Note di Algebra lineare

Prof. Domenico Olanda

Anno accademico 2008-09

Prefazione

Questo volume raccoglie gli appunti di alcune lezioni di algebra lineare e geometria da me svolte presso la Facoltà di Scienze dell'Università "Federico II" di Napoli.

La prima parte è dedicata allo studio degli spazi vettoriali di dimensione finita. La nozione di determinante di una matrice quadrata e le ragioni del suo utilizzo sono gli aspetti essenziali del secondo capitolo. Lo studio e la risoluzione dei sistemi di equazioni lineari è l'argomento sviluppato nel terzo capitolo. Il quarto capitolo è dedicato allo studio dei prodotti scalari di uno spazio vettoriale reale.

Il problema della triangolazione di una matrice quadrata e della diagonalizzazione di un endomorfismo sono gli aspetti essenziali del quinto capitolo.

L' ultimo capitolo, il sesto, è dedicato allo studio delle funzioni lineari simmetriche.

La vastissima letteratura sugli argomenti trattati non giustifica la stesura di queste note le quali hanno solo lo scopo di aiutare gli studenti che hanno seguito le mie lezioni, nella preparazione dell'esame. Sarà utile per lo studente integrare lo studio di questi appunti con la lettura di qualche altro testo sugli stessi argomenti e di livello universitario.

C A P I T O L O I

Spazi vettoriali

1. **Gruppi abeliani.**

Un'operazione (interna) $\perp$ in un insieme S è un'applicazione $\perp : S \times S \longrightarrow S$ di $S \times S$ in S . Denoteremo con $x \perp y$ l'immagine di $\perp$ sulla coppia (x , y) e leggeremo " x composto y " .

L'operazione è detta *associativa* se risulta :

$$\forall\ x,y,z \in S, \qquad (x \perp y) \perp z = x \perp (y \perp z).$$

L'operazione $\perp$ è detta *commutativa* se risulta :

$$\forall\ x,y \in S, \quad x \perp y = y \perp x.$$

Un elemento e di S è detto *neutro* rispetto all'operazione $\perp$ se risulta :

$$\forall\, x \in S, \quad x \perp e = e \perp x = x$$

Evidentemente se esiste l'elemento neutro rispetto all'operazione $\perp$ esso è unico Siano infatti e ed e' due elementi neutri rispetto all'operazione $\perp$, si ha allora $e = e \perp e' = e'$.

Sia $(S, \perp)$ un insieme munito di un'operazione $\perp$ e dotata di elemento neutro e . Un elemento x' è detto *simmetrico* dell'elemento x se risulta :

$$x \perp x' = x' \perp x = e$$

Se l'operazione $\perp$ è associativa il simmetrico di un elemento x se esiste , è unico . Siano infatti x' ed x" due simmetrici dell'elemento x , si ha allora , in base alla definizione

$x \perp x' = x' \perp x = x \perp x'' = x'' \perp x = e$. e conseguentemente :

$$x' = x' \perp e = x' \perp (x \perp x'') = (x' \perp x) \perp x'' = e \perp x'' = x''$$

Sia S un insieme munito di un'operazione $\perp$. La coppia $(S, \perp)$ è detta un ***gruppo*** se sono verificata le seguenti proprietà :

(i) *l' 'operazione $\perp$ è associativa* ;

(ii) *esiste l'elemento neutro per l'operazione $\perp$* .

(iii) *ogni elemento di S è dotato di simmetrico* .

Quando l'operazione $\perp$ è altresì commutativa il gruppo $(S, \perp)$ è detto ***abeliano o commutativo*** .

Quando come segno per l'operazione si usa il simbolo + , si dirà che è stata adottata la *notazione additiva* ed in tal caso l'elemento neutro, se esiste, si indica con **0** e viene detto *zero* ed il simmetrico di un elemento **x** viene indicato con **-x** ed è detto *opposto* di **x** . Se come simbolo per rappresentare l'operazione si usa il simbolo $\cdot$ si dirà che è stata adottata la *notazione moltiplicativa* ; in tal caso l'elemento neutro se esiste viene indicato con **1** e viene detto *unità* ed il simmetrico di un elemento x viene indicato con $\mathbf{x}^{-1}$ o con $\frac{1}{\mathbf{x}}$ e viene detto *inverso* di x .

Diamo ora alcuni esempi di gruppi .

ESEMPIO I . Sia Z = {…..-2 , -1 , 0 , 1 , 2 ….. } l'insieme degli interi relativi e + l'usuale addizione tra interi. Lo studente verifichi che (Z , +) è un gruppo abeliano.

ESEMPIO II. Sia **Q** l'insieme dei numeri razionali e sia + l'usuale addizione tra numeri razionali . Lo studente verifichi che la coppia (**Q**, +) è un gruppo abeliano.

ESEMPIO II' . Sia **Q*** l'insieme dei numeri razionali non nulli e sia $\cdot$ l'usuale moltiplicazione tra numeri razionali . Lo studente verifichi che (**Q***, .) è un gruppo abeliano.

ESEMPIO III. Sia R^n l'insieme delle n-uple ordinate di numeri reali. Definiamo in R^n la seguente operazione di addizione:

$$(x_1, x_2,..., x_n) + (y_1, y_2,..., y_n) = (x_1 + y_1, x_2 + y_2 , ….., x_n + y_n)$$

Lo studente verifichi che (R^n, +) è un gruppo abeliano.

ESEMPIO IV. Una matrice **A** di tipo (m,n) (con m,n $\in$ N) ad elementi reali è una tabella di mn numeri reali disposti su m righe ed n colonne.

Indicando con $\mathbf{a_{ij}}$ il numero che si trova nella riga di posto i e nella colonna di posto j **A** può così scriversi

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots a_{2n} \\ \ldots & \\ a_{m1} & a_{m2} \ldots\ldots\ldots a_{mn} \end{pmatrix}$$

o semplicemente con A = ($\mathbf{a_{ij}}$) .

Indichiamo con M_{mn} (R) l'insieme di tutte le matrici di tipo (m,n) con elementi reali. Definiamo in $M_{n,m}$ (R) la seguente operazione di addizione :

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

Rispetto a tale operazione l'insieme M_{mn} (R) è un gruppo abeliano avente come elemento neutro la matrice $\underline{\mathbf{0}}$ = (0) ad elementi tutti nulli e come opposto di ogni elemento $A=(a_{ij})$ la matrice
$- A = (-a_{ij})$

ESEMPIO V . Sia **R** l'insieme dei numeri reali . Denotiamo con **R** $[x_1,x_2,..,x_n]$ l'insieme dei polinomi di grado al più uno a coefficienti in **R** nelle indeterminate ($x_1,..$ x_n) .
Se $a_o + a_i x_1+...+ a_n x_n$ e $b_o+b_1x_{1+}...+ b_nx_n$

sono due siffatti polinomi si definisce somma dei due il seguente polinomio

$$a_o + b_o + (a_1+ b_1) x_1 +....+ (a_n+ b_n) x_n$$

E' facile verificare che rispetto a tale operazione l'insieme **R** $[x_1,x_2,..,x_n]$ è un gruppo abeliano .

2. Nozione di campo.

Sia K un insieme con almeno due elementi e munito di due operazioni interne che indicheremo rispettivamente con + e ˙

Chiameremo *somma* e *prodotto* le due operazioni + e ˙ . Supporremo che l'operazione + abbia elemento neutro che indicheremo con 0 ed esista un elemento diverso da 0 e che indicheremo con 1 che sia elemento neutro per l'operazione prodotto.

La terna (K, + , ˙) è detta un ***campo*** se sono verificate

le seguenti proprietà :

1) *(K , +) è un gruppo abeliano* ;

2) *il prodotto è associativo e commutativo* ;

3) *ogni elemento diverso da zero ha inverso* ;

4) *per ogni terna a,b,c di K si ha*

$$a(b+c)= ab + ac$$

(proprietà distributiva del prodotto rispetto alla somma).

In un campo valgono le seguenti ulteriori proprietà :

i) Per ogni $a \in K$ risulta $a0 = 0$.

Dimostrazione. $a(b+0) = ab = ab + a0$ e ciò implica $a0 = 0$.

ii) *Il prodotto di due elementi è zero se e solo se uno dei due elementi è zero .*

Si ha cioé

$ab = 0$ *se e solo se* $a = 0$ oppure $b = 0$.

Dimostrazione. Se $a =0$ oppure $b=0$ per la i) è $ab = 0$. Viceversa supponiamo $ab = 0$.

Se $a = 0$ l'asserto è provato ; se $a \neq 0$ moltiplichiamo ambo i membri dell'uguaglianza $ab=0$ per l'inverso a^{-1} di a . In tal modo si ha : $a^{-1}(ab)= a^{-1}0 = 0$; poiché il prodotto è associativo risulta $a^{-1}(ab) = (a^{-1}a)b = 1b = b = 0$ e l'asserto è così provato.

Da quanto ora provato segue che se a e b sono due elementi di K diversi da zero allora il loro prodotto ab è diverso da zero . Se poniamo $K^* = K -\{0\}$ allora è facile controllare che $(K^*, \cdot)$ è un gruppo abeliano.

Sono esempi di **campi** : l'insieme dei *numeri razionali* ; l'insieme dci *numeri reali*; l'insieme dei *numeri complessi ,* (rispetto alle usuali operazioni di addizione e moltiplicazione) .

3. Spazi vettoriali su un campo.

Sia (K , + , $\cdot$) un campo i cui elementi saranno detti *scalari* e sia **V** un insieme i cui elementi saranno detti *vettori* con due operazioni : una di ***addizione*** tra vettori

$$+ : V \times V \longrightarrow V$$

e l'altra di ***moltiplicazione*** *esterna*

$$* : K \times V \longrightarrow V.$$

la quale fa corrispondere ad ogni coppia (α , v) scalare- vettore ancora un vettore indicato con

$$\alpha * v$$

L'insieme V è detto uno ***spazio vettoriale sul campo K***, rispetto alle operazioni + e $*$, se sono verificate le seguenti proprietà :

(3.1) **(V, +) è un gruppo abeliano ;**

(3.2) $\boldsymbol{\alpha * v + \alpha * w = \alpha * (v + w)}$

(3.3) $\boldsymbol{\alpha * v + \beta * v = (\alpha + \beta) * v}$

(3.4) $\boldsymbol{1 * v = v}$ ***(con 1 si è indicata l'unità di K)***

(3.5) $\boldsymbol{(\alpha \cdot \beta) * v = \alpha * (\beta * v)}$

(per ogni coppia di scalari α , β e per ogni coppia di vettori v , w di V).

Nel seguito per semplicità, scriveremo $\alpha\beta$ anziché $\alpha \cdot \beta$ ed αv anziché $\alpha * v$. Indicheremo usualmente con lettere greche gli scalari e con lettere latine i vettori. Il vettore nullo (elemento neutro rispetto alla somma in V) sarà indicato con $\underline{\mathbf{0}}$ distinguendolo così dallo zero di K che sarà indicato con 0.

Diamo ora alcuni esempi di spazi vettoriali :

ESEMPIO 1. Sia K un campo . Nell'insieme K^n ($n \in N$, $n \geq 2$) delle n-ple ordinate di elementi di K introduciamo le seguenti due operazioni di somma e prodotto:

$$(a_1,a_2,...,a_n) + (b_1,b_2,...,b_n) = (a_1 + b_1, a_2 + b_2,..., a_n + b_n)$$

$$\alpha * (a_1 , a_2,..., a_n) = (\alpha a_1 , \alpha a_2 ....\alpha a_n)$$

E' non difficile controllare che rispetto a tali operazioni l'insieme $\mathbf{K^n}$ è uno spazio vettoriale su K . Tale spazio è detto ***spazio vettoriale numerico di dimensione n sul campo K .***

ESEMPIO 2. Nell'insieme $M_{m,n}(K)$ delle matrici di tipo m,n ad elementi nel campo K

definiamo le seguenti due operazioni di addizione e prodotto esterno

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

$$\alpha * (a_{ij}) = (\alpha a_{ij})$$

Lo studente verifichi che rispetto a tali operazioni l'insieme $M_{m,n}(K)$ è uno spazio vettoriale su K .

ESEMPIO 3. Denotiamo con $K[x_1,x_2,..,x_n]$ l'insieme dei polinomi di grado al più uno nelle indeterminate ($x_1,..\quad x_n$) a coefficienti nel campo K. Definiamo in $K[x_1,x_2,..,x_n]$ le seguenti due operazioni di addizione e prodotto esterno :

$$(a_o + a_1x_1+...+ a_nx_n) + (b_o+b_1x_1+...+ b_nx_n) = a_o + b_o + (a_1+ b_1)x_1 +....+ (a_n+ b_n)x_n$$

$$\alpha*(a_o + a_1x_1 +...+ a_nx_n) = \alpha a_o + \alpha a_1x_1 \mathbf{+...+} \alpha a_nx_n$$

Si verifichi che rispetto a tali due operazioni l'insieme $K[x_1,x_2,..,x_n]$ è uno spazio vettoriale sul campo K .

Analizziamo ora alcune proprietà valide in uno spazio vettoriale V su un campo K . Ricordiamo che indicheremo con $\underline{\mathbf{0}}$ il vettore nullo di V, cioè l'elemento neutro rispetto alla somma definita in V, e con 0 lo zero di K .

1. Per ogni vettore v si ha $0\, v = \underline{0}$.

Dimostrazione . Si scelga uno scalare α . Si ha :

$$(\alpha + 0)v = \alpha\, v + 0\, v = \alpha\, v$$

e questa comporta $0v = \underline{0}$.

2. Per ogni scalare α risulta $\alpha\, \underline{0} = \underline{0}$.

Dimostrazione . Sia v un vettore, si ha :

$$\alpha\, (\underline{0}+ v)= \alpha\, \underline{0} + \alpha\, v = \alpha\, v$$

e questa comporta $\alpha\, \underline{0} = \underline{0}$.

3. Per ogni scalare α e per ogni vettore v risulta :

$$\alpha\, v = \underline{0} \qquad \textit{se e solo se è} \qquad \alpha = 0 \text{ oppure } v = \underline{0}$$

Dimostrazione . Se $\alpha = 0$ oppure è $v = \underline{0}$ allora per 1 e 2 è $\alpha\, v = \underline{0}$.

Viceversa supponiamo $\alpha\, v =\underline{0}$ ed $\alpha \neq 0$. Moltiplicando ambo i membri dell'eguaglianza

$\alpha\, v =\underline{0}$ per $\frac{1}{\alpha}$ si ha $\frac{1}{\alpha}\, (\alpha\, v)= \frac{1}{\alpha}\, \underline{0} = \underline{0}$

da cui segue

$$\frac{1}{\alpha}\, (\alpha\, v) = (\frac{1}{\alpha}\, \alpha\,) v = 1v = v = \underline{0}.$$

4. Per ogni scalare α e per ogni vettore v si ha :

$$-(\alpha\, v) = (-\alpha)v = \alpha(-v)$$

Dimostrazione. Da $(\alpha\, v) + (-\alpha)v = 0v = \underline{0}$ segue che è

$(-\alpha)v = -(\alpha\, v)$. Da $\alpha\, v + \alpha(-v) = \alpha\, \underline{0} = \underline{0}$ segue che è $\alpha(-v) = -(\alpha\, v)$.

Si osservi che dalla 4 segue che quando si moltiplica un vettore v per lo scalare -1 si ottiene il vettore -v opposto di v .

Sia V uno spazio vettoriale sul campo K . Un *sottospazio* di V è un suo sottoinsieme non vuoto H verificante le seguenti due proprietà :

$$(i) \quad v, w \in H \Rightarrow v + w \in H$$

$$(ii) \quad \alpha \in K, v \in H \Rightarrow \alpha v \in H$$

Ogni sottospazio contiene almeno il vettore nullo. Infatti poiché H è non vuoto esso possiede almeno un vettore v. Moltiplicando v per 0 si ottiene , per la (ii) , ancora un vettore di H e quindi H possiede il vettore nullo.

Evidentemente scegliendo H = $\{\underline{0}\}$ oppure H = V si realizza un sottospazio. Tali sottospazi sono detti *banali.*

Come si possono costruire sottospazi non banali ? Vediamo.

Siano $v_1, v_2, ..., v_h$, h vettori ed $\alpha_1, \alpha_2, ... \alpha_h$, h scalari . Consideriamo il vettore w dato da

$$w = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_h v_h$$

il vettore w così ottenuto si dice che è *combinazione lineare* dei vettori $v_1, v_2, ..., v_h$ o si dice che w *dipende linearmente* dai vettori $v_1, v_2, ..., v_h$.

Gli scalari $\alpha_1, \alpha_2, ... \alpha_h$ che figurano nella espressione $w = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_h v_h$ sono detti i *coefficienti* della combinazione lineare .

Sia ora H l'insieme dei vettori w ciascuno dei quali sia una combinazione lineare dei vettori

$v_1, ..., v_h$. Evidentemente H è un sottospazio di V ; esso è detto *sottospazio generato dai vettori* $v_1, ..., v_h$ e viene indicato col simbolo $[v_1, v_2, ..., v_h]$.

Quando ogni vettore di V è esprimibile come combinazione lineare dei vettori $v_1, ..., v_h$ cioè quando risulti $V = [v_1,..., v_h]$ allora il sistema $\{ v_1,v_2, ...,v_h \}$ è detto un ***sistema di generatori*** per lo spazio vettoriale V e lo spazio V è detto *finitamente generabile* (in quanto attraverso un numero finito di suoi vettori si possono generare tutti gli altri).

Noi supporremo sempre che lo spazio vettoriale V assegnato sia finitamente generabile e non ridotto al solo vettore nullo.

A titolo di esempio si consideri lo spazio vettoriale $\mathbf{V} = \mathbf{R}^2$ i cui vettori sono le coppie ordinate di numeri reali. Si scelgano in **V** i seguenti sistemi di vettori :

$S_1 = \{(1 , 0) , (0 , 1)\}$

$S_2 = \{ (2 , 0) , (0 , 2) , (2 , 3) \}$

$S_3 = \{ (1 , 0) , (3, 1) \}$

$S_4 = \{ (1 , 0) , (4 , 0) \}$

I vettori di S_1 sono un sistema di generatori in quanto ogni coppia (a , b) risulta una loro combinazione lineare risultando precisamente

$$(a, b) = a (1 , 0) + b (0 , 1)$$

I vettori di S_2 sono anch' essi un sistema di generatori in quanto ogni altra coppia (a , b) risulta una loro combinazione lineare risultando ad esempio

$$(a, b) = \frac{a}{2} (2 , 0) + \frac{b}{2} (0 , 2) + 0 (2 , 3)$$

I vettori di S_3 sono anch' essi un sistema di generatori se ogni altra coppia (a , b) risulta una loro combinazione lineare cioè se sia possibile trovare due numeri α e β tali che risulti :

$$(a, b) = \alpha (1 , 0) + \beta (3 , 1)$$

Questa relazione è equivalente a :

$$(a, b) = (\alpha + 3 \beta ,\ \beta)$$

e quindi basta scegliere $\beta = b$ ed $\alpha = a - 3b$.

I vettori di S_4 non sono un sistema di generatori in quanto le sole coppie che si possono costruire con le coppie $(1 , 0)$, $(4 , 0)$ sono le coppie del tipo $(a , 0)$.

Si osservi che i vettori di S_1 sono un sistema di generatori ma altresì risultano quelli di S_2 .

Nel sistema S_1 i due vettori sono entrambi essenziali perché con un solo dei due non si potrebbe costruire l'altro. Nel sistema S_2 l'ultimo vettore sembra invece svolgere un ruolo marginale per la costruzione degli altri vettori . Come si può decidere ,in presenza di un sistema di generatori, quali vettori siano *indispensabili e quali no* ?

