n \times n$, ottenute dall'identità, sono *matrici elementari*

- I la matrice identica $n \times n$,
- E_{ij} la matrice ottenuta da I scambiando la riga R_i con R_j con $i \neq j$, ad esempio

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

- $E_i(\lambda)$ la matrice ottenuta da I mettendo il numero $\lambda \neq 0$ al posto dell'entrata ii sulla diagonale, ad esempio

$$E_2(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

- $E_{ij}(\lambda)$ la matrice ottenuta da I mettendo il numero $\lambda \neq 0$ al posto dell'entrata ij con $i \neq j$, ad esempio

$$E_{21}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Data una matrice $A \in M_{n,n}(\mathbb{R})$ si ottengono le seguenti operazioni elementari su A come segue:

- $E_{ij}A$ è la matrice A con la riga R_i scambiata con la riga R_j , operazione elementare $R_i \leftrightarrow R_j$

- $E_i(\lambda)A$ è la matrice A con la riga R_i sostituita con λR_i , operazione elementare $R_i \rightarrow \lambda R_i$
- $E_{ij}(\lambda)A$ è la matrice A con la riga R_i sostituita con $R_i + \lambda R_j$, operazione elementare $R_i \rightarrow R_i + \lambda R_j$

Si calcolano facilmente le inverse delle matrici elementari, mediante le operazioni elementari inverse:

- $E_{ij}^{-1} = E_{ij}$ in quanto $R_i \leftrightarrow R_j$ su E_{ij} riproduce I
- $E_i(\lambda)^{-1} = E_i(\lambda^{-1})$ in quanto $R_i \rightarrow \lambda^{-1}R_i$ su $E_i(\lambda)$ riproduce I
- $E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda)$ in quanto $R_i \rightarrow R_i - \lambda R_j$ su $E_{ij}(\lambda)$ produce I

Riduzione di matrici

Sia $A \in M_{n,n}(\mathbb{R})$. Riducendo A mediante matrici elementari E_1, \dots, E_t si ottiene una matrice ridotta per righe ovvero una matrice $J = (r_{ij})$ tale che se $r_{ij} \neq 0$ e $r_{ik} = 0$ per ogni $k < j$ allora anche $r_{hk} = 0$ ogni volta che $h \geq i$ e $k < j$.

Proseguendo la riduzione in modo da eliminare i termini non nulli anche sopra i pivot si può ottenere una matrice totalmente ridotta per righe ovvero una matrice ridotta $J = (r_{ij})$ tale che se $r_{ij} \neq 0$ e $r_{ik} = 0$ per ogni $k < j$ allora $r_{ij} = 1$ e per ogni $h \neq i$ si ha $r_{hj} = 0$. In definitiva si ottiene la seguente formula

$$E_t E_{t-1} \cdots E_2 E_1 A = J \quad (9)$$

dove E_i per $i = 1, \dots, t$ sono matrici elementari e J è totalmente ridotta per righe.

Se $A \in M_{n,n}(\mathbb{R})$ fosse invertibile allora $J = I$ è la matrice identica. Infatti la matrice J ha come primo elemento non nullo di ogni riga 1 e tutti gli altri elementi delle colonne che contengono questi 1 sono nulle. Siccome J non può avere una riga nulla essendo prodotto di matrici invertibili, per la (9), deve essere invertibile¹¹ e quindi $J = I$. Questa argomentazione consente di caratterizzare le matrici invertibili come segue.

¹¹Una matrice con una riga nulla non è invertibile. Infatti anche nel prodotto con una qualsiasi matrice risulterà una riga nulla.

Teorema 2.3 *Sia A una matrice quadrata $n \times n$. Le seguenti affermazioni sono equivalenti:*

1. A è invertibile,
2. A si riduce ad una matrice con n righe non nulle,
3. A si riduce totalmente ad I ,
4. A è un prodotto di matrici elementari.

Dimostrazione:

[1. \Rightarrow 2.] Questo è garantito dall'algorithmo di riduzione; come abbiamo già osservato sopra, la matrice ridotta è invertibile e quindi non può avere una riga nulla.

[2. \Rightarrow 3.] Siccome A è quadrata e la matrice ottenuta riducendo totalmente A non ha righe nulle, allora i pivot (= 1) sono tutti situati sulla diagonale.

[3. \Rightarrow 4.] Nella formula (9) $J = I$, quindi moltiplicando a sinistra per l'inversa del prodotto delle matrici elementari si ottiene

$$A = (E_t E_{t-1} \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_t^{-1} E_{t-1}^{-1}$$

e ora basta ricordare che le inverse delle matrici elementari sono ancora matrici elementari.

[4. \Rightarrow 1.] Se A è prodotto di matrici elementari allora A è invertibile in quanto prodotto di matrici invertibili.

◊

La formula (9) ed il Teorema 2.3 oltre caratterizzare una matrice invertibile forniscono anche un metodo per calcolare l'inversa.

Algoritmo per il calcolo della matrice inversa

Se A è invertibile allora, per la formula (9)

$$E_t E_{t-1} \cdots E_2 E_1 A = I$$

e moltiplicando a destra per A^{-1} si ha

$$E_t E_{t-1} \cdots E_2 E_1 I = A^{-1}$$

ovvero eseguendo su I le stesse operazioni elementari che riducono totalmente A si ottiene A^{-1} .

Quindi affiancando I alla matrice A ed eseguendo la riduzione anche su I giungiamo a ridurre A ad I e I ad A^{-1} .

Supponendo di voler risolvere il nostro sistema (7) con la matrice dei coefficienti A invertibile possiamo quindi calcolare l'inversa mediante una riduzione totale di A .¹² Questo metodo ha però lo svantaggio di non essere ottimale dal punto di vista calcolativo.

Fattorizzazione LU

Siamo interessati a scrivere una matrice A invertibile come prodotto di matrici con molti zeri. A tal fine si può ridurre A ottenendo

$$E_t E_{t-1} \cdots E_2 E_1 A \stackrel{\text{def}}{=} U$$

ovvero U è la matrice così ridotta (non totalmente!). Supponiamo inoltre che A si possa ridurre **senza scambi** di righe. Allora le matrici elementari E_i qui sopra utilizzate per la riduzione sono tutte del tipo $E_{ij}(\lambda)$ con $i > j$ oppure $E_i(\lambda)$.