Per rispondere a questa domanda occorre introdurre la seguente nozione *dipendenza* ed *indipendenza lineare* di un sistema di vettori.

Siano assegnati h vettori $v_1,v_2, ...,v_h$. Il vettore nullo è generato dai vettori $v_1,v_2, ...,v_h$ in modo molto semplice quando si moltiplichi ognuno di essi per 0, si ha cioè :

$$\underline{\mathbf{0}} = 0\, v_1 + 0\, v_2 + \ldots + 0\, v_h$$

Se questa è l'***unica*** possibilità che abbiamo per costruire il vettore nullo a partire dai vettori $v_1,v_2, ...,v_h$ allora tali vettori vengono detti ***linearmente indipendenti***.

Quando i vettori non sono linearmente indipendenti essi vengono detti ***linearmente dipendenti***.

Quindi ribadendo se i vettori $v_1,v_2, ...,v_h$ sono linearmente dipendenti allora esistono h scalari $\alpha_1, \alpha_2, \ldots \alpha_h$ non ***tutti nulli*** , tali che risulti

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{\mathbf{0}}$$

Le proposizioni che seguono aiutano a stabilire se alcuni vettori assegnati siano o meno linearmente dipendenti .

Proposizione 3.1 *I vettori* $v_1,v_2, ...,v_h$ *sono linearmente dipendenti se e solo se uno di essi dipende dai rimanenti.*

Dimostrazione. Supponiamo che i vettori $v_1, v_2, ..., v_h$ siano dipendenti. Esistono allora h scalari $\alpha_1, \alpha_2, ...\ \alpha_h$ non **tutti nulli** , tali che risulti

$$(*) \qquad \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{\mathbf{0}}$$

supposto ad esempio che sia $\alpha_1 \neq 0$ dalla (*) segue :

$$v_1 = -\frac{1}{\alpha_1}(\alpha_2 v_2 + \ldots + \alpha_h v_h) = \beta_2 v_2 + \ldots + \beta_h v_h$$

avendo posto $\beta_i = -\dfrac{\alpha_i}{\alpha_1}$ $i = 2,\ldots, h$, e quindi abbiamo mostrato che uno dei vettori , in questo caso v_1 , dipende dai rimanenti. Viceversa supponiamo che uno dei vettori dipenda dai rimanenti e per fissare le idee supponiamo sia l'ultimo di essi a dipendere dai rimanenti . Sia quindi

$$v_h = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{h-1} v_{h-1} .$$

Da questa relazione segue

$$\underline{\mathbf{0}} = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{h-1} v_{h-1} - v_h = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{h-1} v_{h-1} + (-1) v_h$$

la quale mostra che i vettori sono linearmente dipendenti in quanto il vettore nullo è stato ottenuto con scalari non tutti nulli figurando tra essi lo scalare -1 .

Dalla proposizione ora provata segue questa proprietà che ci sarà spesso utile nel seguito.

Proposizione 3.2 *Siano v e w due vettori entrambi non nulli . I vettori v e w sono dipendenti se e sole essi sono proporzionali.*

Un' altra conseguenza della proposizione 3.1 è la seguente

Proposizione 3.3 *Se uno dei vettori $v_1, v_2, ..., v_h$ è il vettore nullo allora i vettori $v_1, v_2, ..., v_h$ sono linearmente dipendenti. Se due dei vettori $v_1, v_2, ..., v_h$ sono tra loro proporzionali allora i vettori $v_1, v_2, ..., v_h$ sono linearmente dipendenti.*

Ovviamente la proposizione 3.1 che caratterizza i sistemi di vettori linearmente dipendenti equivale alla proposizione che segue e che serve a caratterizzare i sistemi di vettori linearmente indipendenti.

Proposizione 3.4 *I vettori $v_1, v_2, ..., v_h$ sono linearmente indipendenti se e solo se nessuno di essi dipende dai rimanenti.*

Ritornando agli esempi precedenti possiamo allora osservare che :

1. I vettori di S_1 sono indipendenti in quanto non proporzionali .
2. I vettori di S_2 sono dipendenti in quanto il terzo vettore dipende dagli altri due.
3. I vettori di S_3 sono indipendenti in quanto non proporzionali.
4. I vettori di S_4 sono dipendenti in quanto proporzionali.

Le considerazioni che seguono giustificano l'importanza di poter stabilire se alcuni vettori assegnati siano o meno dipendenti.

Siano assegnati h vettori e sia $H = [v_1, v_2, ..., v_h]$ lo spazio da essi generato. Se i vettori $v_1, v_2, ..., v_h$ sono dipendenti, uno di essi , supponiamo v_1 , dipende dai rimanenti ed allora facilmente si riconosce che lo spazio generato da $v_1, v_2, ..., v_h$ coincide con lo spazio W generato dai soli vettori $v_2, ..., v_h$. Al fine di valutare da quali vettori sia costituito H sembra quindi non essenziale la presenza del vettore v_1 che può quindi essere eliminato. Se anche i vettori $v_2, ..., v_h$ fossero dipendenti allora uno di essi supponiamo sia v_2 dipende dai rimanenti . Ma allora lo spazio W generato da $v_2, ..., v_h$ coincide con lo spazio T generato da $v_3, ..., v_h$ e così anche v_2 può essere eliminato avendo constatato che risulta $H = W = T = [v_3, v_4, ..., v_h]$. Tale procedimento iterato si arresterà quando non c'è più un vettore che dipende dai rimanenti e cioè quando i vettori rimasti siano **linearmente indipendenti** .

Queste considerazioni mostrano che se i vettori $v_1, v_2, ..., v_h$ sono un sistema di generatori per lo spazio V essi sono tutti essenziali se essi risultano indipendenti.

Un sistema di ***generatori indipendenti*** è detta una ***base*** dello spazio vettoriale.

Molto importante per ciò che segue è il seguente :

Teorema di Steinitz.

Siano assegnati due sistemi di vettori $S = \{ \mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_m} \}$ e $T = \{ \mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_t} \}$. Se i vettori di S sono linearmente indipendenti ed ognuno di essi dipende dai vettori di T allora il numero dei vettori di S è minore o eguale al numero di vettori di T , risulta cioè $\mathbf{m} \leq \mathbf{t}$.

Dimostrazione. Faremo la dimostrazione ragionando per induzione sulla cardinalità t di T .

Proviamo che il teorema è vero se $t=1$. Se $t=1$ allora T possiede un sol vettore b_1 e noi dobbiamo provare che anche S non può avere più di un vettore . Supponiamo per assurdo che S abbia almeno due vettori a_1,a_2 . Poiché ogni vettore di S dipende dai vettori di T si ha

$a_1 = \alpha b_1$ ed $a_2 = \beta b_1$. Poiché i vettori di S sono indipendenti a_1 non può essere il vettore nullo e quindi è $\alpha \neq 0$. Risulta allora

$b_1 = \frac{1}{\alpha} a_1$ e quindi è $a_2 = \frac{\beta}{\alpha} a_1$. S ha quindi due vettori proporzionali e ciò è assurdo perché i suoi vettori sono linearmente indipendenti. Supponiamo quindi $t > 1$ e vero il teorema per $t-1$. Poichè ogni vettore di S dipende dai vettori di T sussistono le seguenti relazioni :

$$a_1 = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_t b_t$$
$$a_2 = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_t b_t$$
…………..

$$a_m = \gamma_1 b_1 + \gamma_2 b_2 + \ldots + \gamma_t b_t$$

Poiché i vettori di S sono indipendenti a_1 non può essere il vettore nullo e quindi almeno uno degli scalari $\alpha_1, \alpha_2, \ldots \alpha_t$ è non nullo e supponiamo sia $\alpha_1 \neq 0$. Dalla prima relazione si può allora ricavare b_1 come combinazione di a_1 , b_2, …,b_t si ha cioè per b_1 una espressione del tipo

$$b_1 = \delta_1 a_1 + \delta_2 b_2 + \ldots + \delta_t b_t$$

Sostituiamo ora tale espressione di b_1 nelle relazioni

$$a_2 = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_t b_t$$
…………..

$$a_m = \gamma_1 b_1 + \gamma_2 b_2 + \ldots + \gamma_t b_t$$

rimaste e troviamo allora che valgono relazioni di questo tipo

$$a_2 - k_1 a_1 = \zeta_2 b_2 + \ldots + \zeta_t b_t$$
…………..

$$a_m - k_m a_1 = \eta_2 b_2 + \ldots + \eta_t b_t$$

I vettori $w_2 = a_2 - k_l a_1, \ldots, w_m = a_m - k_m a_1$, in numero di m-1 sono indipendenti, in quanto se uno di essi dipendesse dai rimanenti anche in S ci sarebbe un vettore che dipende dai rimanenti, ed inoltre ognuno di essi dipende dai vettori $b_2, \ldots, b_t$ che sono in numero di t-1.

Poiché per t-1 il teorema è vero si ha $m-1 \leq t-1$ e quindi è, come si voleva, $m \leq t$.

Una prima importante conseguenza del teorema ora provato è la seguente

Proposizione 3.5 *Sia V uno spazio vettoriale e sia* $B = \{ e_1, e_2, \ldots, e_n \}$ *una sua base di cardinalità n . Ogni altra base di V ha cardinalità n.*

Dimostrazione. Sia $B' = \{ b_1, b_2, \ldots, b_t \}$ un' altra base di V .Applicando due volte il teorema di Steinitz si ha $n \leq t$ e $t \leq n$ e quindi $t = n$.

Abbiamo così provato che le basi di uno spazio vettoriale finitamente generabile hanno tutte la stessa cardinalità. Detta **n** la cardinalità comune a tutte le basi l'intero **n** è detto la ***dimensione di V***.

Al fine di fornire ulteriori interpetrazioni dell'intero **n** , *dimensione di V* , è molto utile la seguente :

Proposizione 3.6 *Se i vettori* $v_1, v_2, \ldots, v_h$, w *sono dipendenti mentre i vettori* $v_1, v_2, \ldots, v_h$ *sono indipendenti allora il vettore* w *dipende dai vettori* $v_1, v_2, \ldots, v_h$.

Dimostrazione. Per ipotesi poiché i vettori $v_1, v_2, \ldots, v_h$, w sono dipendenti esistono scalari $(\alpha, \alpha_1, \alpha_2, \ldots \alpha_h)$ non tutti nulli per cui risulti :

$$(*) \qquad \alpha w + \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{0}$$

Se fosse $\alpha = 0$ avremmo da (*)

$$0w + \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{0} + \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h =$$

$$= \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{0}$$

e questa per la supposta indipendenza dei vettori $v_1, v_2, \ldots, v_h$ comporterebbe altresì $\alpha_1 = 0$, $\alpha_2 = 0$, ..., $\alpha_h = 0$ e quindi gli scalari $(\alpha, \alpha_1, \alpha_2, \ldots \alpha_h)$ sarebbero tutti nulli contro il supposto. Pertanto risulta $\alpha \neq 0$ e quindi da (*) segue :

$$w = -\frac{1}{\alpha}(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h) = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_h v_h$$

avendo posto $\beta_i = -\dfrac{\alpha_i}{\alpha}$ $\quad$ i= 1,2,.., h.

Possiamo ora provare la seguente

Proposizione 3.7 *L'intero n è dimensione dello spazio vettoriale V se e solo se esso esprime il massimo numero di vettori indipendenti che V possiede.*

Dimostrazione. Supponiamo che lo spazio V abbia dimensione n e sia $B=\{ e_1 , e_2,…., e_n \}$ una sua base. Se $v_1,v_2, ...,v_h$ sono h vettori indipendenti qualsiasi di V , poiché ognuno di essi dipende dai vettori di B, in forza del teorema di Steinitz, risulta $h \leq n$. Pertanto n è il massimo numero di vettori indipendenti di V.

Viceversa supponiamo che V possegga n vettori indipendenti $v_1,v_2, ...,v_n$ e che tale numero sia il *massimo numero* di vettori indipendenti che V possiede. Se w è un qualunque vettore diverso dai vettori $v_1,v_2, ..., v_n$ allora il sistema $v_1,v_2, ..., v_n, w$, ottenuto aggiungendo w ai vettori $v_1,v_2, ...,v_n$, avendo cardinalità n+1 è costituito da vettori linearmente dipendenti. Per la proposizione 3.6, w è quindi combinazione lineare dei vettori $v_1,v_2, ...,v_n$. Poiché anche i singoli vettori v_i dipendono da $v_1,v_2, ...,v_n$ allora $v_1,v_2, ...,v_n$ è un sistema di generatori per V e quindi essendo tali vettori anche indipendenti essi costituiscono una sua base e pertanto V ha dimensione n.

La proposizione 3.6 suggerisce un metodo per costruire insiemi di vettori indipendenti ed una base di V . Vediamo come .

Si consideri un vettore v_1 non nullo . Per la proprietà 3. di pag 9 il vettore v_1 è indipendente. Sia $H_1 = [v_1]$ lo spazio generato da v_1.

Se $H_1 = V$ allora $\{v_1 \}$ è una base di V . Se $H_1 \subset V$ allora sia v_2 un vettore scelto in $V- H_1$. I vettori $\{v_1 , v_2\}$ per la proposizione 3.6 sono indipendenti. Sia $H_2 = [v_1 , v_2]$ lo spazio generato dai vettori $\{v_1 , v_2\}$. Se $H_2 = V$ allora $\{v_1 , v_2\}$ è una base di V. Se invece è $H_2 \subset V$ possiamo scegliere un ulteriore vettore v_3 in $V - H_2$ che aggiunto ai vettori $\{v_1 , v_2\}$ darà luogo ad un sistema di tre vettori $\{v_1 , v_2, v_3\}$ indipendenti. Se lo spazio ha dimensione n tale procedimento sarà iterato n volte e ci consentirà di costruire una base di V.

Abbiamo quindi provato la seguente

Proposizione 3.8 *Sia V uno spazio vettoriale di dimensione n e siano $e_1 , e_2 ,.., e_t$, t , ($t < n$) , vettori indipendenti di V . Si possono allora aggiungere altri n-t vettori opportuni $e_{t+1} , e_{t+2} ,.., e_n$ in modo che $e_1 , e_2 ,.., e_t$ $e_{t+1} , e_{t+2} ,.., e_n$ sia una base di V .*

Al fine di fornire alcune caratterizzazioni delle basi di uno spazio vettoriale ci è utile richiamare alcune semplici definizioni . Sia X un sottoinsieme di un insieme S, ed X sia munito di una certa proprietà "p" . Si dice che X è ***massimale rispetto alla proprietà "p"*** se ogni insieme Y che contenga propriamente X non ha più la proprietà "p" . Si dice che X è ***minimale rispetto alla proprietà "p"*** se ogni sua parte propria non ha più la proprietà "p" .

Siamo ora in grado di provare alcune importanti equivalenze:

Proposizione 3.9 *Per un sistema* $S = \{v_1, v_2,\dots, v_n\}$ *di vettori di uno spazio vettoriale V sono equivalenti le seguenti affermazioni* :

a) S è una base (cioè un sistema di generatori indipendenti)
b) S è massimale rispetto alla proprietà di essere indipendente
c) S è minimale rispetto alla proprietà di essere un sistema di generatori .
d) S è un sistema indipendente di cardinalità massima.
e) S è un sistema di generatori di cardinalità minima.

Dimostrazione. Proviamo che a) e b) sono equivalenti.

Mostriamo che a) implica b). Se aggiungiamo ad S un ulteriore vettore w il sistema $\{v_1 , v_2,\dots , v_n, w\}$ è costituito da vettori dipendenti in quanto ,avendo supposto che $\{v_1 , v_2,\dots, v_n\}$ sono un sistema di generatori , w è combinazione lineare di $\{v_1 , v_2,\dots, v_n\}$. Pertanto l'insieme S rispetto alla proprietà di essere costituito da vettori indipendenti è massimale. Viceversa supponiamo di sapere che l'insieme S sia costituito da vettori indipendenti e sia massimale rispetto a tale proprietà. Se aggiungiamo ad S un ulteriore vettore w il sistema $\{v_1 , v_2,\dots , v_n, w\}$ è costituito da vettori dipendenti per la supposta massimalità di S rispetto alla proprietà di essere costituito da vettori indipendenti. Per la proposizione 3.5, w è allora combinazione lineare dei vettori di S . Per l'arbitrarietà di w e tenendo conto che ogni vettore v_i dipende da $v_1 , v_2,\dots , v_n$ è provato che S è un sistema di generatori.

Proviamo che a) è equivalente a c). Proviamo che a) implica c). Se si priva S= $\{v_1 , v_2,\dots, v_n\}$ di un suo vettore ad esempio di v_1 , i vettori$\{v_2,\dots, v_n\}$ che restano non sono più un sistema di generatori in quanto essendo $\{v_1 , v_2,\dots, v_n\}$ indipendenti nessuno dei suoi vettori può essere generato dai rimanenti. Pertanto S è minimale rispetto alla proprietà di essere un sistema di generatori. Viceversa se sappiamo che i vettori di S sono un sistema di generatori ma minimale rispetto a tale proprietà allora i suoi vettori sono indipendenti . Infatti se fossero dipendenti uno di essi e, per fissare le idee, sia il primo , dipende dai rimanenti. Ma allora come già visto in precedenza risulta $[v_2,\ldots, v_n] =$ $= [v_1 , v_2,\ldots, v_n] = V$. Quindi anche $T =\{ v_2,\ldots, v_n\}$ è un sistema di generatori pur essendo una parte propria di S il che va contro la supposta minimalità di S rispetto alla proprietà di essere un sistema di generatori.

L'equivalenza tra a) e d) è stata già acquisita con la proposizione 3.7.

Proviamo infine l'equivalenza tra a) ed e). Proviamo che a) implica e). Se trovassimo t vettori $w_1 , w_2,\ldots, w_t$ che generano V per il teorema di Steinitz risulta $n \leq t$. Quindi ogni altro sistema di generatori ha cardinalità almeno n . L'intero n esprime quindi la minima cardinalità di un sistema di generatori e quindi l'insieme S come sistema di generatori ha cardinalità minima. Viceversa se sappiamo che S è un sistema di generatori e che come tale ha cardinalità minima allora i vettori $v_1 , v_2,\ldots, v_n$ che lo costituiscono sono indipendenti. Infatti se fossero dipendenti uno di essi e, per fissare le idee, sia il primo , dipende dai rimanenti. Ma allora come già visto in precedenza risulta $[v_2,\ldots, v_n]=[v_1 , v_2,\ldots, v_n]= V$.Quindi anche $T =\{ v_2,\ldots, v_n\}$ è un sistema di generatori pur essendo di cardinalità n-1 mentre avevamo supposto che n fosse la cardinalità minima di un sistema di generatori.

Dalle proposizioni provate segue che se uno spazio vettoriale V finitamente generabile ha dimensione **n** , l' intero **n** può anche essere definito come il ***numero massimo di vettori indipendenti*** che V possiede o come il ***numero minimo di generatori di V*.**

Possiamo ora valutare la dimensione dello spazio vettoriale numerico illustrato nell'esempio I .

ESEMPIO I. $(K^n , + , . K)$. In tale spazio i vettori (1, 0,.. , 0), (0, 1 ,0.. 0), …..,(0, 0,.., 1) sono un sistema di generatori in quanto risulta :

$$a_1 (1,0,..0) + a_2 (0,1,.. \ 0) + ... + a_n (0,0,...1) = (a_1 , a_2, ... , a_n)$$

In particolare si ha che solo

$$0 (1,0,..0) + 0 (0,1,.. \ 0) + ... + 0 (0,0,...1) = (0 , 0, ... 0)$$

il che prova che essi sono anche indipendenti. I vettori (1, 0,.. , 0), (0, 1 ,0.. 0), …..,(0, 0,.., 1)

sono quindi una base , detta ***base canonica*** , di K^n che ha quindi dimensione **n** .