Osserviamo che la matrice U è *triangolare superiore* ovvero ha entrate nulle **sotto** la diagonale mentre ogni matrice $E_{ij}(\lambda)$ con $i > j$ è *triangolare inferiore* ovvero ha entrate nulle **sopra** la diagonale ed infine le $E_i(\lambda)$ sono diagonali. Inoltre il prodotto di queste matrici triangolari inferiori è ancora triangolare inferiore e così le loro inverse. Ponendo

$$L^{-1} \stackrel{\text{def}}{=} E_t E_{t-1} \cdots E_2 E_1$$

¹²Si noti che mentre A^{-1} è unica la riduzione totale di A ad I non è unica.

per quanto detto otteniamo una fattorizzazione¹³

$$A = LU \quad (10)$$

dove la matrice triangolare inferiore L si calcola subito (provare per credere!) come prodotto delle inverse delle matrici elementari.

Nel caso in cui la matrice A non si riduca senza scambi di righe allora si possono operare gli scambi di righe su A che permettono poi di ridurre senza scambi. Si moltiplica A a sinistra per matrici elementari del tipo E_{ij} fino ad ottenere una matrice di permutazione P , prodotto appunto delle matrici di scambio, e una fattorizzazione

$$PA = LU \quad (11)$$

La matrice P è una matrice ottenuta dalla matrice identica facendo una permutazione di righe e così la matrice PA differisce da A solamente per l'ordine delle righe.

Infine, per risolvere un sistema

$$AX = b$$

come in (7), vediamo ora come si può impiegare la fattorizzazione (10). Sostituendo si ottiene

$$LUX = b$$

e ponendo

$$UX \stackrel{\text{def}}{=} Y$$

si risolve il sistema

$$LY = b$$

Siccome L è triangolare inferiore il sistema è subito risolto (anche se L non è ridotta!) ed Y determinata. Poi si risolve $UX = Y$ come al solito. Un vantaggio di questo metodo risulta dal fatto che una volta determinata la fattorizzazione della sola A ogni sistema con A come matrice dei coefficienti si risolve subito.

¹³Tale fattorizzazione risulta anche unica se ci si restringe a matrici L aventi sulla diagonale tutti 1.

Determinanti

Si vuole ora determinare l'invertibilità di una matrice quadrata A prima di procedere alla riduzione. Questo numero che determina l'invertibilità di una matrice sarà detto *determinante* di una matrice.

Caso 2×2

Se il nostro sistema $AX = b$ ha matrice dei coefficienti

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

e supponiamo $a_{11} \neq 0$ allora operando il passo di riduzione

$$R_2 \rightarrow R_2 - \frac{a_{21}}{a_{11}}R_1$$

otteniamo la matrice ridotta

$$J = \begin{pmatrix} a_{11} & a_{12} \\ 0 & \star \end{pmatrix}$$

dove

$$\star = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$

La matrice A è invertibile se e solo se anche $\star \neq 0$. Ponendo

$$\det(A) \stackrel{\text{def}}{=} a_{11}a_{22} - a_{12}a_{21}$$

è facile dedurre dal Teorema 2.3 la seguente equivalenza:

$$\det(A) \neq 0 \Leftrightarrow A \text{ è invertibile!}$$

Infatti, se $a_{11} \neq 0$ si ha che $a_{11}\star = \det(A)$ e se $a_{11} = 0$ allora si ha che $\det(A) = -a_{21}a_{12}$ e quindi, operato lo scambio $R_1 \leftrightarrow R_2$, la matrice ridotta non ha righe nulle se e solo se ha come termini pivotali proprio a_{21} e a_{12} .

Iterazione

Osserviamo che una matrice $A = (a) 1 \times 1$ è invertibile se e solo se $a \neq 0$ e poniamo dunque $\det(A) = a$. Quindi nella nostra definizione di determinante di matrici 2×2 è implicita la possibile definizione per matrici $n \times n$ con $n \geq 3$.

Infatti, data una matrice $A = (a_{ij})$ $n \times n$ chiamiamo *complemento algebrico* A_{ij} di un elemento a_{ij} il numero che si ottiene dal determinante della matrice $(n-1) \times (n-1)$ ottenuta da A sopprimendo la riga R_i e la colonna C_j moltiplicato per $(-1)^{i+j}$.

Poniamo quindi¹⁴

$$\det(A) \stackrel{\text{def}}{=} a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$

Esempio

Si noti che una matrice diagonale

$$\Delta = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

ha determinante

$$\det(\Delta) = a_{11} \cdots a_{nn}$$

Infatti $\det(\Delta) = a_{11}A_{11}$ e per induzione sul formato della matrice si conclude.¹⁵

Determinanti e operazioni elementari

Si calcolano facilmente i determinanti delle matrici elementari:

- $\det(E_{ij}) = -1$
- $\det(E_i(\lambda)) = \lambda$
- $\det(E_{ij}(\lambda)) = 1$

Una noiosa verifica, per induzione, permette anche di dedurre il comportamento del determinante rispetto alla riduzione.

¹⁴Questo si dice anche sviluppo secondo Laplace del determinante. Esistono altri svariati modi per calcolare il determinante.

¹⁵Analogamente, è facile calcolare il determinante di una matrice triangolare inferiore.

Lemma 2.4 *Sia $A \in M_{nn}(\mathbb{R})$ ed E una matrice elementare. Si ha che*

$$\det(EA) = \det(E) \det(A)$$

Si deduce dunque che:

- scambiando due righe il determinante cambia segno
- moltiplicando una riga per un numero λ il determinante viene moltiplicato per λ
- sommando ad una riga un'altra moltiplicata per λ il determinante non cambia.

Determinanti e matrici inverse

Possiamo ora dare la seguente caratterizzazione delle matrici invertibili, da aggiungere alla lista del Teorema 2.3.