Aver stabilito che la dimensione di K^n sia **n** ci consentirà di saper valutare semplicemente anche la dimensione degli altri spazi vettoriali mostrati negli altri esempi II e III.

Per giustificare la nostra affermazione saranno molto utili le considerazioni che seguono.

Concludiamo tale numero illustrando un esempio di spazio vettoriale sul campo reale di dimensione tre che sarà molto utilizzato nel capitolo dedicato alla geometria analitica.

ESEMPIO IV . Si consideri un punto A dello spazio reale S e sia V_A l'insieme di tutti i segmenti orientati AP di primo estremo A, al variare di P in S . Indicheremo con |AP | la lunghezza del segmento AP . Quando P = A il segmento corrispondente AA ha lunghezza zero, sarà chiamato *segmento nullo,* e sarà indicato con $\underline{0}$ **.**
Se v = AP e w = AP' sono due elementi non nulli di V_A si definisce somma di v e w il segmento v + w = AT ottenuto col seguente procedimento

i) *se AP ed AP' hanno direzione diversa , AT è la diagonale del parallelogramma di lati AP ed AP'*

ii) se AP ed AP' hanno la stessa direzione δ e lo stesso verso v allora AT è il segmento che ha la direzione δ e verso v e lunghezza | AT | = |AP | + |AP'|.

iii) se AP ed AP' hanno la stessa direzione δ ma verso opposto allora AT è il segmento nullo se |AP|= |AP'|. Se invece è |AP| ≠ |AP'| (supposto |AP| > |AP'|) allora AT ha la direzione δ il verso di AP e lunghezza | AT | = |AP | - |AP'|.

Se v = AP e w = $\underline{0}$ assumeremo v + $\underline{0}$ = v

Se v = AP ed α è un numero reale . Si definisce α v = AT il segmento così ottenuto .

Il segmento AT è nullo se α = 0 oppure se v = $\underline{0}$.

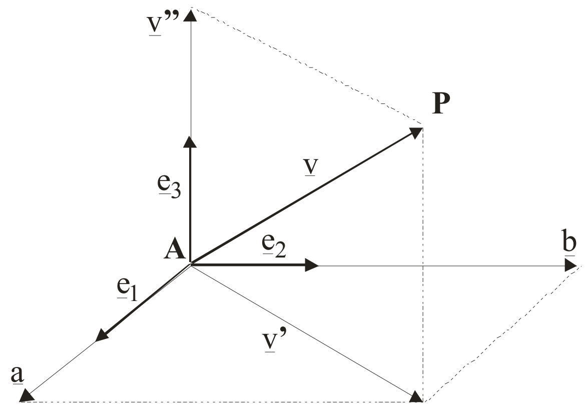
Supposto v = AP non nullo ed α ≠ 0 , allora detta δ la direzione di AP e ν il verso di AP , il segmento AT ha :

direzione δ , *lunghezza* | AT | = | α ||AP | , *verso* ν *se* α > *0 e verso opposto se è* α < *0.*

Si prova che l'insieme V_A con le due operazioni ora definite è uno spazio vettoriale . I segmenti AP saranno in seguito chiamati ***vettori geometrici applicati in A .***

Lo spazio vettoriale V_A ha dimensione tre come ora proveremo.

Siano $e_1 = A\,U_1$, $e_2 = AU_2$ ed $e_3 = AU_3$ tre segmenti ***non nulli e non complanari*** . Tali vettori sono tali che nessuno di essi può essere generato dagli altri due e sono quindi **indipendenti** . Inoltre, riferendoci ai vettori indicati in figura **,**



Si ha

$$AP = v = v' + v'' = a + b + v''$$

Ma è , per opportuni scalari α , β , γ

$$a = \alpha\, e_1 \quad , \qquad b = \beta\, e_2 \quad , \qquad v'' = \gamma\, e_3$$

e quindi è

$$v = a + b + v'' = \alpha\, e_1 + \beta\, e_2 + \gamma\, e_3$$

I tre vettori e_1, e_2 , e_3 sono quindi una base di V_A che ha così *dimensione tre* .

4. Isomorfismi tra spazi vettoriali.

Siano assegnati due spazi vettoriali V e W costruiti sullo stesso campo K .

Una funzione $f : V \longrightarrow W$ tra V e W è detta un ***isomorfismo*** se essa è **biettiva** e **lineare** cioè se valgono per essa le seguenti proprietà :

1. f è biettiva.
2. $f (v + v') = f(v) + f(v')$
3. $f (\alpha v) = \alpha f(v)$

(per ogni coppia di vettori v , v' e per ogni scalare α)

Se una funzione f di V in W verifica solo le proprietà 2. e 3. ma non è biettiva si dice che essa è una *funzione lineare* di V in W.

Gli isomorfismi sono quindi particolari funzioni lineari perché sono quelle biettive .Quando esiste un isomorfismo tra i due spazi vettoriali tali spazi vengono detti tra loro ***isomorfi***.

Si prova facilmente che se f è un isomorfismo tale risulta anche la funzione f^{-1} e che componendo due isomorfismi si ottiene ancora un isomorfismo.

Lo studente verifichi che se si associa ad un polinomio $a_o + a_1x_1+...+ a_nx_n$ la (n+1)-pla $(a_o , a_1,\ldots. a_n)$ dei suoi coefficienti si realizza un **isomorfismo** tra gli spazi vettoriali $K[x_1,...,x_n]$ e K^{n+1}.

Se si associa ad una matrice

$$\begin{pmatrix} a_{11} & a_{12} \ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots a_{2n} \\ \ldots & \\ a_{m1} & a_{m2} \ldots\ldots a_{mn} \end{pmatrix}$$

il vettore numerico

$$(a_{11}\, a_{12} \ldots a_{1n} ,\ a_{21}\, a_{22} \ldots a_{2n} \ , \ldots\ldots\ldots\ldots,\ a_{m1}\, a_{m2} \ldots a_{mn})$$

di K^{mn} che si ottiene disponendo in sequenza ed in orizzontale una dopo l'altra le righe della matrice si realizza un isomorfismo tra gli spazi vettoriali $M_{m,n}(K)$ e K^{mn} .

Vediamo ora se per ogni spazio vettoriale possiamo trovarne un altro “ magari più semplice” ad esso isomorfo. Vediamo.

Sia quindi V uno spazio vettoriale sul campo K di dimensione n e sia B = (e_1 , e_2,…., e_n) una sua base ordinata (*riferimento*). Poiché i vettori e_1 , e_2,…., e_n sono un sistema di generatori per V ogni vettore v risulta una loro combinazione lineare si ha cioè

$$v = x_1e_1 +...+ x_ne_n$$

I numeri $(x_1,…..,x_n)$ che consentono di esprimere v come combinazione di (e_1 , e_2,…., e_n) sono dette le *coordinate* di v nella base fissata.

Mostriamo ora che le coordinate di v sono *univocamente determinate* da v e che quindi v si può scrivere in *unico modo* come combinazione lineare dei vettori e_1 , e_2,…., e_n . Supponiamo quindi che v sia stato ottenuto anche attraverso gli scalari $(y_1,…..,y_n)$ si abbia cioè

$$v = y_1e_1 +...+ y_n e_n$$

Da

$$v = x_1e_1 +...+ x_n e_n = y_1 e_1 +...+ y_n e_n$$

segue

$$(x_1 - y_1) e_1 + (x_2 - y_2) e_2 +…..+ (x_n - y_n) e_n = \underline{0}$$

e questa comporta , per l’indipendenza di e_1 , e_2,…., e_n ,

$$(x_1 - y_1) = (x_2 - y_2) =…..= (x_n - y_n) =0$$

e cioè

$$x_1 = y_1 , \quad x_2 = y_2,……, x_n = y_n$$

Pertanto gli **unici** scalari che danno luogo a v sono i numeri (x_1 , x_2,…., x_n) .

Quanto provato ci consente quindi di costruire una funzione tra V e K^n associando ad ogni vettore v di V la n-pla $(x_1, x_2, \dots, x_n)$ delle sue coordinate

$$f : v \in V \longrightarrow (x_1, x_2, \dots, x_n) \in K^n$$

Tale funzione , come è facile verificare, è biettiva e lineare e quindi è un **isomorfismo** tra V e K^n , detto *coordinazione di V nel riferimento fissato.*

Abbiamo così provato la seguente importante

Proposizione 4.1 *Ogni spazio vettoriale V sul campo K di dimensione finita n è isomorfo allo spazio vettoriale numerico K^n.*

Vediamo ora quali sono i vantaggi di aver acquisito un siffatto risultato.

Tali vantaggi appariranno chiari quando si siano provate alcune proprietà degli isomorfismi che ora andiamo ad illustrare nelle proposizioni che seguono. Da qui in avanti

$$f : V_n \longrightarrow W_m$$

è una **applicazione lineare** tra gli spazi vettoriali V_n e W_m costruiti sullo stesso campo K e dimensione finita n ed m rispettivamente.

Una proprietà notevole delle applicazioni lineari è espressa dal seguente

Teorema fondamentale. *Un applicazione lineare* $f : V_n \longrightarrow W_m$ *è determinata quando si conoscono i valori che essa assume sui vettori di una base ordinata* $(e_1, e_2, \dots, e_n)$ *di* V_n .

Dimostrazione. Sia quindi $(e_1, e_2, \dots, e_n)$ una base ordinata di V_n e supponiamo di conoscere i vettori immagine $f(e_1)$, $f(e_2)$, …,$f(e_n)$ Questa conoscenza ci permetterà di calcolare f su un qualunque vettore v di V_n . Infatti sia v un qualunque vettore di V_n .

Poiché $(e_1, e_2, \dots, e_n)$ è una base esistono n scalari $(\alpha_1, \alpha_2, \dots, \alpha_n)$ per cui si abbia

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$$

Applicando f a tale relazione, tenendo conto della sua linearità, si ha :

$$f(v)= \alpha_1 f(e_1)+ \alpha_2 f(e_2) + \dots + \alpha_n f(e_n)$$

la quale prova l'asserto.

Proposizione 4.2 *Una applicazione lineare f tra V e W trasforma il vettore nullo di V nel vettore nullo di W. Ne segue che se f è iniettiva in particolare se f è un isomorfismo esso trasforma altresì un vettore non nullo di V in un vettore non nullo di W.*

Dimostrazione. Sia v un vettore qualunque di V , per la linearità di f risulta

$$f(v) = f(v + \underline{0})= f(v) + f(\underline{0})$$

da cui segue ovviamente $f(\underline{0}) = \underline{0}$.

Se f è iniettiva ed è $v \neq \underline{0}$ allora $f(v) \neq f(\underline{0}) = \underline{0}$

Proposizione 4.3 *Una applicazione lineare f tra V e W trasforma vettori dipendenti di V in vettori dipendenti di W. Inoltre se f è iniettiva essa trasforma altresì vettori indipendenti di V in vettori indipendenti di W. Un isomorfismo conserva pertanto con la sua inversa la dipendenza e l'indipendenza lineare in quanto trasforma vettori dipendenti di V in vettori dipendenti di W e trasforma vettori indipendenti di V in vettori indipendenti di W.*

Dimostrazione. Siano $v_1, v_2, \dots, v_h$, h vettori dipendenti di V. Poiché i vettori $v_1, v_2, \dots, v_h$, sono dipendenti esistono scalari $(\alpha_1, \alpha_2, \dots \alpha_h)$ non tutti nulli per cui risulti :

$$\alpha_1 v_1+ \alpha_2 v_2 + \dots + \alpha_h v_h = \underline{0}$$

Applicando f ad ambo i membri, tenendo conto della linearità e della proposizione 4.2 si ha

$$\alpha_1 f(v_1)+ \alpha_2 f(v_2) + \dots + \alpha_h f(v_h) = f(\underline{0}) = \underline{0}$$

la quale mostra che anche i vettori trasformati $f(v_1) , f(v_2) , \dots, f(v_h)$ sono dipendenti. Supponiamo f iniettiva e siano $v_1, v_2, \dots, v_h$, h vettori indipendenti di V. Dobbiamo provare che anche i vettori $f(v_1) , f(v_2) , \dots, f(v_h)$ sono indipendenti. Supponiamo quindi che $(\alpha_1, \alpha_2, \dots \alpha_h)$ siano scalari con i quali si abbia

$$\alpha_1 f(v_1)+ \alpha_2 f(v_2) + \dots + \alpha_h f(v_h) = \underline{0}$$

e vediamo se tali scalari sono tutti nulli.

Tale relazione per la linearità di f equivale a

$$f(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h) = \underline{0}$$

Poiché f è iniettiva l'unico vettore che si trasforma nel vettore nullo di W è il vettore nullo di V e così è :

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{0}$$

Per la supposta indipendenza dei vettori $v_1, v_2, \ldots, v_h$ si ha allora $\alpha_1 = \alpha_2 = \ldots \alpha_h = 0$ e ciò prova che i vettori $f(v_1), f(v_2), \ldots, f(v_h)$ sono indipendenti.

Sia

$$f : V \longrightarrow W$$

una applicazione lineare tra gli spazi vettoriali V e W .

Possiamo considerare i sottoinsiemi di V e W seguenti:

$$N = \{ v \in V \ , \ f(v) = \underline{0} \}$$

$$T = f(V) = \{ f(v) \ , v \in V \}$$

Evidentemente N è non vuoto perché di esso fa parte il vettore nullo ed è un **sottospazio** di V , detto il ***nucleo*** dell'applicazione f mentre T è un **sottospazio** di W ed è detto lo ***spazio immagine*** di f .

Tali sottospazi, nucleo ed immagine, possono essere usati per valutare l'iniettività e suriettività della funzione f.

Infatti la funzione f è suriettiva se solo se risulta T = W.

Inoltre

Proposizione 4.4 *La funzione f è iniettiva se e solo se il suo nucleo è ridotto al vettore nullo***.**

Dimostrazione. Se f è iniettiva abbiamo già osservato che l'unico vettore che si trasforma nel vettore nullo è il vettore nullo e quindi è N={ $\underline{0}$ }. Viceversa supponiamo che sia N={ $\underline{0}$ }. Se per due vettori v e v' risulta f(v) = f(v') si ha per la linearità di f , f(v-v') = $\underline{0}$ e così v-v' $\in$ N. Ma per ipotesi è N={ $\underline{0}$ } e quindi v - v' = $\underline{0}$ e cioè è v = v'. Pertanto f è iniettiva .

Utile per ciò che segue è la seguente proposizione:

Proposizione 4.5. *Sia $f : V \longrightarrow W$ una funzione lineare tra gli spazi vettoriali V e W. Se e_1 , e_2,...., e_n sono generatori di V i vettori $f(e_1)$, $f(e_2)$,..., $f(e_n)$ sono un sistema di generatori per lo spazio immagine T = f(V). In particolare se f è un isomorfismo ed (e_1 , e_2,...., e_n) è una **base** di V allora ($f(e_1)$, $f(e_2)$,..., $f(e_n)$) è una **base** di W.*

Dimostrazione. Sia w un qualunque vettore di T= f(V) . Esiste allora un vettore v in V per cui sia w = f(v). Poiché i vettori e_1 , e_2,…., e_n generano V si ha per opportuni scalari $\alpha_1, \alpha_2, \ldots \alpha_n$

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$$

Si ha allora

$$w = f(v) = f(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n) = \alpha_1 f(e_1) + \alpha_2 f(e_2) + \ldots + \alpha_n f(e_n)$$

la quale mostra che i vettori $f(e_1)$, $f(e_2)$,..., $f(e_n)$ sono un sistema di generatori per lo spazio T. La parte finale della proposizione è ovvia ricordando che un isomorfismo trasforma vettori indipendenti in vettori indipendenti.

I sottospazi N e T nucleo ed immagine si *"condizionano"* a vicenda come mostra la seguente

Proposizione 4.6. *Sia $f : V \longrightarrow W$ una funzione lineare tra gli spazi vettoriali V e W. Siano N e T gli spazi nucleo ed immagine di f. Detta n la dimensione di V ,risulta*

$$(*) \qquad \dim N + \dim T = n$$

Dimostrazione. La proprietà (*) è ovvia se la funzione è iniettiva cioè se è N = { $\underline{0}$ }ed è altrettanto vera se è N = V. Supponiamo quindi N non banale e sia h = dimN . Scelti h vettori e_1 , e_2,…., e_h indipendenti in N cioè una sua base aggiungiamo ad essi altri n-h vettori di V, $v_{h+1}, v_{h+2}, \ldots, v_n$ in modo che i vettori e_1 , e_2,…., e_h $v_{h+1}, v_{h+2}, \ldots, v_n$ siano una base di V.

Se ora mostriamo che i vettori $f(v_{h+1}), f(v_{h+2}), ..., f(v_n)$ sono una base per T si ha dimT= n-h e quindi la (*).Cominciamo a provare che sono indipendenti.Siano $\alpha_{h+1}, \alpha_{h+2}, \ldots, \alpha_n$ scalari per i quali risulti

$$\alpha_{h+1} f(v_{h+1}) + \alpha_{h+2} f(v_{h+2}) + \ldots + \alpha_n f(v_n) = \underline{0}$$

Per la linearità di f la relazione scritta equivale a

$$f(\alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n) = \underline{0}$$

la quale mostra che il vettore

$$\alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n$$

appartiene al nucleo. Si ha quindi , per opportuni scalari $\alpha_1, \alpha_2, \ldots \alpha_h$

$$\alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n = \alpha_1 e_1 + \alpha_2 e_2, \ldots + \alpha_h e_h$$

Da questa relazione segue

$$\alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n - \alpha_1 e_1 - \alpha_2 e_2, \ldots - \alpha_h e_h = \underline{0}$$

e quindi per l'indipendenza dei vettori $e_1, e_2, \ldots, e_h$ $v_{h+1}, v_{h+2}, ..., v_n$ si ha come si voleva $\alpha_{h+1} = \alpha_{h+2} = \ldots = \alpha_n = 0$. Proviamo infine che sono un sistema di generatori per T. Sia w un qualunque vettore di T= f(V) . Esiste allora un vettore v in V per cui sia w = f(v). Poiché i vettori $e_1, e_2, \ldots, e_h$ $v_{h+1}, v_{h+2}, ..., v_n$ sono una base di V si ha per opportuni scalari $\alpha_1, \alpha_2, \ldots \alpha_n$

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_h e_h + \alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n$$

da cui segue, tenendo conto della linearità di f e del fatto che i vettori $e_1, e_2, \ldots, e_h$ sono nel nucleo

$$w = f(v) = f(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_h e_h + \alpha_{h+1} v_{h+1} + \alpha_{h+2} v_{h+2} + \ldots + \alpha_n v_n) =$$
$$= \alpha_{h+1} f(v_{h+1}) + \alpha_{h+2} f(v_{h+2}) + \ldots + \alpha_n f(v_n)$$

la quale mostra che i vettori $f(v_{h+1}), f(v_{h+2}), \ldots, f(v_n)$ sono un sistema di generatori per lo spazio T.

Possiamo concludere tale numero provando la seguente importante

Proposizione 4.7 *Due spazi vettoriali V e W costruiti sullo stesso campo e di dimensione finita sono isomorfi se e solo se essi hanno la stessa dimensione.*

Dimostrazione. Se c'è un isomorfismo $f : V \longrightarrow W$ tra V e W abbiamo già visto che se $e_1, e_2, \ldots, e_n$ è una base di V allora $f(e_1), f(e_2), \ldots, f(e_n)$ è una base di W e quindi V e W hanno la stessa dimensione n. Viceversa supponiamo che V e W abbiano entrambi la stessa dimensione n . Se $(e_1, e_2, \ldots, e_n)$ è una base ordinata di V come già visto, associando ad ogni vettore $v = x_1 e_1 + \ldots + x_n e_n$ di V la n-pla $(x_1, x_2, \ldots, x_n)$ delle sue coordinate

$$f : v \in V \longrightarrow (x_1, x_2, \ldots, x_n) \in K^n$$

si realizza un isomorfismo tra V e K^n . Analogamente se $(w_1, w_2, \ldots, w_n)$ è una base ordinata di W , associando ad ogni vettore $w = y_1 w_1 + \ldots + y_n w_n$ di W la n-pla $(y_1, y_2, \ldots, y_n)$ delle sue coordinate

$$g : w \in W \longrightarrow (y_1, y_2, \ldots, y_n) \in K^n$$

si realizza un isomorfismo tra W e K^n . L' applicazione

$$g^{-1} \circ f : V \longrightarrow W$$

essendo una funzione composta da isomorfismi è allora un isomorfismo tra V e W.