Proposizione 2.5 *Sia $A \in M_{nn}(\mathbb{R})$. Si ha che $\det(A) \neq 0$ se e solo se A è invertibile.*

Dimostrazione: Se A è invertibile allora è prodotto di matrici elementari e applicando il Lemma 2.4 si ottiene $\det(A) \neq 0$. Viceversa, dalla (9), una riduzione totale di A , applicando il Lemma 2.4 si ha che $\det(J) \neq 0$ quindi $J = I$. ◊

Determinanti e prodotti di matrici

Siano $A, B \in M_{nn}(\mathbb{R})$ e sia AB la matrice prodotto righe per colonne. Possiamo ora dedurre la seguente formula:

$$\det(AB) = \det(A) \det(B) \tag{12}$$

Infatti, se A è invertibile allora è prodotto di matrici elementari e la formula si deduce dal Lemma 2.4 iterato. Se A non è invertibile bisogna verificare che $\det(AB) = 0$. Utilizzando la (9) si ha che $E_t \cdots E_1 A = J$ dove J è ridotta e ha almeno una riga di zeri. Ora $AB = E_1^{-1} \cdots E_t^{-1} JB$ e JB ha una riga di zeri. Quindi $\det(AB) = \det(E_1^{-1}) \cdots \det(E_t^{-1}) \det(JB) = 0$ poichè abbiamo

già osservato che una matrice con una riga di zeri non è invertibile e per la Proposizione 2 deve avere determinante nullo.

Un immediato corollario della formula (12) è che

$$\det(A^{-1}) = \det(A)^{-1}$$

Teoremi di Laplace per righe

Il I Teorema di Laplace dice che possiamo calcolare il determinante sviluppando rispetto ad una qualsiasi riga. Il II Teorema di Laplace ha come sua applicazione una formula per l'inversa, si veda (14) più avanti.

Teorema 2.6 (I e II Teorema di Laplace) *Sia A una matrice $n \times n$. Sia i un indice fissato ($1 \leq i \leq n$) si ha*

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

Sia j un'altro indice fissato ($1 \leq i, j \leq n$ e $i \neq j$) si ha

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0$$

Queste formule sono facilmente deducibili dal calcolo del determinante della matrice $E_{1i}A$ e da quello della matrice ottenuta da A sostituendo R_j con R_i , ripetutamente.

Il III Teorema di Laplace è una generalizzazione del primo a più righe e permette un facile calcolo del determinante di matrici a blocchi.

Matrici trasposte

Data una matrice $A = (a_{ij})$ si definisce la matrice trasposta $A^T \stackrel{\text{def}}{=} (a_{ji})$ quella ottenuta da A scambiando righe con colonne ovvero le colonne di A^T sono le righe di A . Si vede facilmente che:

- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Una matrice tale che $A^T = A$ si dice simmetrica ad esempio ogni matrice diagonale è simmetrica. Si vede subito che la matrice di scambio è simmetrica e inoltre:

- $E_{ij}^T = E_{ij}$
- $E_i(\lambda)^T = E_i(\lambda)$
- $E_{ij}(\lambda)^T = E_{ji}(\lambda)$

Dunque i determinanti delle matrici elementari coincidono con il determinante delle loro trasposte. In generale, se la matrice A è quadrata:

$$\det(A) = \det(A^T) \quad (13)$$

Infatti, se A è invertibile allora è prodotto di matrici elementari $A = E_t \cdots E_1$ e si ha $\det(A^T) = \det(E_1^T \cdots E_t^T) = \det(E_1) \cdots \det(E_t) = \det(A)$. Se A non è invertibile neanche A^T è invertibile. Infatti si verifica facilmente dalla definizione d'inversa che

$$(A^T)^{-1} = (A^{-1})^T$$

Dalla formula (13) seguono i Teoremi di Laplace per colonne ad esempio il determinante di una matrice si può calcolare anche sviluppando rispetto ad una qualsiasi colonna. Infine, sia $C \stackrel{\text{def}}{=} (A_{ij})$ la matrice costituita dai complementi algebrici di una matrice A invertibile. Si ha che:

$$A^{-1} = \frac{C^T}{\det(A)} \quad (14)$$

Infatti, come si verifica immediatamente, per il I e II Teorema di Laplace facendo il prodotto con A , si ottiene l'identica.

Caratteristica

Al fine di rispondere al Problema 1.5 e quindi risolvere i sistemi con matrici dei coefficienti non invertibili o non quadrate prendiamo ora in esame sottomatrici quadrate invertibili. Sia A una matrice $m \times n$. Si dice minore di ordine p di A il determinante di una qualunque sottomatrice quadrata $p \times p$ contenuta in A .

Definizione 2.7 Diciamo che una matrice A $m \times n$ ha caratteristica o rango p se

- esiste un minore di ordine p non nullo e
- tutti i minori (se esistono!) di ordine $p + 1$ sono nulli.

Denotiamo $\rho(A) = p$, in questo caso, la caratteristica di A .

Ad esempio, una matrice quadrata A $n \times n$ ha rango massimo n se e solo se A è invertibile.

Se A è una matrice $m \times n$ ed E è una qualunque matrice elementare abbiamo che

$$\rho(EA) = \rho(A)$$

Questo segue da un'analisi caso per caso delle operazioni elementari corrispondenti su A e dal fatto che la non nullità dei minori di A si mantiene, come abbiamo già osservato (ad esempio utilizzando (12)). Si ricava anche facilmente, dalla formula (13) che

$$\rho(A) = \rho(A^T)$$

Possiamo ora enunciare il seguente:

Teorema 2.8 (Rouché-Capelli) *Sia A una matrice $m \times n$ e sia $AX = b$ un sistema in n -incognite, come in (7). Sia $(A \mid b)$ la matrice $m \times (n + 1)$ completa del sistema. Si ha che il sistema ammette soluzioni se e solo se $\rho(A) = \rho(A \mid b)$. In tal caso, il numero di equazioni significative di un qualunque sistema ridotto equivalente è uguale alla caratteristica di A .*

La dimostrazione, tenuto conto di quanto detto sopra, segue dall'algoritmo di riduzione. Infatti, la caratteristica è invariante per operazioni elementari e inoltre si ha che una matrice ridotta per righe ha caratteristica p se e solo se ha p -righe non nulle.

Teorema di Kronecker

In una matrice A con m -righe diciamo che una riga R_k è ottenibile come combinazione lineare delle restanti righe se si può esprimere come segue

$$R_k = \lambda_1 R_1 + \cdots + \widehat{R_k} + \cdots + \lambda_m R_m$$

dove $\widehat{R_k}$ indica che nella somma è omessa la riga k -esima e λ_i per $i = 1, \dots, m$ sono numeri.¹⁶

Se A è una matrice $m \times n$ e $\rho(A) = p$ si può vedere che (per il Teorema 2.8!):

¹⁶Se A è quadrata e questo accade allora $\det(A) = 0$.

- se $p < m$ allora esiste una riga che è combinazione lineare delle rimanenti
- se $p < n$ allora esiste una colonna che è combinazione lineare delle rimanenti.

Questa osservazione dimostra il seguente:

Teorema 2.9 (Kronecker) *Per una matrice A $m \times n$ si ha che $\rho(A) = p$ se e solo se*

- *esiste un minore di ordine p non nullo, corrispondente ad una sottomatrice M e*
- *tutti i minori (se esistono!) di ordine $p+1$ corrispondenti a sottomatrici contenenti M sono nulli.*

Basta dimostrare che queste ipotesi su A implicano che $\rho(A) = p$. E infatti le righe fuori da M sono dunque combinazioni lineari delle righe in M e dopo una opportuna riduzione queste righe si annullano.

Anche se questo teorema permette una sostanziale diminuzione dei calcoli volti a trovare la caratteristica di una matrice, il metodo della riduzione risulta comunque vantaggioso.

3 Spazi Vettoriali

Siano X' ed X'' due soluzioni di un sistema $\boxed{\text{omogeneo}}$ $AX = 0$. Si noti che una loro combinazione lineare (come colonne) $\lambda'X' + \lambda''X''$ è ancora una soluzione. Infatti $A(\lambda'X' + \lambda''X'') = \lambda'AX' + \lambda''AX'' = 0$. Per un sistema omogeneo S in n -incognite si ha dunque che l'insieme delle soluzioni $\text{Sol}(S) \subseteq \mathbb{R}^n$ è chiuso rispetto alle operazioni di somma e prodotto per uno scalare. Inoltre, questa proprietà, come l'algoritmo di riduzione di Gauss, non necessita dell'insieme numerico \mathbb{R} . Ad esempio, se la matrice A ha entrate numeri razionali allora le soluzioni del sistema possono essere soluzioni razionali, ovvero $\text{Sol}(S) \subseteq \mathbb{Q}^n$.

Questi fatti individuano peculiarità dell'insieme delle soluzioni di sistemi omogenei che sono del tutto analoghe alle proprietà già descritte per le matrici e che vogliamo ora isolare.

Denotiamo con k un $\boxed{\text{campo degli scalari}}$ ad esempio \mathbb{Q}, \mathbb{R} oppure \mathbb{C} i numeri razionali, reali o complessi.

Definizione 3.1 *Sia V un insieme con due operazioni:*

- *la somma di due elementi $u, v \in V$ denotata $u + v$ e*
- *il prodotto di un elemento $v \in V$ per uno scalare $\lambda \in k$ che denotiamo λv .*

Se la somma è commutativa, associativa, esistono gli opposti ed un elemento neutro che denotiamo 0 ed il prodotto per uno scalare è associativo, distributivo rispetto alle somme (anche in k) e tale che $1v = v$ per ogni $v \in V$ allora diciamo che V è un $\boxed{k\text{-spazio vettoriale}}$.

Se V è un k -spazio vettoriale diremo che gli elementi di V , $v \in V$, sono $\boxed{\text{vettori}}$.

Un sottoinsieme $W \subseteq V$ è un $\boxed{\text{sottospazio}}$ se è chiuso rispetto alla somma ed il prodotto per uno scalare in V ovvero:

- se $u, v \in W$ allora $u + v \in W$ e
- se $v \in W$ e $\lambda \in k$ allora $\lambda v \in W$.

Diremo che dati v_1, \dots, v_n vettori di V il vettore v

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

è combinazione lineare di v_1, \dots, v_n . Denotiamo

$$L\{v_1, \dots, v_n\} \stackrel{\text{def}}{=} \{\lambda_1 v_1 + \dots + \lambda_n v_n \mid \lambda_1, \dots, \lambda_n \in k\}$$

il sottospazio generato da v_1, \dots, v_n .

Quando per un sottoinsieme $W \subseteq V$ risulta che $W = L\{w_1, \dots, w_r, \dots\}$ per certi vettori $w_1, \dots, w_r, \dots \in W$ diciamo che w_1, \dots, w_r, \dots è una (tra le tante!) successione di generatori per W .

Supponiamo ora di avere due sottospazi $V_0, V_1 \subset V$ di un k -spazio vettoriale V . Si vede facilmente che la loro intersezione insiemistica $V_0 \cap V_1$ è ancora un sottospazio ma l'unione insiemistica $V_0 \cup V_1$ non è in generale un sottospazio! Definiamo dunque lo spazio somma $V_0 + V_1$ come il più piccolo sottospazio di V che contiene V_0 e V_1 . Insiemeisticamente si ha che

$$V_0 + V_1 \stackrel{\text{def}}{=} \{v_0 + v_1 \mid v_0 \in V_0, v_1 \in V_1\}$$

Se $V_0 = L\{u_1, \dots, u_r\}$ e $V_1 = L\{v_1, \dots, v_s\}$ si ha dunque che

$$V_0 + V_1 = L\{u_1, \dots, u_r, v_1, \dots, v_s\}$$

Esempi

L'insieme delle matrici $M_{m,n}(k)$ con le operazioni di somma e prodotto per uno scalare è un k -spazio vettoriale. In particolare si ha che

$$k^n = M_{1,n}(k)$$

il k -spazio dei vettori riga¹⁷ (analogamente lo spazio dei vettori colonna è $M_{m,1}(k)$). Come già osservato si ha che il sottoinsieme delle soluzioni di un sistema omogeneo è un sottospazio

$$W = \text{Sol}(S) \subseteq k^n$$

¹⁷In particolare, se $k = \mathbb{C}$ allora \mathbb{C}^n è un \mathbb{C} spazio vettoriale. Siccome $\mathbb{R} \subset \mathbb{C}$ si ha che \mathbb{C} (e dunque \mathbb{C}^n) è anche un \mathbb{R} -spazio vettoriale e inoltre $\mathbb{C} = L\{1, i\}$. Analogamente $\mathbb{Q} \subset \mathbb{R}$ e quindi \mathbb{R} è anche un \mathbb{Q} -spazio vettoriale ma non esiste una successione finita di generatori per \mathbb{R} come \mathbb{Q} -spazio vettoriale.

il sottospazio delle soluzioni di S .