Come conseguenza di questo teorema possiamo allora valutare la dimensione degli spazi vettoriali illustrati negli esempi II e III del n.3. Avendo già osservato che $K[x_1,x_2,..,x_n]$ è isomorfo a K^{n+1} e che $M_{m,n}(R)$ è isomorfo a K^{mn} si ha per quanto ora provato che

$$\dim K[x_1,x_2,..,x_n] = n+1 \qquad e \qquad \dim M_{m,n}(R) = mn.$$

Concludiamo tale numero con una proposizione di cui faremo un grande uso nelle applicazioni successive.

Proposizione 4.8 *Sia $f : V \longrightarrow W$ un isomorfismo tra gli spazi vettoriali V e W. Un vettore v di V è combinazione lineare dei vettori $v_1, v_2, \ldots, v_h$ se e solo se il vettore $f(v)$ è combinazione lineare dei vettori trasformati $f(v_1), f(v_2) \ldots, f(v_h)$.*

Dimostrazione. Supponiamo che v sia combinazione lineare dei vettori $v_1, v_2, \ldots, v_h$ si abbia cioè $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h$ Applicando f e tenendo conto della sua linearità si ha che è $f(v) = \alpha_1 f(v_1) + \alpha_2 f(v_2) + \ldots + \alpha_h f(v_h)$.

Viceversa supponiamo che il vettore $f(v)$ sia combinazione lineare dei vettori $f(v_1), f(v_2) \ldots, f(v_h)$ si abbia cioè

$$f(v) = \alpha_1 f(v_1) + \alpha_2 f(v_2) + \ldots + \alpha_h f(v_h).$$

Tale relazione per la linearità di f equivale a

$$f(v) = f(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h)$$

e questa per la iniettività di f comporta

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h \quad .$$

5. I sottospazi di uno spazio vettoriale.

In questo numero V è uno spazio vettoriale di dimensione finita n costruito su un campo K. Sia $\mathcal{H}$ la famiglia di tutti i suoi sottospazi . La famiglia $\mathcal{H}$ ha le seguenti proprietà , di facile dimostrazione.

1. *per ogni* $H, T \in \mathcal{H}$,

$$H \subseteq T \;\Rightarrow\; \dim H \le \dim T$$
$$H \subset T \;\Rightarrow\; \dim H < \dim T$$

2. *l'intersezione di una famiglia di sottospazi è un sottospazio.*

In generale l'unione di sottospazi non è un sottospazio. Mostriamo ciò con un esempio.

Nello spazio vettoriale $V = \mathbf{R}^2$ delle coppie ordinate di numeri reali si considerino i due seguenti sottoinsiemi H e T .

$H = \{(a, 0), \ a \in \mathbf{R}\}$

$T = \{(0, b), \ b \in \mathbf{R}\}$

Facilmente si riconosce che H e T sono sottospazi mentre il sottoinsieme $X = H \cup T$ non è un sottospazio risultando ad esempio
$(2, 0) \in X$, $(0, 5) \in X$ mentre $(2,0) + (0,5) = (2 , 5) \notin X$.

La proprietà 2 consente la seguente definizione. Sia X un sottoinsieme di V. Si chiama ***sottospazio generato da X*** il sottospazio [X] che si ottenga intersecando tra loro tutti i sottospazi che contengono X.Tale sottospazio è ovviamente il più *piccolo sottospazio* (rispetto all'inclusione) che contiene X.

Ha interesse considerare lo spazio X quando X sia l'unione di due sottospazi H e T . Mostreremo ora che tale sottospazio $[H \cup T]$ che si ottiene intersecando tra loro tutti i sottospazi che contengono H e T coincide col seguente sottoinsieme di V

$$L = \{ a + b \ , \ a \in H \ , \ b \in T\}$$

il quale contiene tutti i vettori che si ottengono sommando tra loro un vettore di H ed un vettore di T. Evidentemente L è un sottospazio di V.

Quando si scelga $b = \underline{0}$ e si faccia variare a in H si riconosce che tra i vettori di L ci sono in particolare tutti quelli di H. Analogamente quando si scelga $a = \underline{0}$ e si faccia variare b in T si riconosce che tra i vettori di L ci sono in particolare tutti quelli di T. Pertanto L contiene sia H che T. Inoltre se un sottospazio J contiene H e T allora contiene anche tutti i vettori $a + b$ con $a \in H$ e $b \in T$ e quindi contiene L. L è pertanto il più piccolo sottospazio che contiene H e T e quindi coincide con lo spazio $[H \cup T]$ da essi generato.Lo spazio $[H \cup T]$ viene anche indicato col simbolo $H \underline{\cup} T$ o $H + T$.

La proprietà ora provata per i sottospazi H e T può essere estesa facilmente ad un numero finito di t sottospazi $H_1 , H_2 , \ldots, H_t$ con $t > 2$. Indicando con $H_1 + H_2 + \ldots + H_t$ lo spazio generato da $H_1 \cup H_2 \cup \ldots \cup H_t$ si prova facilmente che esso coincide col sottospazio L seguente

$$L = \{ a_1 + a_2 + .. + a_t \ , \ a_1 \in H_1 , a_2 \in H_2 , \ldots, a_t \in H_t \}$$

Proviamo infine la seguente importante proprietà :

3. *per ogni H , T* $\in \mathcal{H}$ *si ha* (***formula di Grassmann***).

$$\text{(3.1)} \qquad \dim H + \dim T = \dim (H \cap T) + \dim (H + T)$$

Dimostrazione. Possiamo supporre che nessuno dei due sottospazi H e T sia contenuto nell'altro altrimenti la (3.1) è ovvia.

Poniamo $h = \dim H$ e $t = \dim T$ e siano $(e_1, e_2, \dots, e_h)$ e $(w_1, w_2, \dots, w_t)$ rispettivamente una base di H e una di T. Mostriamo che se $H \cap T = \{ \underline{0} \}$ allora i vettori $(e_1, e_2, \dots, e_h \; w_1, w_2, \dots, w_t)$ costituiscono una base di H+T e così la (3.1) è provata .

Un vettore v di H + T è del tipo $v = a + b$ con $a \in H$ e $b \in T$. Essendo $(e_1, e_2, \dots, e_h)$ e $(w_1, w_2, \dots, w_t)$ basi di H e T si ha :

$$v = \underline{a} + \underline{b} = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_h e_h + \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_t w_t$$

la quale mostra che i vettori $(e_1, e_2, \dots, e_h \; w_1, w_2, \dots, w_t)$ sono un sistema di generatori per H+T. Se essi risultano altresì indipendenti allora sono una base. Siano quindi $\alpha_1, \alpha_2, \dots, \alpha_h$, $\beta_1, \beta_2, \dots, \beta_t$ scalari tali che risulti

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_h e_h + \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_t w_t = \underline{0} .$$

Da questa segue

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_h e_h = -(\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_t w_t).$$

Posto $a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_h e_h$ e $b = \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_t w_t$

Ovviamente $a \in H$ e $b \in T$. Inoltre $-b \in T$. Da $a = -b$ segue allora che $a \in H \cap T$ e $b \in H \cap T$. Ma essendo per ipotesi $H \cap T = \{ \underline{0} \}$ si ha $a = \underline{0}$ e $b = \underline{0}$. Ma

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_h e_h = \underline{0} \quad \text{e} \quad b = \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_t w_t = \underline{0}$$

comportano essendo $(e_1, e_2, \dots, e_h)$ e $(w_1, w_2, \dots, w_t)$ vettori indipendenti

$$\alpha_1 = \alpha_2 = \dots = \alpha_h = \beta_1 = \beta_2 = \dots = \beta_t = 0$$

come si voleva provare.

Supponiamo quindi che sia $H \cap T$ diverso dal vettore nullo e sia $i = \dim H \cap T$. Siano $(e_1, e_2, \dots, e_i)$ vettori indipendenti di $H \cap T$ cioè una base di $H \cap T$. Usando la proposizione 3.7 si possono trovare h -i vettori $v_{i+1}, v_{i+2}, \dots, v_h$ di $H - (H \cap T)$, scelti in modo che $(e_1, e_2, \dots, e_i \; v_{i+1}, v_{i+2}, \dots, v_h)$ sia una base di H e si possono trovare t -i vettori $w_{i+1}, w_{i+2}, \dots, w_t$ di $T - H \cap T$ scelti in modo che $(e_1, e_2, \dots, e_i \; w_{i+1}, w_{i+2}, \dots, w_t)$ sia una base di T. Se proviamo che i vettori $(e_1, e_2, \dots, e_i \; v_{i+1}, v_{i+2}, \dots, v_h \; w_{i+1}, w_{i+2}, \dots, w_t)$ sono una base per H + T , essendo in numero di $i + h - i + t - i = h + t - i$ si ha

$$\dim (H + T) = h + t - i = \dim H + \dim T - \dim H \cap T$$

e cioè la (3.1).

Sia v un vettore di $H+T$. Il vettore v è del tipo $v = a + b$ con $a \in H$ e $b \in T$. Essendo $(e_1 , e_2,\ldots, e_i\ v_{i+1} , v_{i+2} ,\ldots , v_h)$ una base di H e $(e_1 , e_2,\ldots, e_i\ w_{i+1} , w_{i+2} ,\ldots , w_t)$ una base di T si ha

$$v = a + b = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_i e_i + \alpha_{i+1} v_{i+1} + \alpha_{i+2} v_{i+2} + \ldots + \alpha_h v_h + \gamma_1 e_1 + \gamma_2 e_2 + \ldots + \\ + \gamma_i e_i + \beta_{i+1} w_{i+1} + \beta_{i+2} w_{i+2} + \ldots + \beta_t w_t$$

la quale mostra che v è combinazione lineare dei vettori $(e_1 , e_2,\ldots, e_i\ v_{i+1} , v_{i+2} ,\ldots , v_h\ w_{i+1} , w_{i+2} ,\ldots , w_t)$ i quali sono quindi un sistema di generatori per $H + T$. Mostriamo che sono indipendenti . Siano $\alpha_1 ,\alpha_2 ,\ldots , \alpha_h\ \beta_{i+1},\beta_{i+2} , \ldots, \beta_t$ scalari tali che risulti

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_i e_i + \alpha_{i+1} v_{i+1} + \alpha_{i+2} v_{i+2} + \ldots + \alpha_h v_h + \beta_{i+1} w_{i+1} + \beta_{i+2} w_{i+2} + \ldots + \beta_t w_t = \underline{0}$$

Da questa segue

(**) $\beta_{i+1}w_{i+1} + \beta_{i+2} w_{i+2} + \ldots + \beta_t w_t = - (\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_i e_i + \alpha_{i+1} v_{i+1} + \alpha_{i+2} v_{i+2} + \ldots + \alpha_h v_h)$

Poniamo

$$b = \beta_{i+1}w_1 + \beta_{i+2} w_2 + \ldots + \beta_t w_t \qquad \text{ed}$$

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_i e_i + \alpha_{i+1} v_{i+1} + \alpha_{i+2} v_{i+2} + \ldots + \alpha_h v_h$$

Ora $b \in T$ ed $a \in H$. Quindi dall'essere $b = -a$ segue che $b \in H \cap T$.

Poiché $e_1 , e_2 \ldots e_i$ è una base di $H \cap T$. si ha , per opportuni scalari $\delta_1 , \delta_2 , \ldots, \delta_i$

$$b = \beta_{i+1}w_{i+1} + \beta_{i+2} w_{i+2} + \ldots + \beta_t w_t = \delta_1 e_1 + \delta_2 e_2 + \ldots + \delta_i e_i$$

da questa segue

$$\delta_1 e_1 + \delta_2 e_2 + \ldots + \delta_i e_i - \beta_{i+1}w_{i+1} - \beta_{i+2} w_{i+2} - \ldots - \beta_t w_t = \underline{0}$$

e questa comporta, essendo $(e_1, e_2, \ldots, e_i\ w_{i+1}, w_{i+2}, \ldots, w_t)$ una base,

$$\delta_1 = \delta_2 = \ldots = \delta_i = \beta_{i+1} = \beta_{i+2} = \ldots = \beta_t = 0$$

Ma se $\beta_{i+1} = \beta_{i+2} = \ldots = \beta_t = 0$ da (**) segue

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_i e_i + \alpha_{i+1} v_{i+1} + \alpha_{i+2} v_{i+2} + \ldots + \alpha_h v_h = \underline{0}$$

dalla quale segue, essendo $(e_1, e_2, \ldots, e_i\ v_{i+1}, v_{i+2}, \ldots, v_h)$ una base,

$$\alpha_1 = \alpha_2 = \ldots = \alpha_i = \alpha_{i+1} = \alpha_{i+2} = \ldots = \alpha_h = 0 .$$

Avendo provato che $\alpha_1 = \alpha_2 = \ldots = \alpha_h = \beta_{i+1} = \beta_{i+2} = \ldots = \beta_t = 0$ i vettori

$(e_1, e_2, \ldots, e_i\ v_{i+1}, v_{i+2}, \ldots, v_h\ w_{i+1}, w_{i+2}, \ldots, w_t)$ sono indipendenti e l'asserto è provato.

Osserviamo esplicitamente che nella dimostrazione ora fatta abbiamo provato che se H e T hanno in comune il solo vettore nullo unendo una base di H ad una base di T si ottiene un insieme di vettori linearmente indipendente , base per lo spazio generato da H e T.

Tale proprietà con un semplice processo di induzione sul numero t di sottospazi può essere così generalizzata .

Proposizione 5.1 *Siano $H_1, H_2, \ldots, H_t$, t $(t \geq 2)$ sottospazi di uno spazio vettoriale V_n di dimensione n sul campo K. Se ogni H_i interseca nel solo vettore nullo lo spazio generato dai rimanenti sottospazi allora unendo una base di H_1, con una base di H_2 e .. con una di H_t si ottiene un insieme di vettori linearmente indipendente , base per lo spazio $H_1 + H_2 + \ldots + H_t$ generato da $H_1 \cup H_2 \cup \ldots \cup H_t$.*

Concludiamo tale numero introducendo una utile nozione. Due sottospazi H e T di uno spazio vettoriale V_n sono detti *supplementari* se risulta

$$H \cap T = \{ \underline{0} \} \qquad \text{e} \qquad H + T = V_n$$

Per la formula di Grassmann già provata , se H e T sono supplementari risulta :

$$\dim H + \dim T = n$$

Si possono costruire sottospazi supplementari ?
La proposizione che segue da risposta al quesito posto.

Proposizione 5.2 *Sia* $e_1, e_2, \dots, e_n$ *una base di* V_n *.I sottospazi*

$$H = [\, e_1, e_2, \dots, e_t \,] \quad , \quad T = [\, e_{t+1}, e_{t+2}, \dots, e_n \,]$$

generati rispettivamente da $e_1, e_2, \dots, e_t$ *e da* $e_{t+1}, e_{t+2}, \dots, e_n$ *sono tra loro supplementari.*

Dimostrazione. Poiché $\{ e_1, e_2, \dots, e_n \}$ è una base di V_n allora per ogni vettore $\underline{v}$ di V_n si ha :

(*) $$\underline{v} = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t + \alpha_{t+1} e_{t+1} + \alpha_{t+2} e_{t+2} + \dots + \alpha_n e_n$$

Posto $a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t$ e $b = \alpha_{t+1} e_{t+1} + \alpha_{t+2} e_{t+2} + \dots + \alpha_n e_n$ la relazione (*) mostra che è $v = a + b$ con $a \in H$ e $b \in T$ e che quindi è $H + T = V_n$ .Proviamo infine che è anche $H \cap T = \{ \underline{0} \}$. Sia v un vettore di $H \cap T$. Poichè $e_1, e_2, \dots, e_t$ è una base di H e $e_{t+1}, e_{t+2}, \dots, e_n$ è una base di T, si ha per opportuni scalari $\alpha_1, \alpha_2 \dots \alpha_t$, $\alpha_{t+1}, \alpha_{t+2} \dots \alpha_n$

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t = \alpha_{t+1} e_{t+1} + \alpha_{t+2} e_{t+2} + \dots + \alpha_n e_n$$

Da questa segue

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t - \alpha_{t+1} e_{t+1} - \alpha_{t+2} e_{t+2} + \dots - \alpha_n e_n = \underline{0}$$

e questa comporta , essendo i vettori $e_1, e_2, \dots, e_n$ indipendenti, $\alpha_1 = \alpha_2 = \alpha_t = \alpha_{t+1} = \alpha_{t+2} = \dots = \alpha_n = 0$.
Si ha quindi $v = \underline{0}$ e ciò mostra per l'arbitrarietà di v in $H \cap T$ che è $H \cap T = \{ \underline{0} \}$.

La proposizione che segue "inverte" in un certo senso la proposizione ora provata

Proposizione 5.3 *Siano H e T due sottospazi supplementari di dimensioni t e $n-t$. Se $\{ e_1, e_2, \dots, e_t \}$ è una base di H e $\{ e_{t+1}, e_{t+2} \dots, e_n \}$ è una base di T allora $\{ e_1, e_2, \dots, e_t\ e_{t+1}, e_{t+2} \dots, e_n \}$ è una base di V_n.*

Dimostrazione . Per la dimostrazione è sufficiente provare che i vettori $e_1, e_2, \dots, e_t, e_{t+1}, e_{t+2} \dots, e_n$ sono indipendenti . Siano quindi $\alpha_1, \alpha_2, \dots \alpha_t, \alpha_{t+1}, \alpha_{t+2} \dots, \alpha_n$ scalari per i quali risulti

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t + \alpha_{t+1} e_{t+1} + \alpha_{t+2} e_{t+2} + \dots + \alpha_n e_n = \underline{0}$$

Da questa segue

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t = -\alpha_{t+1} e_{t+1} - \alpha_{t+2} e_{t+2} - \dots - \alpha_n e_n$$

Posto

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t \quad \text{e} \quad b = -\alpha_{t+1} e_{t+1} - \alpha_{t+2} e_{t+2} - \dots - \alpha_n e_n$$

si ha che $a \in H$ e $b \in T$ ma da $a = b$ segue allora che è $a = b \in H \cap T$. Ma essendo H e T supplementari si ha $H \cap T = \{ \underline{0} \}$ e quindi è $a = b = \underline{0}$.

Se il vettore

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_t e_t = \underline{0}$$

si ha per l'indipendenza di $e_1, e_2, \dots, e_t$, $\alpha_1 = \alpha_2 = \dots = \alpha_t = 0$

e da

$$b = -\alpha_{t+1} e_{t+1} - \alpha_{t+2} e_{t+2} - \dots - \alpha_n e_n = \underline{0}$$

si ha per l'indipendenza di $e_{t+1}, e_{t+2} \dots, e_n$ che è $\alpha_{t+1} = \alpha_{t+2} = \dots = \alpha_n = 0$ e ciò prova l'asserto.

Quando si tenga conto della proposizione 5.2 e della proposizione 3.7 del capitolo I è facile verificare che sussiste la seguente proposizione

Proposizione 5.4. *Sia H un sottospazio dello spazio vettoriale V_n. E' sempre possibile costruire un sottospazio T supplementare di H .*

Concludiamo con una proprietà importante degli spazi supplementari.