L'insieme $k[X]$ dei polinomi in una variabile a coefficienti in k è un k -spazio vettoriale con le operazioni di somma di polinomi e prodotto per uno scalare. Se $k = \mathbb{R}$ consideriamo solitamente $\mathbb{R}[X]$ come sottoinsieme dell'insieme delle funzioni (continue) di variabile reale $\mathcal{C}(\mathbb{R})$. L'insieme $\mathcal{C}(\mathbb{R})$ è un \mathbb{R} -spazio vettoriale mediante le operazioni di somma di funzioni e prodotto per una costante ed $\mathbb{R}[X]$ è un suo sottospazio. Si ha che la successione infinita $1, X, \dots, X^n, \dots$ genera $\mathbb{R}[X]$ come \mathbb{R} -spazio vettoriale ovvero si ha

$$\mathbb{R}[X] = L\{1, X, \dots, X^n, \dots\}$$

Si noti che non è possibile estrarre da tale successione una successione finita che generi tutto $\mathbb{R}[X]$ ovvero fissato $n > 0$ a piacere il sottospazio

$$\mathbb{R}_n[X] \stackrel{\text{def}}{=} L\{1, X, \dots, X^n\} \subset \mathbb{R}[X]$$

è il sottospazio dei polinomi di grado $\leq n$.

Da questo esempio si deduce che non esiste sempre un successione finita di generatori per uno spazio vettoriale arbitrario.

Problema 3.2 *Ogni successione finita di generatori ha sempre lo stesso numero di vettori? Cos'è la dimensione di uno spazio vettoriale?*

Basi

Supponiamo che V sia un k -spazio vettoriale e che v_1, \dots, v_n sia una successione finita di vettori che generano tutto V : diciamo allora che V è di *tipo finito* su k .

Diciamo che i vettori v_1, \dots, v_n sono *linearmente dipendenti* se esistono scalari $\lambda_1, \dots, \lambda_n$ **non tutti nulli** in k tali che

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

Diremo che v_1, \dots, v_n sono *linearmente indipendenti* se **non** sono linearmente dipendenti ovvero qualora si supponga che

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

allora si ha che $\lambda_1 = \dots = \lambda_n = 0$.

Si noti che i vettori v_1, \dots, v_n sono linearmente dipendenti se e solo se ne esiste uno tra essi che è combinazione lineare dei restanti. Infatti, se $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ ed esiste $\lambda_k \neq 0$ allora si ottiene v_k come combinazione lineare $v_k = -\lambda_k^{-1}(\lambda_1 v_1 + \dots + \widehat{v_k} \dots + \lambda_n v_n)$ dove $\widehat{v_k}$ indica che si è ommesso v_k nella somma. Viceversa se un certo v_k è combinazione lineare dei restanti allora si ha che v_1, \dots, v_n sono linearmente dipendenti.

Sia $V = L\{v_1, \dots, v_n\}$ un k -spazio di tipo finito. Se i v_1, \dots, v_n sono linearmente dipendenti allora esiste un certo vettore, supponiamo sia v_n , che è combinazione lineare dei restanti e si ha ovviamente che

$$V = L\{v_1, \dots, v_n\} = L\{v_1, \dots, v_{n-1}\}$$

Scartando successivamente i vettori linearmente dipendenti si può sempre giungere ad una successione di generatori costituita da vettori linearmente indipendenti!

Definizione 3.3 Una successione v_1, \dots, v_n di vettori linearmente indipendenti che generano uno spazio vettoriale V su k si dice base.

Teorema 3.4 Tutte le basi di uno spazio vettoriale (di tipo finito) sono costituite dallo stesso numero di vettori.

Dimostrazione: Sia $V = L\{v_1, \dots, v_n\}$. Mostriamo che $n + 1$ -vettori qualunque di V , siano w_1, \dots, w_{n+1} , sono sempre linearmente dipendenti. Scriviamo $w_i = a_{i1}v_1 + \dots + a_{in}v_n$ come combinazioni lineari con $a_{ij} \in k$. La matrice $A \stackrel{\text{def}}{=} (a_{ij})$ è una matrice $(n + 1) \times n$ ed il sistema omogeneo con n -equazioni

$$A^T X = 0$$

nelle incognite $X^T = (X_1, \dots, X_{n+1})$ ha sicuramente una soluzione non nulla (e quindi infinite!) per il Teorema 2.8. Sia $(\lambda_1, \dots, \lambda_{n+1})$ tale soluzione. Calcoliamo

$$\begin{aligned} & \lambda_1 w_1 + \dots + \lambda_{n+1} w_{n+1} = \\ & = \lambda_1 (a_{11}v_1 + \dots + a_{1n}v_n) + \dots + \lambda_{n+1} (a_{n+11}v_1 + \dots + a_{n+1n}v_n) = \\ & (\lambda_1 a_{11} + \dots + \lambda_{n+1} a_{n+11})v_1 + \dots + (\lambda_1 a_{1n} + \dots + \lambda_{n+1} a_{n+1n})v_n = 0 \end{aligned}$$

infatti $\lambda_1 a_{1i} + \dots + \lambda_{n+1} a_{n+1i} = 0$ per ogni $i = 1, \dots, n$ in quanto $(\lambda_1, \dots, \lambda_{n+1})$ è una soluzione del suddetto sistema. Quindi w_1, \dots, w_{n+1} sono linearmente dipendenti.

Supponiamo ora che v_1, \dots, v_n siano linearmente indipendenti ovvero fissiamo una base di V . Supponiamo ora che u_1, \dots, u_{n-1} generino V ovvero $V = L\{u_1, \dots, u_{n-1}\}$. Applicando il precedente ragionamento a v_1, \dots, v_n si avrebbe che sono dipendenti, contro l'ipotesi, quindi V non può esser generato con $n - 1$ -vettori, da cui la tesi. \odot

Possiamo ora definire la dimensione di uno spazio vettoriale.

Dimensione

Un k -spazio vettoriale (di tipo finito) V ha *dimensione* n se una (e quindi ogni!) sua base è costituita da n -vettori. In questo caso indichiamo

$$\dim_k(V) = n$$

la sua dimensione. Ad esempio, si vede facilmente che

$$\dim_k(M_{m,n}(k)) = mn$$

infatti esiste sempre una *base canonica* fatta dalle matrici e_{ij} che hanno 1 al posto ij e zero altrimenti. In particolare $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$, come ci si aspetta.¹⁸

Possiamo ora elencare alcune proprietà relative a basi e dimensione:

- $\dim_k(V) = n$ se e solo se n è il numero **massimo** di vettori linearmente indipendenti di V
- $\dim_k(V) = n$ se e solo se n è il numero **minimo** di vettori che generano V , inoltre
- se $W \subseteq V$ è un sottospazio allora $\dim_k(W) \leq \dim_k(V)$ e
- se $\dim_k(W) = \dim_k(V)$ allora $W = V$.

¹⁸La dimensione dipende anche da k : si ha che $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ ma $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.

Coordinate

Per lo spazio k^n la base canonica di n -vettori, che denotiamo e_1, \dots, e_n , è tale che il vettore riga e_i ha 1 al posto i -esimo. Un vettore generico $x = (x_1, \dots, x_n)$ si scrive come combinazione lineare

$$x = x_1 e_1 + \dots + x_n e_n$$

chiaramente in modo unico! Se v_1, \dots, v_n è una base di uno spazio vettoriale V allora ogni vettore $v \in V$ si scrive **in modo unico** come combinazione lineare dei vettori della base. Infatti, se

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n = \lambda'_1 v_1 + \dots + \lambda'_n v_n$$

allora

$$(\lambda_1 - \lambda'_1)v_1 + \dots + (\lambda_n - \lambda'_n)v_n = 0$$

ma siccome v_1, \dots, v_n sono indipendenti si ha che $\lambda_i = \lambda'_i$ per $i = 1, \dots, n$.

Se indichiamo sinteticamente con $B : v_1, \dots, v_n$ la base fissata chiameremo *coordinate* di v rispetto a B il vettore colonna

$$v_B \stackrel{\text{def}}{=} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

se $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Tramite questa osservazione possiamo sempre costruire un'applicazione bigettiva

$$\varphi_B : V \rightarrow k^n \quad v \mapsto \varphi_B(v) \stackrel{\text{def}}{=} (\lambda_1, \dots, \lambda_n)$$

tra un k -spazio vettoriale V di dimensione n , fissata una sua base B , e k^n .¹⁹ Questa applicazione è tale che $\varphi_B(v_1) = e_1, \dots, \varphi_B(v_n) = e_n$ e vettori in V sono linearmente indipendenti se e solo se lo sono le loro coordinate in k^n .

¹⁹Si noti che inoltre le coordinate del vettore somma $u+v$ sono la somma delle coordinate e la moltiplicazione per uno scalare λv corrisponde alla moltiplicazione delle coordinate per lo scalare λ . L'introduzione delle coordinate consente di lavorare direttamente in k^n .

Se $W \subseteq V$ è un sottospazio, sia $W = L\{w_1, \dots, w_r\}$, si ha una matrice delle coordinate

$$W_B \stackrel{\text{def}}{=} (w_{1B} \mid \dots \mid w_{rB})$$

dove le r -colonne sono le coordinate dei vettori rispetto alla base B di V . Si ha che

$$\dim_k L\{w_1, \dots, w_r\} = \rho(W_B) \quad (15)$$

Questo fatto riassume quanto osservato precedentemente.

Vogliamo ora calcolare la dimensione del sottospazio delle soluzioni di un sistema omogeneo. Sia $W = \text{Sol}(S) \subseteq \mathbb{R}^n$ dove S è un sistema omogeneo $AX = 0$. Si ha che

$$\dim_{\mathbb{R}} \text{Sol}(S) = n - \rho(A) \quad (16)$$

Infatti, per il Teorema 2.8, $\rho(A) = p$ sono le equazioni significative e si hanno quindi $n - p$ -parametri liberi. Una soluzione è dunque una n -upla dove compaiono p -termini pivotali determinati da $n - p$ -parametri. Assegnando il valore 1 ad un parametro e 0 ai restanti si possono costruire $n - p$ -soluzioni linearmente indipendenti, che sono anche generatori per il sottospazio delle soluzioni.

Vettori geometrici

I vettori geometrici del piano e dello spazio²⁰ costituiscono esempi paradigmatici di spazi vettoriali. Ricordiamo che un vettore geometrico è una classe di equipollenza di segmenti orientati. Questo significa che $\vec{v} = \{\vec{AB}, \vec{CD}, \dots\}$ dove i segmenti orientati \vec{AB} , \vec{CD} , etc. sono tali che $ABCD$ è un parallelogramma. Questo significa che i segmenti orientati hanno stesso modulo (anche intensità o norma), direzione e verso ovvero

- la lunghezza dei segmenti AB , CD , etc. coincide e viene denotata con $|\vec{v}|$ oppure $\|\vec{v}\|$, la norma del vettore geometrico;
- i segmenti AB , CD , etc. stanno su rette parallele, e quindi determinano una direzione;
- i segmenti \vec{AB} , \vec{CD} , etc. sono orientati nel senso che i loro estremi sono coppie **ordinate** di punti.

²⁰Piano e spazio euclideo ovvero soddisfacenti i postulati della geometria euclidea.

L'insieme dei vettori geometrici del piano e dello spazio verranno denotate \mathcal{V}_2 e \mathcal{V}_3 rispettivamente.