Se H e T sono supplementari sappiamo che è per definizione $V_n = H + T$ e che quindi ogni vettore v dello spazio si ottiene come somma di un vettore a di H e di un vettore b di T . Noi vogliamo mostrare che la decomposizione di $\underline{v}$ come somma di un vettore a di H e di un vettore b di T è unica . Si ha cioè :

se $v = a + b$ e $v = a' + b'$ allora è $a = a'$ e $b = b'$.

Infatti da $a + b = a' + b'$ segue $a - a' = b' - b$ e questa comporta che $a - a' \in H \cap T$ e $b' - b \in H \cap T$. Ma poiché è $H \cap T = \{ \underline{0} \}$ si ha $a - a' = \underline{0}$ e $b' - b = \underline{0}$ e quindi come si voleva $a = a'$ e $b' = b$.

A titolo di esempio si consideri lo spazio vettoriale R^2 delle coppie ordinate di numeri reali. In tale spazio i sottospazi $H = \{ (a, 0) , \underline{a} \in R \}$ e $T = \{ (0 , b) , \underline{b} \in R \}$ sono supplementari ed in accordo con la proprietà sopra illustrata ogni vettore (a, b) di R^2

$$(a , b) = (a , 0) + (0 , b)$$

si scrive in un sol modo come somma di un vettore di H e di uno di T.

C A P I T O L O II

Matrici e determinanti

1. Introduzione.

Abbiamo visto al capitolo precedente che se in uno spazio vettoriale V di dimensione n e costruito sul campo K si fissa un riferimento $(e_1 , e_2,\dots, e_n)$ è possibile costruire un isomorfismo tra V e K^n. Tale isomorfismo associa ad ogni vettore $v= x_1 e_1 +...+ x_n e_n$ di V la n-pla $(x_1 , x_2,\dots, x_n)$ delle sue coordinate nel riferimento scelto.

Poichè un isomorfismo conserva la dipendenza e l'indipendenza lineare , allora stabilire se h vettori $v_1 , v_2,\dots, v_h$ di V siano dipendenti o indipendenti equivale a spostare tale indagine sugli h vettori numerici delle loro coordinate. Ma c'è un metodo "semplice e veloce" per stabilire se tali vettori numerici sono dipendenti o indipendenti ? La risposta a questa domanda viene data introducendo la nozione di **determinante** di una matrice quadrata , nozione di cui ora parleremo.

Preliminarmente è però essenziale fare le seguenti osservazioni.

Siano assegnati un certo numero di vettori numerici di ordine n e siano .

$$(a_1,..., a_n) ,\ (b_1,..., b_n), \dots\dots, (w_1,...., w_n) \quad .$$

Se supponiamo che essi siano dipendenti , esisteranno opportuni scalari α, $\beta,\dots, \gamma$ *non tutti nulli* tali che risulti :

$$\alpha\, (a_1,..., a_n) + \beta\, (b_1,..., b_n) + \dots\dots + \gamma\, (w_1,...., w_n) = (0,\dots,0).$$

Fissati s posti $(i_1,i_2,...,i_s)$ tra i posti da 1 ad n si considerino i vettori numerici di ordine s ottenuti considerando in ognuno dei vettori numerici assegnati solo le componenti di posto

$i_i,i_2,..,i_s$, cioè consideriamo i vettori ,ora di lunghezza s , seguenti

$$(a_{i1},a_{i2},...,a_{is}) \quad (b_{i1},b_{i2},...,b_{is}) \quad \dots\ (w_{i1},w_{i2},...,w_{is})$$

E' ovvio che anche essi sono dipendenti in quanto con gli stessi scalari α, $\beta,\dots, \gamma$ non tutti nulli sopra adottati si ha ancora :

$$\alpha\, (a_{i1},a_{i2},...,a_{is}) + \beta\, (b_{i1},b_{i2},...,b_{is}) + \dots + \gamma\, (w_{i1},w_{i2},...,w_{is}) = (0,0 , \dots,0).$$

Usando un linguaggio poco preciso ma espressivo possiamo riassumere l'osservazione fatta dicendo che : se si " ***accorciano*** " *dei vettori numerici dipendenti essi restano dipendenti* .

Ne segue che se si " ***allungano*** " *dei vettori numerici indipendenti essi restano indipendenti.*

Mostriamo quanto detto con un esempio . Nello spazio V = $\mathbf{R}^4$ delle quaterne ordinate di numeri reali i vettori (1 , 0 , 2 , 1) , (0 , 3 , 1 , 1) , (4 , 3 , 9 , 5) sono dipendenti risultando :

$$8 (1 , 0 , 2 , 1) + 2 (0 , 3 , 1 , 1) - 2 (4 , 3 , 9 , 5) = (0, 0, 0, 0)$$

Le tre coppie (2 ,1) , (1, 1) , (9, 5) ottenute “accorciando” le tre quaterne date (considerando di ognuna solo gli ultimi due numeri) sono anch’ esse dipendenti risultando ancora

$$8 (2 , 1) + 2 (1 , 1) - 2 (9 , 5) = (0, 0)$$

Sempre nello stesso spazio le due quaterne (1 , 0 , 2 , 1) , (0 , 3 , 1 , 1) sono indipendenti in quanto non proporzionali e così le due sestine

(1 , 0 , 2 , 1 , **2 , 2**) , (0 , 3 , 1 , 1 , **7 , 5**)

ottenute aggiungendo ad ognuna di esse ulteriori due numeri sono ancora indipendenti.

Diamo ora la nozione di determinante di una matrice quadrata.

2 . Determinante di una matrice quadrata.

Ad una matrice **quadrata** A d’ordine n ad elementi in un campo K ,si può associare uno scalare , elemento di K , detto ***determinante di A***, e denotato con $|A|$ o detA al seguente modo :

Se n=1 e cioè è A = (a) allora si pone detA = a .

Vediamo come si calcola il determinante quando è $n \geq 2$.

Sia quindi A una matrice quadrata d’ordine n con $n \geq 2$. Indichiamo la matrice A al seguente modo :

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} \end{pmatrix}$$

Se a_{ij} è un elemento della matrice A , chiameremo ***complemento di*** $\boldsymbol{a_{ij}}$ il determinante della matrice quadrata d'ordine n-1 ottenuta cancellando in A la riga i e la colonna j .Tale complemento quando venga moltiplicato per $(-1)^{i+j}$ è detto ***complemento algebrico di*** $\boldsymbol{a_{ij}}$ e viene denotato col simbolo A_{ij} . Si può provare che si ottiene lo stesso elemento di K , sia che si esegua questo calcolo

$$a_{i1} A_{i1} + a_{i2} A_{i2} + \ldots\ldots + a_{in}A_{in}$$

lungo una qualunque riga i = 1,2,….,n di A e sia che si esegua questo calcolo

$$a_{1j} A_{1j} + a_{2j} A_{2j} + \ldots\ldots + a_{nj}A_{nj}$$

lungo una qualunque colonna j = 1,2,…,n di A . Tale scalare viene chiamato ***determinante*** di A . Riepilogando , per calcolare il determinante di A , si deve scegliere una riga o una colonna e poi eseguire ***la somma dei prodotti degli elementi della linea scelta per i rispettivi complementi algebrici .***

Ora se la matrice A ha ordine due se n = 2 e cioè è

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

si ha $A_{11} = a_{22}$ ed $A_{12} = - a_{21}$

e pertanto risulta :

$$\det A = a_{11} A_{11} + a_{12} A_{12} = a_{11} a_{22} - a_{12} a_{21} .$$

Avendo definito il determinante di A quando A ha ordine uno siamo stati in grado di calcolare il determinante di A quando A ha ordine due . Ma allora sapremo calcolare il determinante anche quando A ha ordine tre in quanto i complementi algebrici dei suoi elementi si otterranno attraverso il calcolo di determinanti di matrici d'ordine due. Per le stesse ragioni, sapendo calcolare il determinante di A quando essa ha ordine tre sapremo calcolare il determinante di A anche quando ha ordine quattro e così via. Sappiamo quindi calcolare il determinante di A quando essa ha ordine n se sappiamo calcolare il determinante di una matrice d'ordine n-1.

A titolo di esempio si voglia calcolare il determinante della matrice C reale d'ordine tre seguente :

$$C = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \\ 4 & 2 & 8 \end{pmatrix}$$

Sviluppando il calcolo lungo la seconda riga si ha :

$$\det C = 2\,(-18) + 4\,(6) = -12 .$$

E' utile osservare che nel calcolo del determinante della matrice C abbiamo scelto opportunamente la seconda riga in quanto su tale riga uno degli elementi è zero e ciò ha ridotto quindi il numero degli addendi da calcolare .

Se A è una matrice quadrata d'ordine n ed occorre calcolare il suo determinante è auspicabile che su qualche riga o colonna figurino molti zeri perché ciò riduce il numero di calcoli da effettuare. C'è una proprietà dei determinanti che viene incontro a questa nostra esigenza Enunceremo tale proprietà omettendo la sua dimostrazione .

Sussiste allo scopo la seguente :

Proposizione 2.1 *Il determinante di una matrice quadrata d'ordine n non cambia se si aggiunge ad una sua riga (o colonna) una combinazione lineare delle altre righe (o delle altre colonne).*

Per il calcolo del determinante della matrice

$$C = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \\ 4 & 2 & 8 \end{pmatrix}$$

possiamo applicare la proprietà ora enunciata per semplificarne il calcolo. A tale scopo possiamo sommare alla terza riga di C la seconda riga moltiplicata per -2 e si ottiene la seguente matrice B

$$B = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \\ 0 & 2 & 0 \end{pmatrix}$$

la quale per il teorema precedente ha lo stesso determinante di C . Sviluppando il determinante di B secondo l'ultima riga si ha detB = det C = 2 (-6) = - 12 .

E' evidente che se in una matrice una sua riga o colonna ha gli elementi tutti nulli allora il determinante di tale matrice è zero. Ciò premesso in forza della proposizione 2.1 si ha questa ovvia proprietà:

Proposizione 2.2 . *Se in una matrice due righe (o colonne) sono uguali o proporzionali allora il determinante è zero.*

I teoremi che seguono sono molto importanti in quanto chiariscono il perché sia molto utile saper calcolare questo scalare che abbiamo chiamato determinante . Proviamo la seguente

Proposizione 2.3 . *Se le righe (o colonne) di una matrice quadrata A sono vettori dipendenti allora risulta detA = 0 .*

Dimostrazione. Supponiamo che le righe di A siano dipendenti . Allora una di tali righe è combinazione lineare delle altre e supponiamo ad esempio che l'ultima riga sia combinazione lineare delle altre , risulti cioè $\underline{a}_n = \alpha_1\underline{a}_1 + \alpha_2\underline{a}_2 + \ldots + \alpha_{n-1}\underline{a}_{n-1}$. Se si somma ad $\underline{a}_n$ la seguente combinazione lineare delle altre $-\alpha_1\underline{a}_1 - \alpha_2\underline{a}_2 + \ldots - \alpha_{n-1}\underline{a}_{n-1}$ sappiamo che si ottiene una matrice B che ha lo stesso determinante di A . Ma l'ultima riga di B ha tutti gli elementi eguali a zero e quindi è detA=detB=0.

Dal teorema ora provato segue il seguente importante corollario:

Proposizione 2.4 . *Se il determinante di una matrice quadrata A è diverso da zero allora le righe e le colonne di A sono linearmente indipendenti .*

Vediamo qualche utile applicazione delle cose dette. Supponiamo si voglia stabilire se le tre quaterne (1,1,0,1) , (0,1,2,1) , (1,0,0,3) siano o meno linearmente indipendenti . Consideriamo la matrice A che si ottiene assumendo come sue righe le quaterne date .

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

In tale matrice la sottomatrice quadrata

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix}$$

costituita dalle prime tre colonne di A ha il determinante diverso da zero (infatti esso è due)

e quindi per la proposizione 2.4 le sue righe sono indipendenti . Poiché le righe di A sono un allungamento delle righe di B e poiché allungando vettori indipendenti essi restano indipendenti si conclude che le quaterne assegnate sono linearmente indipendenti.

Concludiamo tale numero con qualche utile osservazione. E' apparso chiaro dalla definizione che il calcolo del determinante di una matrice quadrata A d'ordine n non sia agevole quando l'ordine n è abbastanza grande. Ci sono però alcune matrici particolari per le quali il calcolo del determinante non presenta difficoltà. Vediamo.

Una matrice $A = (a_{ij})$ quadrata d'ordine n è detta ***triangolare*** (*bassa*) se per ogni $j > i$, risulta $a_{ij} = 0$ mentre è detta ***triangolare*** (*alta*) se per ogni $j < i$ è $a_{ij} = 0$.

$$A = \begin{pmatrix} a_{11} & 0 \ldots\ldots\ldots 0 \\ a_{21} & a_{22}\ 0 \ldots\ldots 0 \\ a_{31} & a_{32}\ a_{33}\ 0 \ldots 0 \\ .. & \\ a_{n1} & a_{n2} \ldots\ldots a_{nn} \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots a_{1n} \\ 0 & a_{22} \ldots\ldots a_{2n} \\ 0 & 0 \quad a_{33} \ldots a_{3n} \\ \ldots & \\ 0 & 0 \quad 0 \ldots 0 \ldots a_{nn} \end{pmatrix}$$

(triangolare bassa) *(triangolare alta)*

Gli elementi $(a_{11}, a_{22}, \ldots, a_{nn})$ della matrice A sono detti gli elementi della ***diagonale*** *(principale)* di A .

Una matrice $A (a_{ij})$ quadrata d'ordine n è detta ***diagonale*** se risulta per ogni $i \neq j$, $a_{ij} = 0$.

Con un semplice ragionamento di induzione si prova facilmente che se $A =(a_{ij})$ è una matrice quadrata d'ordine n ed A è triangolare o diagonale allora è

$$\det A = a_{11} . a_{22} \ldots . a_{nn}$$

cioè il suo determinante è il ***prodotto degli elementi della sua diagonale principale***.

3. Prodotto di matrici.

Siano $\underline{a} = (a_1 , a_2 ,\ldots\ldots , a_n)$, $\underline{b} = (b_1 , b_2 ,\ldots\ldots , b_n)$ due n-ple ordinate di elementi di un campo K .

Si definisce prodotto scalare di tali due n-ple, lo scalare

$$\underline{a} \times \underline{b} = a_1 b_1 + a_2 b_2 + \ldots\ldots a_n b_n$$

che si ottiene eseguendo *la somma dei prodotti degli elementi di egual posto*.

Siano $A = (a_{ij})$ e $B = (b_{ij})$ due matrici ad elementi nel campo K . La matrice A sia di tipo (m , n) e quella B di tipo (n , p) (quindi il numero di colonne di A deve esser eguale al numero di righe di B). In tale situazione si può definire prodotto delle due matrici (***righe per colonne***) la matrice C = AB di tipo (m , p) il cui generico elemento c_{ij} si ottiene eseguendo il prodotto scalare della riga i-sima di A con la colonna j-sima di B :

$$c_{ij} = \underline{a}_i \times \underline{b}^j$$

Sia ora $M = M_{n,R}$ lo spazio vettoriale di tutte le matrici quadrate d'ordine n sul campo K . Se A e B sono due elementi di M il loro prodotto (che è eseguibile) è ancora un elemento di M. Tale prodotto non ha la proprietà commutativa . Mostriamo ciò con un esempio.

Siano A e B le seguenti matrici quadrate d'ordine tre :

$$A = \begin{pmatrix} 1, & 1, 0 \\ 0, & 0, 0 \\ 0, & 0, 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1, & 0, 0 \\ 1, & 0, 0 \\ 0, & 0, 0 \end{pmatrix}$$

Eseguendo il loro prodotto si ha :

$$AB = \begin{pmatrix} 2, & 0, 0 \\ 0, & 0, 0 \\ 0, & 0, 0 \end{pmatrix}$$

Mentre risulta :

$$BA = \begin{pmatrix} 1, & 1, & 0 \\ 1, & 1, & 0 \\ 0, & 0, & 0 \end{pmatrix}$$

Nell'insieme M la seguente matrice I (detta ***matrice identica***)

$$I = (\delta_{ij}) \quad \text{con} \quad \delta_{ij} = 0 \ \text{se} \ i \neq j \quad \text{e} \quad \delta_{ii} = 0$$

$$I = \begin{pmatrix} 1, & 0, \ldots, 0 \\ 0, & 1, \ldots, 0 \\ \ldots & \\ 0, & 0, \ldots, 1 \end{pmatrix}$$

è ***l'elemento neutro*** rispetto al prodotto tra matrici. Risulta infatti per ogni matrice A di M

$$AI = IA = A.$$

Poiché la matrice I è diagonale si ha :

$$\det I = 1.$$

Sia A una matrice quadrata d'ordine n . L' ***inversa*** della matrice A, se esiste , è una matrice A^{-1} anch'essa quadrata d'ordine n , che abbia la seguente proprietà :

$$A\,A^{-1} = A^{-1}\,A = I\,.$$

Allo scopo di determinare quali matrici sono dotate di inversa ci è utile provare preliminarmente la seguente proprietà :

Sia A una matrice quadrata d'ordine n . Risulta :

1. $a_{i1}\,A_{j1} + a_{i2}\,A_{j2} + \ldots\ldots + a_{in}A_{jn} = \det A \qquad \text{se} \quad i = j$

2. $a_{i1} A_{j1} + a_{i2} A_{j2} + \ldots\ldots + a_{in} A_{jn} = 0$ se $i \neq j$

Occorre provare la 2, essendo la 1. vera per la definizione di determinante.

All'uopo consideriamo la matrice A

$$A = \begin{pmatrix} a_{11} \ a_{12} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{1n} \\ \ldots. \\ a_{i1} \ a_{i2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{in} \\ .. \\ a_{j1} \ a_{j2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{jn} \\ .. \\ a_{n1} \ a_{n2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{nn} \end{pmatrix}$$

e la matrice ausiliaria B , eguale alla matrice A tranne che nella riga di posto j dove presenta di nuovo la riga $\underline{a}_i$

$$B = \begin{pmatrix} a_{11} \ a_{12} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{1n} \\ \ldots. \\ a_{i1} \ a_{i2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{in} \\ .. \\ a_{i1} \ a_{i2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{in} \\ .. \\ a_{n1} \ a_{n2} \ \ldots\ldots\ldots\ldots\ldots\ldots a_{nn} \end{pmatrix}$$

Per tali matrici , (essendo in tutto eguali tranne che nella riga di posto j), risulta

$$A_{j1} = B_{j1} \ , \ A_{j2} = B_{j2} \ , \ \ldots \ \ldots, \ A_{jn} = B_{jn}$$

E quindi è , come si voleva :

$$0 = \det B = b_{j1} B_{j1} + b_{j2} B_{j2} + \ .. + \ b_{jn} B_{jn} = a_{i1} A_{j1} + a_{i2} A_{j2} + \ \ldots + a_{in} A_{jn}$$

Un' altra utile proprietà , di cui omettiamo la dimostrazione , è la seguente :

Per ogni coppia di matrici A e B in M risulta

3. $$\det(AB) = \det A \det B$$

Attraverso l'uso delle proprietà 1., 2. ,3 ora stabilite siamo ora in grado di provare la seguente importante:

Proposizione 3.1 *Una matrice quadrata A ammette inversa se e solo se essa ha il determinante diverso da zero.*

Dimostrazione . Se la matrice A ammette inversa esiste una matrice A^{-1} tale che risulti $A A^{-1} = I$. Si ha $\det(A A^{-1}) = \det A \det A^{-1} = \det I = 1$ da cui segue $\det A \neq 0$.

Viceversa supponiamo sia $\det A = k \neq 0$. Tenendo conto delle 1 e 2 la matrice :

$$A^{-1} = \frac{1}{k} \begin{pmatrix} A_{11} & A_{21} & \dots\dots\dots\dots\dots & A_{n1} \\ A_{12} & A_{22} & \dots\dots\dots\dots\dots & A_{n2} \\ \dots\dots\dots & & & \\ & & & \\ A_{1n} & A_{2n} & \dots\dots\dots\dots\dots & A_{nn} \end{pmatrix}$$

risulta l'inversa della matrice A .

4 . Rango di una matrice.

Sia assegnata una matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \dots\dots & a_{1n} \\ a_{21} & a_{22} \dots\dots & a_{2n} \\ \dots & & \\ a_{m1} & a_{m2} \dots\dots & a_{mn} \end{pmatrix}$$

una matrice di tipo (m ,n) sul campo K. Indicheremo le sue m righe coi simboli $\underline{a}_1 , \underline{a}_2,\dots, \underline{a}_m$ e le sue n colonne coi simboli $\underline{a}^1 , \underline{a}^2,\dots, \underline{a}^n$ Supponiamo di aver individuato nella matrice A , p sue righe , per esempio le prime p , con queste proprietà :

i) $\underline{a}_1 , \underline{a}_2,\dots, \underline{a}_p$ *sono indipendenti*

ii) *ogni altra riga risulta una loro combinazione lineare.*

Se $\underline{a}_{i_1} , \underline{a}_{i_2},\dots, \underline{a}_{i_t}$ sono t righe indipendenti di A poiché ognuna di esse dipende da

$\underline{a}_1 , \underline{a}_2,\dots, \underline{a}_p$ allora , per il teorema di Steinitz risulta $t \leq p$.

Pertanto il *numero massimo di righe indipendenti di A è p* . Individuare nella matrice A un gruppo di righe con le proprietà i) e ii) equivale quindi a determinare quale sia il *numero massimo di righe indipendenti* che A possiede.

In modo analogo supponiamo che nella matrice A ci siano **s** colonne per esempio le prime s con le seguenti proprietà :

j) $\underline{a}^1 , \underline{a}^2,\dots, \underline{a}^s$ *sono indipendenti*

jj) *ogni altra colonna risulta una loro combinazione lineare.*

Se $\underline{a}^{j_1} , \underline{a}^{j_2},\dots, \underline{a}^{j_t}$ sono t colonne indipendenti di A, poiché ognuna di esse dipende da

$\underline{a}^1 , \underline{a}^2,\dots, \underline{a}^s$ allora , per il teorema di Steinitz risulta $t \leq s$.

Pertanto il *numero massimo di colonne indipendenti di A è s* . Individuare nella matrice A un gruppo di colonne con le proprietà j) e jj) equivale quindi a determinare quale sia il *numero massimo di colonne indipendenti* che A possiede. Mostreremo ora che p = s e cioè che *il massimo numero di righe indipendenti di A eguaglia il massimo numero di colonne indipendenti di A* .

Il ***numero massimo di righe (o colonne) indipendenti*** di A è detto il ***rango*** di A **.**

Proviamo quindi la seguente

Proposizione 4.1 *Il massimo numero di righe indipendenti di una matrice A eguaglia il numero massimo di colonne indipendenti di A.*

Dimostrazione. Possiamo supporre che la matrice A non abbia tutti gli elementi eguali a zero altrimenti l'asserto è ovvio.

Siano $\underline{a}_{i_1}, \underline{a}_{i_2}, \ldots, \underline{a}_{i_p}$ p righe di A con le proprietà i) e ii) e siano

$\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_t}$ t colonne di A con le proprietà j) e jj). Mostreremo che è p=t e ciò proverà l'asserto.

Poiché le righe $\underline{a}_{i_1}, \underline{a}_{i_2}, \ldots, \underline{a}_{i_p}$ generano tutte le altre si ha

$$(\underline{a}_1) = \alpha_1(\underline{a}_{i_1}) + \alpha_2(\underline{a}_{i_2}) + \ldots + \alpha_p(\underline{a}_{i_p})$$

$$(\underline{a}_2) = \beta_1(\underline{a}_{i_1}) + \beta_2(\underline{a}_{i_2}) + \ldots + \beta_p(\underline{a}_{i_p})$$

…..

……

$$(\underline{a}_m) = \gamma_1(\underline{a}_{i_1}) + \gamma_2(\underline{a}_{i_2}) + \ldots + \gamma_p(\underline{a}_{i_p})$$

Ora se volessimo descrivere la colonna di posto j di A dovremmo prendere l'elemento di posto j nella prima riga , prendere l'elemento di posto j della seconda riga , prendere l'elemento di posto j dell'ultima riga .

$$(\ldots \underline{a}_1 \ldots \overset{j}{\square} \ldots) = \alpha_1(\ldots \underline{a}_{i_1} \ldots \overset{j}{\square} \ldots) + \alpha_2(\ldots \underline{a}_{i_2} \ldots \overset{j}{\square} \ldots) + \ldots + \alpha_p(\ldots \underline{a}_{i_p} \ldots \overset{j}{\square} \ldots)$$

$$(\ldots \underline{a}_2 \ldots \square \ldots) = \beta_1(\ldots \underline{a}_{i_1} \ldots \square \ldots) + \beta_2(\ldots \underline{a}_{i_2} \ldots \square \ldots) + \ldots + \beta_p(\ldots \underline{a}_{i_p} \ldots \square \ldots)$$

…..

…..

$$(\ldots \underline{a}_m \ldots \square \ldots) = \gamma_1(\ldots \underline{a}_{i_1} \ldots \square \ldots) + \gamma_2(\ldots \underline{a}_{i_2} \ldots \square \ldots) + \ldots + \gamma_p(\ldots \underline{a}_{i_p} \ldots \square \ldots)$$

e quindi si ha

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ .. \\ .. \\ a_{mj} \end{pmatrix} = a_{i_1 j} \begin{pmatrix} \alpha_1 \\ \beta_1 \\ .. \\ .. \\ \gamma_1 \end{pmatrix} + a_{i_2 j} \begin{pmatrix} \alpha_2 \\ \beta_2 \\ .. \\ .. \\ \gamma_2 \end{pmatrix} + \ldots.. + a_{i_p j} \begin{pmatrix} \alpha_p \\ \beta_p \\ .. \\ .. \\ \gamma_p \end{pmatrix}$$

Abbiamo così provato che ogni colonna di A è combinazione lineare dei p vettori

$$\begin{pmatrix} \alpha_1 \\ \beta_1 \\ .. \\ .. \\ \gamma_1 \end{pmatrix}, \begin{pmatrix} \alpha_2 \\ \beta_2 \\ .. \\ .. \\ \gamma_2 \end{pmatrix}, \ldots.., \begin{pmatrix} \alpha_p \\ \beta_p \\ .. \\ .. \\ \gamma_p \end{pmatrix}$$

e conseguentemente essendo le colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_t}$ indipendenti si ha per il teorema di Steinitz $t \leq p$.

Nella matrice trasposta A_T di A le righe di A sono le colonne di A_T e le colonne di A sono le righe di A_T

Applicando alla matrice trasposta quanto già provato per A si ha allora $p \leq t$ e l'asserto è così provato.

Illustreremo ora un teorema molto utile in quanto esso fornisce un metodo per il calcolo del rango di una matrice.Prima però occorre dare alcune semplici definizioni.

Sia A una matrice di tipo m ,n sul campo K.

Si scelgano h righe $\underline{a}_{i_1}, \underline{a}_{i_2}, \ldots, \underline{a}_{i_h}$ di A ed h colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_h}$ di A. La sottomatrice H di A i cui elementi sono quelli che si trovano contemporaneamente sulle righe e colonne scelte sarà da noi chiamata un *minore d'ordine h* di A.

Quindi la prima riga di H ha per elementi gli elementi della riga $\underline{a}_{i_1}$ che occupano i posti $j_1, j_2, \ldots, j_h$. La seconda riga di H ha per elementi gli elementi della riga $\underline{a}_{i_2}$ che occupano i posti $j_1, j_2, \ldots, j_h$. L'ultima riga di H ha per elementi gli elementi della riga $\underline{a}_{i_h}$ che occupano i posti $j_1, j_2, \ldots, j_h$ .Facciamo un esempio.

Sia A la seguente matrice

$$A=\begin{pmatrix} 1, & 2, & 5, & 1, & 0 \\ 3, & 3, & 2, & 0, & 4 \\ 1, & 0, & 3, & 4, & 2 \end{pmatrix}$$

Se si scelgono la prima e la terza riga e la quarta e quinta colonna di A la matrice H che tali scelte determinano è

$$H=\begin{pmatrix} 1, & 0 \\ 4, & 2 \end{pmatrix}$$

Se si scelgono la prima e la seconda riga e la prima e la seconda colonna di A la matrice L che tali scelte determinano è

$$L=\begin{pmatrix} 1, & 2 \\ 3, & 3 \end{pmatrix}$$

Sia H il minore d'ordine h ottenuto scegliendo le righe $\underline{a}_{i_1}$, $\underline{a}_{i_2}$,…., $\underline{a}_{i_h}$ di A e le colonne $\underline{a}^{j_1}$, $\underline{a}^{j_2}$,…., $\underline{a}^{j_h}$ di A. Fissate una ulteriore riga $\underline{a}_i$ ed una ulteriore colonna $\underline{a}^j$ con $i \neq i_1, i_2, \ldots, i_h$ e $j \neq j_1, j_2, \ldots, j_h$ il minore $H_{i,j}$ d'ordine h+1 ottenuto scegliendo le righe $\underline{a}_{i_1}$, $\underline{a}_{i_2}$,…., $\underline{a}_{i_h}$, $\underline{a}_i$ di A e le colonne $\underline{a}^{j_1}$, $\underline{a}^{j_2}$,…., $\underline{a}^{j_h}$ $\underline{a}^j$ di A si chiama un *orlato* del minore H.

Riferendoci agli esempi precedenti, si ha ad esempio

$$L_{3,4} = \begin{pmatrix} 1, & 2, & 1 \\ 3, & 3, & 0 \\ 1, & 0, & 4 \end{pmatrix} \qquad H_{2,1} = \begin{pmatrix} 1, & 0, & 1 \\ 4, & 2, & 3 \\ 0, & 4, & 3 \end{pmatrix}$$

Siamo ora in grado di enunciare e provare il seguente importante

Teorema degli orlati. *Sia A una matrice di tipo m,n sul campo K. Sia H un minore d'ordine h di A ottenuto scegliendo le righe* $\underline{a}_{i_1}, \underline{a}_{i_2}, \dots, \underline{a}_{i_h}$ *di A e le colonne* $\underline{a}^{j_1}, \underline{a}^{j_2}, \dots, \underline{a}^{j_h}$ *di A. Se risulta :*

1. *detH=* $\delta \neq 0$
2. *det*$H_{i,j} = 0$ *per ogni i e j* $(i \neq i_1, i_2, \dots, i_h \quad e \quad j \neq j_1, j_2, \dots, j_h)$

allora :

i) *le righe* $\underline{a}_{i_1}, \underline{a}_{i_2}, \dots, \underline{a}_{i_h}$ *sono indipendenti ed ogni altra riga risulta una loro combinazione lineare*

ii) *le colonne* $\underline{a}^{j_1}, \underline{a}^{j_2}, \dots, \underline{a}^{j_h}$ *sono indipendenti ed ogni altra colonna risulta una loro combinazione lineare.*

Ne segue che il ***rango di A è h****.*

Dimostrazione . Poiché H ha il determinante diverso da zero allora le sue righe e le sue colonne sono indipendenti. Le righe $\underline{a}_{i_1}, \underline{a}_{i_2}, \dots, \underline{a}_{i_h}$ di A sono un allungamento delle righe di H e quindi sono anch'esse indipendenti. Analogamente le colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \dots, \underline{a}^{j_h}$ di A sono un allungamento delle colonne di H e quindi sono anch'esse indipendenti. Mostriamo che ogni riga $\underline{a}_i$ di A è combinazione lineare delle righe $\underline{a}_{i_1}, \underline{a}_{i_2}, \dots, \underline{a}_{i_h}$.

Per le ipotesi fatte le matrici $H_{i,1}$, $H_{i,2}, \dots, H_{i,n}$ ottenute orlando H con la riga di posto i e tutte le colonne da 1 ad n hanno tutte il determinante eguale a zero.

$$\det H_{i,1} = \det\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_h} & a_{i_1 1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_h j_1} & \cdots & a_{i_h j_h} & a_{i_h 1} \\ a_{i j_1} & \cdots & a_{i j_h} & a_{i1} \end{pmatrix} = 0$$

$$\det H_{i,2} = \det\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_h} & a_{i_1 2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_h j_1} & \cdots & a_{i_h j_h} & a_{i_h 2} \\ a_{i j_1} & \cdots & a_{i j_h} & a_{i2} \end{pmatrix} = 0$$

$\vdots$

$$\det H_{i,n} = \det\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_h} & a_{i_1 n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_h j_1} & \cdots & a_{i_h j_h} & a_{i_h n} \\ a_{i j_1} & \cdots & a_{i j_h} & a_{in} \end{pmatrix} = 0$$

ora le matrici $H_{i,1}$, $H_{i,2}, \ldots, H_{i,n}$ differiscono solo nell'ultima colonna per cui i complementi algebrici degli elementi dell'ultima colonna di $H_{i,1}$ sono gli stessi di quelli degli elementi dell'ultima colonna di $H_{i,2}$ e di quelli degli elementi dell'ultima colonna di $H_{i,n}$. Indichiamo tali complementi algebrici con $\gamma_1 , \gamma_2 , \ldots, \gamma_h , \delta$. Sviluppando i determinanti di $H_{i,1}$, $H_{i,2}, \ldots, H_{i,n}$ secondo gli elementi dell'ultima colonna si hanno le seguenti relazioni scalari :

$$a_{i_1 1}\gamma_1 + \ldots + a_{i_h 1}\gamma_h + a_{i1}\delta = 0$$

$$a_{i_1 2}\gamma_1 + \ldots + a_{i_h 2}\gamma_h + a_{i2}\delta = 0$$

....

...

$$a_{i_1 n}\gamma_1 + \ldots + a_{i_h n}\gamma_h + a_{in}\delta = 0$$

Tali relazioni sono equivalenti alla seguente relazione vettoriale

$$\delta\, \underline{a}_i = -(\gamma_1 \underline{a}_{i_1} + \ldots + \gamma_h \underline{a}_{i_h})$$

e dividendo per δ , che è diverso da zero , si ha

$$\underline{a}_i = -\frac{1}{\delta}(\gamma_1 \underline{a}_{i_1} + \ldots + \gamma_h \underline{a}_{i_h})$$

la quale mostra che la riga $\underline{a}_i$ è combinazione lineare delle righe $\underline{a}_{i_1}, \ldots, \underline{a}_{i_h}$.

Avendo provato che le righe $\underline{a}_{i_1}, \ldots, \underline{a}_{i_h}$ sono indipendenti e che generano tutte le altre allora è h il massimo numero di righe indipendenti. Poiché il massimo numero di righe indipendenti eguaglia il massimo numero di colonne indipendenti allora anche le colonne indipendenti $\underline{a}^{j_1}$, $\underline{a}^{j_2}$,…., $\underline{a}^{j_h}$ sono in numero massimo e quindi ogni altra colonna risulta una loro combinazione lineare. Il teorema è così completamente provato.

Il teorema ora provato oltre a fornire un metodo per il calcolo del rango ha questa importante conseguenza espressa dalla seguente proposizione.

Proposizione 4.2 *Sia A una matrice quadrata d'ordine n . Le righe e le colonne di A sono dipendenti se e solo se il suo determinante è eguale a zero.*

La proposizione 4.2 è ovviamente equivalente alla seguente

Proposizione 4.3 *Sia A una matrice quadrata d'ordine n . Le righe e le colonne di A sono indipendenti se e solo se il suo determinante è diverso da zero.*

C A P I T O L O III

Sistemi di equazioni lineari

1.Sistemi di equazioni lineari

In questo capitolo studieremo i sistemi di equazioni lineari , cioè sistemi S di questo tipo

$$S\begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

dove sia i ***coefficienti*** a_{ij} ed i ***termini noti*** c_i sono elementi di un fissato campo K.

Quando i termini noti $c_1, c_2, \ldots, c_n$ sono tutti eguali a ***zero*** il sistema è detto ***omogeneo.***

Una ***soluzione*** del sistema è una n-pla ordinata (y_1 , $y_2, \ldots, y_n$) di elementi di K soluzione delle equazioni del sistema cioè verificante

$$\begin{cases} a_{11}y_1 + a_{12}y_2 .. + a_{1n}y_n = c_1 \\ a_{21}y_1 + a_{22}y_2 .. + a_{2n}y_n = c_2 \\ .. \\ a_{m1}y_1 + a_{m2}y_2 .. + a_{mn}y_n = c_m \end{cases}$$

Quando il sistema ha almeno una soluzione esso è detto ***compatibile*** **.**Quando esso non ha soluzioni è detto ***incompatibile.*** Al fine di determinare condizioni che assicurino la compatibilità del sistema S sono utili le seguenti considerazioni.

Assegnato un sistema

$$S\begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

possiamo considerare le seguenti due matrici : la prima , denotata con **A** , detta ***matrice incompleta*** , o dei ***coefficienti*** è quella che ha per elementi i coefficienti delle incognite

$$\mathbf{A} = \begin{pmatrix} a_{11}\ ,\ a_{12},\ldots a_{1n} \\ a_{21},\ \ a_{22}\ ..\ a_{2n} \\ .. \\ a_{m1},\ a_{m2},.\ \ a_{mn} \end{pmatrix}$$

La seconda , detta ***matrice completa*** , e denotata con **A'**

$$A' = \begin{pmatrix} a_{11}\ ,\ a_{12},\ldots a_{1n}\ \ c_1 \\ a_{21},\ \ a_{22}\ ..\ a_{2n}\ c_2 \\ .. \\ a_{m1},\ a_{m2},.\ \ a_{mn} c_m \end{pmatrix}$$

contiene oltre i coefficienti anche i termini noti disposti come sua ultima colonna**.**

Indichiamo con $\underline{A}^1$, $\underline{A}^2$, …., $\underline{A}^n$ le n colonne di A e con $\underline{C}$ la colonna dei termini noti.

$$\underline{A}^1 = \begin{pmatrix} a_{11} \\ a_{21} \\ .. \\ a_{m1} \end{pmatrix} \quad \underline{A}^2 = \begin{pmatrix} a_{12} \\ a_{22} \\ .. \\ a_{m2} \end{pmatrix} \quad \ldots\ldots \quad \underline{A}^n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ .. \\ a_{mn} \end{pmatrix} \quad \underline{C} = \begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_m \end{pmatrix}$$

Con tali notazioni il sistema S può essere scritto in modo equivalente nella seguente ***forma vettoriale*** :

$$S: \qquad \underline{A}^1 x_1 + \underline{A}^2\ x_2 + \ \ldots + \underline{A}^n\ x_n\ = \underline{C}$$

Se (y_1 , y_2,…., y_n) è una soluzione del sistema allora si ha

(1.1) $$\underline{A}^1 y_1 + \underline{A}^2\ y_2 + \ \ldots + \underline{A}^n\ y_n\ = \underline{C}$$

e questa mostra che il vettore $\underline{C}$ dei termini noti è combinazione lineare dei vettori $\underline{A}^1$, $\underline{A}^2$, …., $\underline{A}^n$

colonne di A **.** Viceversa se il vettore $\underline{C}$ dei termini noti è combinazione lineare dei vettori $\underline{A}^1$, $\underline{A}^2$, …., $\underline{A}^n$, colonne di A cioè se si ha

(1.2) $$\underline{A}^1 y_1 + \underline{A}^2 y_2 + \dots + \underline{A}^n y_n = \underline{C}$$

allora (y_1 , y_2,…., y_n) è una soluzione del sistema.
Ci sono quindi *tante soluzioni quanti sono i modi* in cui $\underline{\mathbf{C}}$ si pùò ottenere come combinazione lineare dei vettori $\underline{A}^1$, $\underline{A}^2$, …., $\underline{A}^n$.