La somma di vettori geometrici $\vec{u} + \vec{v}$ (definita mediante la regola del parallelogramma) ed il prodotto per uno scalare $\lambda\vec{v}$ con $\lambda \in \mathbb{R}$ (naturalmente definita in modo tale che $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$ e inoltre $\vec{v} + (-1)\vec{v} = \vec{0}$ sia il vettore nullo²¹) munisce \mathcal{V}_2 e \mathcal{V}_3 di una struttura di \mathbb{R} -spazio vettoriale. Si vede (non facilmente!) che:

- due vettori geometrici non nulli $\vec{u} = \{\vec{AB}, \vec{CD}, \dots\}$ e $\vec{v} = \{\vec{EF}, \vec{GH}, \dots\}$ sono linearmente dipendenti se e solo se i segmenti AB , etc. e EF , etc. sono paralleli;
- tre vettori geometrici non nulli $\vec{u} = \{\vec{AB}, \vec{OD}, \dots\}$, $\vec{v} = \{\vec{EF}, \vec{OH}, \dots\}$ e $\vec{w} = \{\vec{IL}, \vec{ON}, \dots\}$ sono linearmente dipendenti se e solo se i punti O, D, H e N sono complanari;
- quattro vettori in \mathcal{V}_3 sono sempre linearmente dipendenti.

Da quanto detto segue che $\dim(\mathcal{V}_2) = 2$ e $\dim(\mathcal{V}_3) = 3$. Fissiamo, ad esempio, una base $\mathcal{B} : \vec{v}_1, \vec{v}_2, \vec{v}_3$ di \mathcal{V}_3 . Si ha dunque che ogni vettore $\vec{v} \in \mathcal{V}_3$ si scrive in modo unico mediante una combinazione lineare $\vec{v} = x\vec{v}_1 + y\vec{v}_2 + z\vec{v}_3$ dove la terna (x, y, z) sono le coordinate rispetto a \mathcal{B} . Come già osservato in generale, questo permette d'identificare \mathcal{V}_3 ad \mathbb{R}^3 , in particolare, una volta fissata una base. Ma inoltre il nostro spazio dei vettori geometrici è dotato di altre proprietà metriche e, alcune tra queste, si generalizzano naturalmente a tutti gli spazi \mathbb{R}^n ed anche a spazi vettoriali non di tipo finito.

Ortogonalità

In \mathcal{V}_2 e \mathcal{V}_3 si definisce un'operazione che permette di tradurre algebricamente il concetto di ortogonalità. Siano $\vec{u} = \{\vec{AB}, \vec{OP}, \dots\}$, $\vec{v} = \{\vec{CD}, \vec{OQ}, \dots\}$ due vettori non nulli. Definiamo l'angolo \widehat{uv} tra \vec{u} e \vec{v} l'angolo, compreso tra 0 e π , determinato da POQ .

Definizione 3.5 Definiamo il prodotto scalare tra due vettori geometrici non nulli \vec{u} e \vec{v} il numero reale

$$\vec{u} \cdot \vec{v} \stackrel{\text{def}}{=} |\vec{u}||\vec{v}| \cos(\widehat{uv})$$

²¹Il vettore nullo $\vec{0}$ è rappresentato da un qualunque punto.

Se uno tra \vec{u} e \vec{v} è nullo poniamo $\vec{u} \cdot \vec{v} \stackrel{\text{def}}{=} 0$.

In particolare, si noti che

$$\vec{u} \perp \vec{v} \Leftrightarrow \vec{u} \cdot \vec{v} = 0$$

e inoltre

$$\vec{v} \cdot \vec{v} = |\vec{v}|^2$$

in modo che $\vec{v} \cdot \vec{v} = 0$ implica che $\vec{v} = 0$. Possiamo sempre accorciare un vettore $\vec{v} \neq 0$, dividendolo per il suo modulo otteniamo il suo versore ovvero il seguente vettore

$$\text{vers}(\vec{v}) \stackrel{\text{def}}{=} \frac{\vec{v}}{|\vec{v}|}$$

tale che $|\text{vers}(\vec{v})| = 1$ e $\text{vers}(\vec{v})$ ha stessa direzione e stesso verso di \vec{v} . Inoltre si ha che l'angolo \widehat{uv} è identico all'angolo tra $\text{vers}(\vec{u})$ e $\text{vers}(\vec{v})$.

Osserviamo che $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ e le operazioni di somma e prodotto per uno scalare sono compatibili con l'operazione di prodotto scalare ovvero

$$\vec{u} \cdot (\lambda\vec{v} + \mu\vec{w}) = \lambda(\vec{u} \cdot \vec{v}) + \mu(\vec{u} \cdot \vec{w}).$$

Diciamo che una base di \mathcal{V}_2 o di \mathcal{V}_3 è ortonormale se è costituita da vettori di modulo unitario due a due ortogonali. Ovvero $\vec{v}_1, \vec{v}_2, \vec{v}_3$ è una base ortonormale di \mathcal{V}_3 se e solo se $\vec{v}_i \cdot \vec{v}_j = \delta_{ij}$ dove δ_{ij} è il delta di Kronecker.²² Spesso una tale base viene denotata con $\vec{i}, \vec{j}, \vec{k}$. Se abbiamo due vettori \vec{u} e \vec{v} con coordinate (x, y, z) e (x', y', z') rispettivamente ad una base ortonormale possiamo calcolare facilmente il loro prodotto scalare

$$\begin{aligned} \vec{u} \cdot \vec{v} &= (x\vec{i} + y\vec{j} + z\vec{k}) \cdot (x'\vec{i} + y'\vec{j} + z'\vec{k}) = \\ &= xx'(\vec{i} \cdot \vec{i}) + (xy' + x'y)(\vec{i} \cdot \vec{j}) + \dots + zz'(\vec{k} \cdot \vec{k}) = \\ &= xx' + yy' + zz' \end{aligned}$$

come somma dei prodotti delle coordinate. La norma di un vettore risulta una ben nota formula

$$|\vec{u}| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{x^2 + y^2 + z^2}$$

²²Il delta di Kronecker è un simbolo che rappresenta la matrice identica $I = (\delta_{ij})$.

Infine, ricordiamo che altre formulazioni classiche della geometria analitica euclidea possono essere rinvenute, in termini di coordinate rispetto ad una base ortonormale, con l'impiego del prodotto scalare²³ e mediante la distanza tra vettori $d(\vec{u}, \vec{v}) \stackrel{\text{def}}{=} \|\vec{u} - \vec{v}\| = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}$.