Abbiamo così stabilito questo importante ***criterio di compatibilità*** del sistema S.

Proposizione 1.1 *Il sistema*

$$S\begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

ha soluzioni se e soltanto se il vettore $\underline{C}$ *dei termini noti si pùò ottenere come combinazione lineare dei vettori colonna* $\underline{A}^1$ *,* $\underline{A}^2$ *, …,* $\underline{A}^n$ *della matrice A .*

Usando la proposizione 1.1 possiamo ora provare un altro teorema (Rouchè-Capelli) il quale fornisce un altro criterio di compatibilità del sistema . Tale criterio è però più facilmente utilizzabile nelle applicazioni.

Proposizione 1.2 *Il sistema*

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

ha soluzioni se e soltanto se le due matrici A ed A' del sistema hanno lo stesso rango.

Dimostrazione. Indichiamo con p il rango di A e con p' il rango di A'.Dobbiamo far vedere che il sistema S ha soluzioni se e soltanto se risulta p = p' .

Supponiamo dapprima che sia p = p' e proviamo che il sistema S ha soluzioni. Siano $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$, p colonne indipendenti della matrice A. Tali colonne fanno parte anche della matrice A' e poiché per ipotesi anche A' ha rango p , tali colonne indipendenti sono in numero massimo. Ne, segue che ogni altra colonna di A' è combinazione lineare delle colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$. In particolare la colonna $\underline{c}$ dei termini noti dipende dalle colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$ e quindi da tutte le colonne di A. Per la proposizione 1.1 se la colonna $\underline{c}$ dei termini dipende da tutte le colonne di A , S ha soluzioni.

Supponiamo ora che S abbia soluzioni e mostriamo che è p = p' . Siano $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$ p colonne indipendenti della matrice A ed in numero massimo cioè tali che ogni altra colonna di A risulti una loro combinazione lineare. Se mostriamo che ogni colonna di A' risulta combinazione lineare delle colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$ allora tali colonne sono un sistema massimo di colonne indipendenti di A' che ha così anche essa rango p . Una qualunque colonna $\underline{a}^{j}$ di A' diversa dalla colonna $\underline{c}$ essendo anche colonna di A dipende da $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$. Poiché per ipotesi il sistema ha soluzioni , per la proposizione 1.1. la colonna $\underline{c}$ dipende dalle colonne di A . Ma poiché le colonne di A dipendono da $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$ allora anche $\underline{c}$ dipende dalle colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$. Avendo provato che ogni colonna di A' dipende dalle colonne $\underline{a}^{j_1}, \underline{a}^{j_2}, \ldots, \underline{a}^{j_p}$ resta provato che tali colonne sono colonne indipendenti di A' ed in numero massimo per cui p è il rango di A' . Risulta quindi p' = p ed il teorema è provato.

Facciamo un esempio per mostrare l'utilità del teorema ora provato.

Si voglia ad esempio stabilire per quale valore del parametro reale h il seguente sistema S ha soluzioni .

$$S = \begin{cases} 2x + y + hz = 4 \\ x - 2y + 2z = 1 \\ 3x + y + z = 5 \end{cases}$$

Le due matrici A ed A' del sistema S sono

$$A = \begin{pmatrix} 2 & 1 & h \\ 1 & -2 & 2 \\ 3 & 1 & 1 \end{pmatrix} \qquad A' = \begin{pmatrix} 2 & 1 & h & 4 \\ 1 & -2 & 2 & 1 \\ 3 & 1 & 1 & 5 \end{pmatrix}$$

La matrice

$$L = \begin{pmatrix} 2 & 1 & 4 \\ 1 & -2 & 1 \\ 3 & 1 & 5 \end{pmatrix}$$

estratta da A' ha determinante diverso da zero risultando $\det L = 4$ e pertanto A' ha rango tre . La matrice A ha rango tre se e solo se risulta $\det A = 7h - 3 \neq 0$. Pertanto per $h \neq \frac{3}{7}$ il sistema ha soluzioni , risultando $p = p' = 3$, mentre per $h = \frac{3}{7}$ il sistema non ha soluzioni risultando $p = 2$ e $p' = 3$.

Le proposizioni 1.1 e 1.2 ora provate forniscono criteri per stabilire la compatibilità del sistema ma non offrono metodi di determinazione delle soluzioni quando si sia accertato che queste esistono.

Vedremo ora come si trovano le soluzioni di un sistema S quando si sia stabilito che le soluzioni esistono.

Cominciamo col caso più favorevole .

Supponiamo che il sistema

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{n1}x_1 + a_{n2}x_2 .. + a_{nn}x_n = c_n \end{cases}$$

abbia **n equazioni ed n incognite** . In tal caso la matrice A incompleta del sistema avendo n righe ed n colonne è quadrata e quindi è possibile calcolare il suo determinante . Se risulta

$$\det A \neq 0$$

allora il sistema ha soluzioni in quanto risulta **p = p' = n** . Proveremo ora che nelle ipotesi in cui siamo il sistema ***ha una sola soluzione*** . Per provare ciò è utile scrivere il sistema in una forma equivalente detta *forma matriciale*.

Denotata al solito con

$$A = \begin{pmatrix} a_{11}\,,\ a_{12},\ldots a_{1n} \\ a_{21},\ \ a_{22}\ ..\ a_{2n} \\ .. \\ a_{n1},\ a_{n2},.\ \ a_{nn} \end{pmatrix}$$

la matrice incompleta del sistema e con $\underline{X}$ e $\underline{C}$ i seguenti vettori (*delle incognite e dei termini noti*)

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{pmatrix} \qquad\qquad \underline{C} = \begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_n \end{pmatrix}$$

si riconosce facilmente che il sistema **S** può essere scritto in modo equivalente nella seguente forma matriciale :

$$A\ \underline{X} = \underline{C}$$

$$\mathbf{S:}\quad \begin{pmatrix} a_{11},\ a_{12},\ldots a_{1n} \\ a_{21},\ \ a_{22}\,..\ a_{2n} \\ .. \\ a_{n1},\ a_{n2},.\ \ a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_n \end{pmatrix}$$

Poiché per ipotesi è $\det A = \delta \neq 0$ la matrice A ammette inversa . Denotata con A^{-1} la matrice inversa di A , moltiplicando ambo i membri di $A\ \underline{X} = \underline{C}$, per A^{-1} si ha che la soluzione del sistema è unica ed è eguale a

$$\underline{X} = A^{-1}\ \underline{C} \tag{1.2}$$

Possiamo ora rendere espliciti i valori $(x_1, x_2, \ldots, x_n)$ della soluzione trovata.
Ricordiamo che la matrice inversa di A ha nella riga di posto **i** , i complementi algebrici degli elementi della colonna di posto **i** di A , ciascuno diviso per detA. Dalla (1.2) qui scritta esplicitamente :

$$\begin{pmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} A_{11},\ A_{21},\ldots A_{n1} \\ A_{12},\ \ A_{22}\,..\ A_{n2} \\ .. \\ A_{1n},\ \ A_{2n},\ldots A_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_n \end{pmatrix}$$

segue quindi , per ogni $i = 1,2,\ldots,n$,

$$x_i = \frac{A_{1i}\ c_1 + A_{2i}\ c_2 + \ldots + A_{ni}\ c_n}{\delta} \tag{1.3}$$

Indichiamo ora con B^i la matrice che si ottiene dalla matrice A sostituendo la sua i-sima colonna con la colonna $\begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_n \end{pmatrix}$ dei termini noti .

$$
\begin{array}{cc}
 & \cdots\cdots\; i \;\cdots\cdots \\
B^i = & \begin{pmatrix} a_{11}, & a_{12},\ldots & c_1 \ldots & a_{1n} \\ a_{21}, & a_{22} \ldots & c_2 \ldots & a_{2n} \\ \ldots & & & \\ a_{n1}, & a_{n2},\ldots & c_n \ldots & a_{nn} \end{pmatrix}
\end{array}
$$

Poiché le matrici A e B^i differiscono solo nella colonna di posto i , allora i complementi algebrici degli elementi della colonna i-sima di B^i sono eguali ai complementi algebrici degli elementi della colonna i-sima di A. Ne segue che al numeratore della (1.3) ciascun elemento c_i della colonna i-sima di B^i è moltiplicato per il suo complemento algebrico. Il numeratore della (1.3) è quindi il determinante della matrice B^i .

Abbiamo così provato il seguente importante risultato

Proposizione 1.3 *Sia assegnato un sistema*

$$
S = \begin{cases} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n = c_2 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 \ldots + a_{nn}x_n = c_n \end{cases}
$$

*avente **n** equazioni in **n** incognite . Se risulta **det A ≠ 0** , il sistema ha una **sola** soluzione*

(y_1 , y_2,...., y_n) che si determina al seguente modo (regola di Cramer) :

$$
y_1 = \frac{\det B^1}{\det A}, \quad y_2 = \frac{\det B^2}{\det A}, \quad \ldots\ldots\ldots\ldots, \quad y_n = \frac{\det B^n}{\det A}
$$

(*avendo al solito indicato con B^i (i= 1,2,...,n) la matrice che si ottiene dalla matrice A* *sostituendo la sua i-sima colonna con la colonna* $\begin{pmatrix} c_1 \\ c_2 \\ .. \\ c_n \end{pmatrix}$ *dei termini noti) .*

Facciamo un esempio. Si vogliano determinare le soluzioni del seguente sistema

$$S = \begin{cases} x + y + z = 4 \\ x - y + z = 2 \\ x + y - 3z = 0 \end{cases}$$

Le due matrici del sistema sono

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -3 \end{pmatrix} \qquad A' = \begin{pmatrix} 1 & 1 & 1 & 4 \\ 1 & -1 & 1 & 2 \\ 1 & 1 & -3 & 0 \end{pmatrix}$$

Essendo $\det A = 8$ per la proposizione 1.3 tale sistema ha una sola soluzione che si determina con la regola di Cramer .Risulta allora

$$x = \frac{1}{8} \det \begin{pmatrix} 4 & 1 & 1 \\ 2 & -1 & 1 \\ 0 & 1 & -3 \end{pmatrix} = \frac{16}{8} = 2$$

$$y = \frac{1}{8} \det \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & -3 \end{pmatrix} = \frac{8}{8} = 1$$

$$z = \frac{1}{8} \det \begin{pmatrix} 1 & 1 & 4 \\ 1 & -1 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \frac{8}{8} = 1$$

Rimangono ora da esaminare i sistemi di equazioni ***compatibili*** in cui il numero **m** di equazioni è diverso dal numero **n** delle incognite e quelli in cui pur essendo **m=n** risulta **detA = 0**.

Prima di passare all'esame di questi casi sono molto utili le seguenti considerazioni.

Supponiamo di avere un sistema di equazioni lineari

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n + c_1 = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n + c_2 = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n + c_m = 0 \\ \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n + \alpha = 0 \end{cases}$$

in cui una delle equazioni che qui abbiamo rappresentato , *per distinguerla* , con

$$\alpha_1 x_1 + \alpha_2 x_2 .. + \alpha_n x_n + \alpha = 0$$

sia combinazione lineare di tutte le altre . Si abbia cioè

(1.4) $$\alpha_1 x_1 + \alpha_2 x_2 .. + \alpha_n x_n + \alpha = \lambda_1 (a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n + c_1) + \lambda_2 (a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n + c_2) + ... + \lambda_m (a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n + c_m)$$

Consideriamo il sistema

$$S' = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n + c_1 = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n + c_2 = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n + c_m = 0 \end{cases}$$

ottenuto privando il sistema S dell' equazione $\alpha_1 x_1 + \alpha_2 x_2 .. + \alpha_n x_n + \alpha = 0$.

Ogni soluzione $(y_1 , y_2, …, y_n)$ del sistema S è evidentemente soluzione anche del sistema S'.

Viceversa se (y_1 , y_2,…., y_n) è una soluzione del sistema S' , stante la (1.4) , essa è anche soluzione del sistema S . In conclusione eliminando da S una equazione che sia combinazione lineare delle altre si ottiene un sistema che ha *una equazione in meno ma che ha le stesse soluzioni del sistema S* .

Al fine di determinare le soluzioni di un assegnato sistema S compatibile ci si può quindi limitare a considerare **solo** quelle equazioni di S che siano ***indipendenti e tali che ogni altra equazione sia una loro combinazione lineare.***

Ma come si fa a selezionare queste "***fortunate equazioni*** " ? Vediamo.

Ricordiamo che lo spazio vettoriale $K[x_1 , x_2,\dots, x_n]$ dei polinomi di primo grado in n variabili e lo spazio vettoriale numerico K^{n+1} sono isomorfi , un isomorfismo essendo l'applicazione

$$f: \ (a_1x_1+ a_2 x_2 +\dots+ a_n x_n + c) \longrightarrow (a_1, a_2, \dots, a_n, c)$$

che associa al polinomio $a_1x_1 + a_2x_2..+ a_nx_n + c$ la n+1 –pla $(a_1, a_2, .. a_n, c)$ dei suoi coefficienti.

Sia ora assegnato un sistema S ***compatibile*** di m equazioni in n incognite.

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. \ + a_{1n}x_n \ + c_1 = 0 \\ a_{21}x_1 + a_{22}x_2 .. \ + a_{2n}x_n \ + c_2 = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n \ + c_m = 0 \end{cases}$$

e sia

$$A' = \begin{pmatrix} a_{11} , \ a_{12} ,\dots a_{1n} \ \ c_1 \\ a_{21} , \ \ a_{22} .. \ a_{2n} \ c_2 \\ .. \\ a_{m1} , \ a_{m2} ,. \ \ a_{mn} c_m \end{pmatrix}$$

la sua matrice completa. Ora le righe della matrice A' sono le immagini tramite l'isomorfismo f delle equazioni del sistema per cui un numero massimo di equazioni indipendenti del sistema si otterrà in corrispondenza ad un numero massimo di righe indipendenti di A' .

Ora poiché il sistema S è per ipotesi compatibile , le due matrici A ed A' del sistema hanno lo stesso rango e sia p il rango di A ed A'. Se H è la sottomatrice quadrata d'ordine p di A con determinante diverso da zero (ma che ha tutti i suoi orlati con determinante nullo) che ci ha permesso di stabilire il rango di A allora le p righe di A che concorrono alla formazione di H sono un sistema massimo di righe indipendenti di A . Le stesse righe nella matrice A' costituiscono anche in A' che anch'essa rango p , un sistema massimo di righe indipendenti di A'

Le equazioni corrispondenti a tali righe sono quindi indipendenti ed inoltre ogni altra equazione risulta una loro combinazione lineare. Attraverso il calcolo del rango di A (uguale a quello di A') utilizzando il teorema degli orlati , abbiamo così stabilito quali siano le equazioni del sistema S che debbono essere mantenute e quali invece possono essere eliminate, senza alterare l'insieme delle soluzioni del sistema. Quando si sia effettuata questa scelta il sistema ha sempre la seguente forma :

ci sono nel sistema S , ***p equazioni ed n incognite con p ≤ n*** . Vediamo perché .

Inizialmente per il sistema, ci sono tre possibili casi :

Caso 1 . **m = n** ma è **det A = 0**
In tal caso il rango p di A è minore di n e quindi mantenendo le p equazioni indipendenti si hanno

p equazioni ed n incognite con p < n

Caso 2. **m > n** Ci siano cioè più equazioni che incognite . In tal caso la matrice **A** ha una forma di rettangolo " alto ma stretto" con m righe ed n colonne . Poiché il rango esprime il massimo numero di colonne indipendenti risulta p ≤ n . Se p = n si ricade nel caso esaminato nella proposizione 1.3 e pertanto possiamo supporre che sia p < n. Anche in questo caso mantenendo le p equazioni indipendenti si hanno

p equazioni ed n incognite con p < n.

Caso 3. **m < n .** Ci siano cioè meno equazioni rispetto al numero di incognite **.** . In tal caso la matrice **A** ha una forma di rettangolo " basso ma largo " con m righe ed n colonne. Poiché il rango p di A esprime il massimo numero di righe indipendenti risulta p ≤ m < n. Anche in questo caso mantenendo le p equazioni indipendenti si hanno

p equazioni ed n incognite con p < n.

Dopo quanto detto dobbiamo quindi prendere in considerazione ***un sistema S compatibile con p equazioni indipendenti ed n incognite con p < n***

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n + c_1 = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n + c_2 = 0 \\ .. \\ a_{p11}x_1 + a_{p2}x_2 .. + a_{pn}x_n + c_p = 0 \end{cases}$$

e vedere come si trovano le sue soluzioni.

Essendo le p equazioni del sistema indipendenti la matrice

$$A = \begin{pmatrix} a_{11} \;\; a_{12} \ldots\ldots\ldots\ldots\ldots\ldots\ldots a_{1n} \\ a_{21} \;\; a_{22} \ldots\ldots\ldots\ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots \\ a_{p1} \;\; a_{p2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots a_{pn} \end{pmatrix}$$

incompleta del sistema ha rango p e quindi essa possiede una sottomatrice quadrata H d'ordine p con il determinante diverso da zero corrispondente a p colonne indipendenti di A.

Per rendere semplice la scrittura, quindi solo per semplicità , supponiamo che siano le prime p colonne di A a determinare con le sue p righe tale sottomatrice H.

$$A = \begin{pmatrix} a_{11} \;\; a_{12} \ldots\ldots a_{1p} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} \;\; a_{22} \ldots\ldots a_{2p} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots \\ a_{p1} \;\; a_{p2} \ldots \;\; a_{pp} \ldots\ldots\ldots\ldots a_{pn} \end{pmatrix}$$

Se ora assegniamo alle incognite (x_{p+1} , x_{p+2} , ……, x_n) i valori (y_{p+1} , y_{p+2} , ……, y_n)

Il sistema S assume la seguente forma :

$$S' = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1p}\,x_p = - a_{1p+1}y_{p+1} - .... - a_{1n}y_n - c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2p}x_p = - a_{2p+1}y_{p+1} - \quad - a_{2n}y_n - c_2 \\ .. \\ a_{p11}x_1 + a_{p2}x_2 .. + a_{pp}x_p = - a_{pp+1}y_{p+1} - ... - a_{pn}y_n - c_p \end{cases}$$

cioè diventa un sistema con p equazioni e p incognite e con la matrice H dei suoi coefficienti avente il determinante diverso da zero. Tale sistema per la proposizione 1.3 ha una sola soluzione (y_1 , y_2 , ……, y_p) e così la n-pla

$$\zeta = (y_1 ,\ y_2 , \ldots\ldots, y_p\ y_{p+1} ,\ y_{p+2} , \ldots\ldots, y_n)$$

è una soluzione del sistema S . Poiché i primi p valori della soluzione ζ sono **determinati** dai valori y_{p+1} , y_{p+2} , ……, y_n che ζ ha nei suoi ultimi n-p posti si hanno queste due utili conseguenze .

i) *due soluzioni* ζ *e* η *che abbiano eguali i valori di posto p+1 , p+2 ,..., n coincidono.*

ii) *indicando con* $\mathcal{S}$ *l'insieme delle soluzioni del sistema S , la funzione*

$$f : (y_{p+1} ,\ y_{p+2} , \ldots\ldots, y_n) \in K^{n-p} \longrightarrow \zeta = (y_1 ,\ y_2 , \ldots\ldots, y_p\ y_{p+1} ,\ y_{p+2} , \ldots\ldots, y_n) \in \mathcal{S}$$

è biettiva .