Parallelismo

In \mathcal{V}_3 si definisce ulteriormente una operazione che associa ad una coppia di vettori \vec{u} e \vec{v} un terzo vettore di \mathcal{V}_3 . Ricordiamo che il prodotto vettoriale tra due vettori geometrici linearmente indipendenti $\vec{u} = \{\vec{AB}, \vec{OP}, \dots\}$ e $\vec{v} = \{\vec{CD}, \vec{OQ}, \dots\}$ è il vettore denotato $\vec{u} \wedge \vec{v} \stackrel{\text{def}}{=} \{\vec{OR}, \vec{EF}, \dots\}$ tale che

- il segmento orientato \vec{OR} si determina sulla retta per O e ortogonale al piano individuato dai tre punti P , O e Q secondo la regola della mano destra e tale che
- $|\vec{u} \wedge \vec{v}| \stackrel{\text{def}}{=} |u||v| \sin(\widehat{uv})$

Se \vec{u} e \vec{v} sono linearmente dipendenti (ovvero paralleli) si pone $\vec{u} \wedge \vec{v} \stackrel{\text{def}}{=} 0$, in particolare si ha sempre $\vec{u} \wedge \vec{u} = 0$. Da quanto detto segue che $-\vec{u} \wedge \vec{v} = \vec{v} \wedge \vec{u}$ e inoltre

$$\vec{u} \wedge (\lambda \vec{v} + \mu \vec{w}) = \lambda(\vec{u} \wedge \vec{v}) + \mu(\vec{u} \wedge \vec{w}).$$

Possiamo quindi calcolare $\vec{u} \wedge \vec{v}$ in coordinate rispetto ad una base ortonormale

$$\begin{aligned} \vec{u} \wedge \vec{v} &= (x\vec{i} + y\vec{j} + z\vec{k}) \wedge (x'\vec{i} + y'\vec{j} + z'\vec{k}) = \\ &= xx'(\vec{i} \wedge \vec{i}) + (xy' - yx')(\vec{i} \wedge \vec{j}) + \dots + zz'(\vec{k} \wedge \vec{k}) = \\ &= (xy' - yx')(\vec{i} \wedge \vec{j}) + (yz' - zy')(\vec{j} \wedge \vec{k}) + (zx' - xz')(\vec{k} \wedge \vec{i}) \end{aligned}$$

Si ha che $\vec{i} \wedge \vec{j} = \vec{k}$ oppure $\vec{i} \wedge \vec{j} = -\vec{k}$. Nel primo caso la base ortonormale si dice destrorsa e in questo caso si ha che $\vec{j} \wedge \vec{k} = \vec{i}$ e $\vec{k} \wedge \vec{i} = \vec{j}$. Quindi $\vec{u} \wedge \vec{v}$ ha coordinate

$$(yz' - zy', zx' - xz', xy' - yx')$$

²³Un'equazione cartesiana della retta, fissato un vettore $\vec{v} \in \mathcal{V}_2$ con coordinate (a, b) rispetto ad una base ortonormale, si trova dall'insieme dei vettori \vec{u} ortogonali a \vec{v} ovvero $\{(x, y) \in \mathbb{R}^2 / ax + by = 0\}$. Analogamente, fissato un vettore $\vec{v} \in \mathcal{V}_3$ con coordinate (a, b, c) , l'insieme dei vettori \vec{u} ortogonali a \vec{v} ovvero $\{(x, y, z) \in \mathbb{R}^3 / ax + by + cz = 0\}$ rappresenta un piano.

rispetto ad una base ortonormale destrorsa di \mathcal{V}_3 . Si noti che il vettore $\vec{u} \wedge \vec{v}$ risulta²⁴ ortogonale sia ad \vec{u} che a \vec{v} .

Proiezioni di vettori geometrici

In \mathcal{V}_2 possiamo facilmente definire la proiezione di un qualunque vettore \vec{v} su un vettore non nullo $\vec{v}_1 \neq \vec{0}$ mediante la formula

$$\vec{p} \stackrel{\text{def}}{=} \frac{\vec{v} \cdot \vec{v}_1}{|\vec{v}_1|^2} \vec{v}_1 = |\vec{v}| \cos(\widehat{v\vec{v}_1}) \text{vers}(\vec{v}_1) \quad (17)$$

Questo risulta geometricamente, come segmenti orientati $\vec{v} = \{\vec{AB}, \vec{OP}, \dots\}$ e $\vec{v}_1 = \{\vec{CD}, \vec{OP}_1, \dots\}$, dalla proiezione ortogonale del punto P sulla retta per O e P_1 . Analogamente possiamo definire la riflessione mediante la formula

$$\vec{r} \stackrel{\text{def}}{=} 2\vec{p} - \vec{v} \quad (18)$$

sfruttando la relazione $(\vec{v} + \vec{r})/2 = \vec{p}$.

In \mathcal{V}_3 possiamo inoltre definire la proiezione (e la riflessione) di un qualunque vettore \vec{v} su un sottospazio V_2 (un piano) generato da due vettori non paralleli \vec{v}_1 e \vec{v}_2 . Anche questo risulta geometricamente dalla proiezione ortogonale di un punto su un piano, e inoltre, se la base \vec{v}_1, \vec{v}_2 di V_2 è ortonormale, abbiamo che tale proiezione si calcola come

$$(\vec{v} \cdot \vec{v}_1)\vec{v}_1 + (\vec{v} \cdot \vec{v}_2)\vec{v}_2$$

Infine, se $\vec{v}_1, \vec{v}_2, \vec{v}_3$ è una base ortonormale di \mathcal{V}_3 allora un qualunque vettore $v \in \mathcal{V}_3$ si scrive

$$v = (\vec{v} \cdot \vec{v}_1)\vec{v}_1 + (\vec{v} \cdot \vec{v}_2)\vec{v}_2 + (\vec{v} \cdot \vec{v}_3)\vec{v}_3$$

in modo unico, come si verifica facilmente. Dunque i numeri $\vec{v} \cdot \vec{v}_i$ per $i = 1, 2, 3$ rappresentano le coordinate x, y, z di v e $\vec{p}_i = (\vec{v} \cdot \vec{v}_i)\vec{v}_i$ sono le proiezioni sui vettori \vec{v}_i . Queste osservazioni saranno essenziali, in seguito, per costruire basi ortonormali.

²⁴Provare per credere, facendo il calcolo del prodotto scalare: $xyz' - xzy' + yzx' - yxz' + zxy' - zyx' = 0$, etc.