Essendo f biettiva si ha $| K^{n-p} | = | \mathcal{S} |$ e cioè le soluzioni del sistema S sono tante quante le (n-p)-ple ordinate di elementi di K.

Se K è infinito come spesso accade (in quanto esso spesso sarà il campo reale) si dice perciò che il sistema **S** ha $\infty^{\mathbf{n-p}}$ **soluzioni .**

Poiché ogni soluzione è determinata sulla base della scelta dei valori y_{p+1}, y_{p+2}, ……, y_n assegnati alle incognite x_{p+1}, x_{p+2}, ……, x_n, è evidente che le soluzioni del sistema (spesso ***infinite***) non avendo un nesso tra loro debbono essere calcolate una per una applicando di volta in volta la regola di Cramer. Questa difficoltà come ora vedremo può essere superata dopo aver discusso della risoluzione dei sistemi omogenei.

2. Sistemi omogenei.

In questo numero studieremo i sistemi di equazioni lineari omogenei

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = 0 \end{cases}$$

cioè quelli che hanno i termini noti tutti eguali a zero.

Sono ovvie le seguenti osservazioni.

a) Il sistema S omogeneo è compatibile una sua soluzione (detta ***soluzione banale***) essendo la n-pla (0,0,….,,0) .Ne segue che le due matrici A ed A' hanno sempre lo stesso rango che indichiamo con **p** .

b) se il rango **p** è eguale ad **n** il sistema ammette, in accordo con la proposizione 1.3, soltanto la soluzione banale (0,0,….,,0).

Possiamo quindi supporre che sia **p < n** in modo che il sistema ammetta $|K|^{n-p}$ soluzioni**.**

Denotiamo al solito con $\mathcal{S}$ l'insieme delle soluzioni del sistema **S** .

Ciò che distingue i sistemi omogenei rispetto a quelli non omogenei è la "*qualità*" dell'insieme $\mathcal{S}$ delle soluzioni.

Infatti se il sistema è non omogeneo l'insieme $\mathcal{S}$ è un sottoinsieme di K^n .

Se invece il sistema è omogeneo si ha una importante proprietà espressa dalla seguente

Proposizione 2.1 Sia

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. \ + a_{1n}x_n \ = 0 \\ a_{21}x_1 + a_{22}x_2 .. \ + a_{2n}x_n \ = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n \ = 0 \end{cases}$$

un sistema omogeneo con rango p. L'insieme $\mathcal{S}'$ *delle sue soluzioni è un sottospazio di* K^n *di dimensione n-p.*

Dimostrazione. Per rendere più semplice la dimostrazione usiamo la forma matriciale

$$A\,\underline{X} = \underline{0}$$

(dove al solito è $A = \begin{pmatrix} a_{11} & a_{12} \ .... a_{1n} \\ a_{21} & a_{22} \ .... a_{2n} \\ ... & \\ a_{m1} & a_{m2} \ .... a_{mn} \end{pmatrix}$ $\quad \underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{pmatrix}$ $\quad \underline{0} = \begin{pmatrix} 0 \\ 0 \\ .. \\ 0 \end{pmatrix}$)

per rappresentare il sistema dato.

Se $\underline{Y} = \begin{pmatrix} y_1 \\ y_2 \\ .. \\ y_n \end{pmatrix}$ e $\underline{Z} = \begin{pmatrix} z_1 \\ z_2 \\ .. \\ z_n \end{pmatrix}$ sono due soluzioni del sistema allora è

$A\underline{Y} = 0$ e $A\underline{Z} = 0$ da cui segue $A(\underline{Y}+\underline{Z}) = A\underline{Y} + A\underline{Z} = 0$ e questa mostra che anche $\underline{Y} + \underline{Z}$ è una soluzione del sistema. Se $\underline{Y}$ è una soluzione del sistema si ha $A\underline{Y} = 0$ se α è uno scalare qualsiasi si ha $A(\alpha\ \underline{Y}) = \alpha\, A\underline{Y} = 0$ e ciò mostra che anche $\alpha\,\underline{Y}$ è una soluzione .

Avendo provato che l'insieme $\mathcal{S}$ delle soluzioni di S ha le seguenti proprietà :

1) $\underline{Y} \in \mathcal{S}$, $\underline{Z} \in \mathcal{S} \Rightarrow \underline{Y} + \underline{Z} \in \mathcal{S}$

2) $\alpha \in K$, $\underline{Y} \in \mathcal{S} \Rightarrow \alpha \underline{Y} \in \mathcal{S}$

resta provato che $\mathcal{S}$ è un sottospazio di K^n **.** Proveremo ora che esso ha dimensione **n-p .**

Poiché A ha rango **p** *per semplicità* supporremo che le prime **p** righe e che le prime **p** colonne di A siano indipendenti . L'insieme delle soluzioni del sistema S coincide allora con l'insieme delle soluzioni del sistema *ridotto* **S'** costituito dalle prime p equazioni di **S .**

$$\mathbf{S'} = \begin{cases} a_{11}x_1 + a_{12}\, x_2 + \ldots . + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}\, x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots \\ a_{p1}x_1 + a_{p2}\, x_2 + \ldots + a_{pn}x_n = 0 \end{cases}$$

Per quanto detto al numero 1 di questo capitolo **,** ogni soluzione di **S'** è determinata sulla base della scelta dei valori y_{p+1} , y_{p+2} , ……, y_n assegnati alle incognite x_{p+1} , x_{p+2} , ……, x_n .
Sia allora

$\underset{\sim}{\zeta}_{p+1} = (\alpha_1 , \alpha_2 , .., \alpha_p , 1 , 0,\ldots ,0)$

la soluzione di **S'** ottenuta assegnando alle incognite x_{p+1} , x_{p+2} , ……, x_n i valori
(1, 0, 0,…,0)

$\underset{\sim}{\zeta}_{p+2} = (\beta_1 , \beta_2 , .., \beta_p , 0 , 1,\ldots ,0)$

la soluzione di **S'** ottenuta assegnando alle incognite x_{p+1} , x_{p+2} , ……, x_n i valori
(0, 1, 0,…,0)

e ….

$\underline{\zeta}_n = (\gamma_1 , \gamma_2 , .., \gamma_p , 0 , 0,\dots ,1)$

la soluzione di **S'** ottenuta assegnando alle incognite x_{p+1} , x_{p+2} , ……, x_n i valori (0, 0, ,…,1)

Tali soluzioni sono n-ple indipendenti perché sono indipendenti gli n-p vettori numerici (1, 0, 0,…,0) , (0, 1, 0,…,0) ,,,,,,,,,, (0, 0, ,…,1)

Sia ora $\underline{Y} = (y_1 , y_2 , .., y_p \; y_{p+1} , y_{p+2} , \dots\dots, y_n)$ una qualunque soluzione di **S.** Poiché abbiamo provato che $\mathcal{S}$ è un sottospazio allora sono ancora soluzioni di **S** le seguenti n-ple

$$y_{p+1}\underline{\zeta}_{p+1} , \qquad y_{p+2}\underline{\zeta}_{p+2} , \qquad \dots\dots.., \qquad y_n \underline{\zeta}_n$$

Ma sempre perché $\mathcal{S}$ è un sottospazio , sommando tra loro queste soluzioni si ottiene ancora una soluzione . Ma la soluzione

$$\underline{\zeta} = y_{p+1}\underline{\zeta}_{p+1} + y_{p+2}\underline{\zeta}_{p+2} + , \dots.., + y_n \underline{\zeta}_n$$

che si ottiene eseguendo questa somma , presenta anch'essa agli ultimi **n-p** posti i valori y_{p+1} , y_{p+2} , ……, y_n

$$\underline{\zeta} = (\dots. \qquad \dots , y_{p+1} , y_{p+2} , \dots\dots, y_n)$$

al pari di $\underline{Y}$. Ne segue che è $\underline{\zeta} = \underline{Y}$ e ciò prova, per l'arbitrarietà di $\underline{Y}$ che le soluzioni

$\underline{\zeta}_{p+1}$, $\underline{\zeta}_{p+2}$, ….., $\underline{\zeta}_n$ sono oltre che indipendenti anche un sistema di generatori per lo spazio $\mathcal{S}$. Lo spazio $\mathcal{S}$ ha quindi dimensione **n-p** e l'asserto è provato.

La proposizione ora provata non solo fornisce **la dimensione** dello spazio $\mathcal{S}$ delle soluzioni del sistema S ma offre anche un ***metodo per determinare una sua base.***

Facciamo un esempio.

Si vogliano determinare le soluzioni del seguente sistema omogeneo:

$$\mathbf{S} = \begin{cases} 2x + y + z + 2t = 0 \\ x + y + z + t = 0 \end{cases}$$

La matrice di tale sistema

$$\mathbf{A} = \begin{pmatrix} 2, 1, 1, 2 \\ 1, 1, 1, 1 \end{pmatrix}$$

ha rango due risultando det $\begin{pmatrix} 2, 1 \\ 1, 1 \end{pmatrix} = 1$ e pertanto lo spazio $\mathcal{S}$ delle soluzioni è un sottospazio di $\mathbf{R}^4$ di dimensione due. Una base di $\mathcal{S}$ si ottiene considerando le due soluzioni di S che si ottengono assegnando alle incognite z e t una prima volta i valori (1 , 0) e successivamente i valori (0 , 1). Troviamo quindi queste due particolari soluzioni .

Ponendo $z = 1$ e $t = 0$, **S** diventa

$$\begin{cases} 2x + y = -1 \\ x + y = -1 \end{cases}$$

Applicando a tale sistema la regola di Cramer si trova x = 0 ed y = -1. Quindi la prima soluzione di S è la quaterna $\zeta_1 = (0 , -1 , 1, 0)$

Ponendo $z = 0$ e $t = 1$, **S** diventa

$$\begin{cases} 2x + y = -2 \\ x + y = -1 \end{cases}$$

Applicando a tale sistema la regola di Cramer si trova $x = -1$ ed $y = 0$. Quindi la seconda soluzione di **S** è la quaterna $\zeta_2 = (-1 , 0 , 0, 1)$

Lo spazio $\mathcal{S}$ delle soluzioni di **S** è quindi lo spazio generato dalle due quaterne trovate.

$$\mathcal{S} = [(0 , -1 , 1, 0) ,(-1 , 0 , 0, 1)]$$

Ciò che ora faremo giustifica l'interesse mostrato per la risoluzione di un sistema omogeneo.

Sia assegnato un sistema di equazioni lineari non omogeneo

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

e supponiamo sia compatibile in modo da avere interesse a ricercare le sue soluzioni.

Ponendo nel sistema **S** i termini noti $c_1 , c_2 ,\dots , c_n$ eguali a zero si ottiene un sistema omogeneo

$$\mathbf{S_o} \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = 0 \end{cases}$$

che è detto ***associato*** ad **S** .

Abbiamo gia' visto che nel caso sia **p=p' = n** il sistema **S** ha una sola soluzione che si trova con la regola di Cramer. Se invece è **p < n** allora il sistema **S** ammette infinite soluzioni ognuna delle quali va trovata assegnando a certe n-p incognite dei valori arbitrari ed applicando al corrispondente sistema la regola di Cramer . Sembrerebbe così che per trovare tutte le soluzioni di

S sia necessario applicare la regola di Cramer *molte volte* ! Per fortuna non è così come ora mostreremo.

Ci viene incontro la seguente

Proposizione 2.2 *Sia assegnato un sistema*

$$\mathbf{S} = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = c_2 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = c_m \end{cases}$$

non omogeneo e compatibile . Tutte le soluzioni di S si ottengono sommando ad una sua soluzione tutte le soluzioni del sistema

$$\mathbf{S_o} = \begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = 0 \end{cases}$$

omogeneo associato ad S .

Dimostrazione. Per rendere più semplice la dimostrazione usiamo la forma matriciale $A\underline{X} = \underline{C}$ e $A\underline{X} = \underline{0}$ per rappresentare il sistema **S** ed il suo omogeneo associato. Indichiamo inoltre con $\underline{\zeta}$ una soluzione di **S** , con $\mathcal{S}$ l'insieme di tutte le soluzioni di **S** e con $\mathcal{S}_0$ il sottospazio delle soluzioni di $\mathbf{S_o}$.

Se $\underline{\eta}$ è una soluzione di $\mathbf{S_o}$ si ha $A\,\underline{\eta} = \underline{0}$. Poiché $\underline{\zeta}$ è una soluzione di **S** si ha $A\,\underline{\zeta} = \underline{C}$ ed allora si ha

$$A(\underline{\zeta} + \underline{\eta}) = A\,\underline{\zeta} + A\,\underline{\eta} = \underline{C} + \underline{0} = \underline{C}$$

Abbiamo così mostrato che addizionando alla soluzione $\underline{\zeta}$ una soluzione di $\mathbf{S_o}$ si ottiene ancora una soluzione di **S .** Se mostriamo che una qualunque soluzione di **S** si ottiene in questo modo allora il teorema è provato. Sia quindi $\underline{\zeta}'$ una soluzione di **S** . Poiché $\underline{\zeta}$ e $\underline{\zeta}'$ sono soluzioni di S si ha $A\,\underline{\zeta} = \underline{C}$ e $A\,\underline{\zeta}' = \underline{C}$ e quindi

$$A(\underset{\sim}{\zeta}' - \underset{\sim}{\zeta}) = A\underset{\sim}{\zeta}' - A\underset{\sim}{\zeta} = \underline{C} - \underline{C} = 0$$

Pertanto $\underset{\sim}{\eta} = \underset{\sim}{\zeta}' - \underset{\sim}{\zeta}$ è una soluzione di $\mathbf{S_o}$ che sommata a $\underset{\sim}{\zeta}$ fornirà la soluzione $\underset{\sim}{\zeta}'$ scelta.

In definitiva sommando a $\underset{\sim}{\zeta}$ tutte le soluzioni $\underset{\sim}{\eta}$ di $\mathbf{S_o}$ si ottengono tutte le soluzioni di **S** e l'asserto è così provato.

La proposizione ora provata risolve il problema che avevamo posto circa l'incombenza di dover applicare la regola di Cramer molte volte per trovare le soluzioni di S.

Infatti per trovare $\underset{\sim}{\zeta}$ si applicherà la regola di Cramer una sola volta e per trovare le soluzioni di $\mathbf{S_o}$ basterà trovare una sua base e per la sua determinazione si dovrà applicare la regola di Cramer altre **n-p** volte.

Concludiamo tale numero esaminando come determinare in modo rapido le soluzioni di un particolare sistema omogeneo che spesso incontreremo nelle applicazioni successive.

Si consideri un sistema omogeneo $\mathbf{S_o}$ con ***n-1 equazioni indipendenti ed n incognite*** .

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 .. \qquad + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 .. \qquad + a_{2n}x_n = 0 \\ .. \\ a_{n-1,1}x_1 + a_{n-1,2}\, x_2 + \ldots + a_{n-1,n}x_n = 0 \end{cases}$$

Poichè il rango è n-1 la matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots & a_{1n} \\ a_{21} & a_{22} \ldots\ldots & a_{2n} \\ & & \cdots \\ \ldots & & \\ a_{n-1,1}, & a_{n-1,2}, \ldots & a_{n-1,n} \end{pmatrix}$$

del sistema possiede almeno una sua sottomatrice quadrata d'ordine n-1 con il determinante diverso da zero. Ora le soluzioni del sistema formano un sottospazio di dimensione 1 per cui per determinare tutte le sue soluzioni basta determinare una sua soluzione $(y_1, y_2, \ldots, y_n)$ ***non nulla***.

Denotiamo con L_1, L_2, …, L_n le n matrici che si ottengono dalla matrice A cancellando rispettivamente la prima , la seconda ,…, l'ultima colonna. Ognuna di tali matrici è quadrata d'ordine n-1 e quindi di ognuna di esse si può calcolare il determinante. Indichiamo con

$$\lambda_1 = \det L_1 \ , \quad \lambda_2 = \det L_2 \ , \ \dots \ , \quad \lambda_n = \det L_n$$

Come già osservato almeno uno dei numeri λ_1 , λ_2 , … , λ_n non è zero . Vogliamo provare che la n-pla non nulla

$$(y_1 , y_2 ,\dots, y_n) = (\lambda_1 \ , \ - \lambda_2 \ , \ \dots \ , \ (-1)^{n-1} \lambda_n)$$

che si ottiene considerando gli n numeri (λ_1, λ_2, … , λ_n) ma **presi a segno alterno** è una soluzione del sistema . Occorre provare che (λ_1, - λ_2 , … , $(-1)^{n-1}\lambda_n$) è soluzione di ogni equazione del sistema e che quindi se in ogni equazione del sistema si sostituiscono al posto delle incognite (x_1, x_2, …., x_n) i numeri (λ_1, - λ_2 , … , $(-1)^{n-1}\lambda_n$) si ottiene zero.

Vediamo.

Consideriamo le n-1 matrici **quadrate d'ordine n** , M_1. M_2 , … , M_{n-1}, così ottenute:

M_1 ha come **prima riga** la ***prima riga*** di A e le altre n-1 righe eguali alle righe di A.

M_2 ha come **prima riga** la ***seconda riga*** di A e le altre n-1 righe eguali alle righe di A.

…

….

M_{n-1} ha come **prima riga** la ***riga di posto n-1*** di A e le altre n-1 righe eguali alle righe di A.

$$M_1 = \begin{pmatrix} a_{11} & a_{12}\dots\dots & a_{1n} \\ a_{11} & a_{12}\dots\dots & a_{1n} \\ a_{21} & a_{22}\dots\dots & a_{2n} \\ & & \dots \\ \dots & & \\ a_{n-1,1}, & a_{n-1,2}, \dots & a_{n-1,n} \end{pmatrix}$$

$$M_2 = \begin{pmatrix} a_{21} & a_{22} \ldots\ldots & a_{2n} \\ & & \cdots \\ a_{11} & a_{12} \ldots\ldots & a_{1n} \\ a_{21} & a_{22} \ldots\ldots & a_{2n} \\ & & \cdots \\ \ldots & & \\ a_{n-1,1}, & a_{n-1,2}, \ldots & a_{n-1,n} \end{pmatrix}$$

.

.

.

$$M_{n-1} = \begin{pmatrix} a_{n-1,1}, & a_{n-1,2}, \ldots & a_{n-1,n} \\ a_{11} & a_{12} \ldots\ldots & a_{1n} \\ a_{21} & a_{22} \ldots\ldots & a_{2n} \\ & & \cdots \\ \ldots & & \\ a_{n-1,1}, & a_{n-1,2}, \ldots & a_{n-1,n} \end{pmatrix}$$

Ognuna di tali matrici avendo ciascuna due righe eguali ha il **determinante eguale a zero**.

Sviluppando det M_i $i = 1,2,.., n-1$ secondo gli elementi della prima riga si ha :

$$\det M_i = a_{i1} h_1 + a_{i2} h_2 + \ldots, + a_{in} h_n = 0$$

avendo indicato con $(h_1, h_2, \ldots, , h_n)$ i complementi algebrici degli elementi della prima riga . Ma i numeri (λ_1 , $-\lambda_2$, … , $(-1)^{n-1}$ λ_n) sono esattamente i complementi algebrici degli elementi della prima riga e quindi si ha l'asserto.

Concludiamo con un esempio .

Si vogliano determinare le soluzioni del seguente sistema omogeneo.

$$S = \begin{cases} 2x + 2y + 3z = 0 \\ 2x + y - 4z = 0 \end{cases}$$

Questo sistema ha due equazioni indipendenti e tre incognite e quindi le terne che lo soddisfano costituiscono un sottospazio di $\mathbf{R}^3$ di dimensione 1. Una terna non nulla che quindi fornisce una base per lo spazio delle soluzioni si ottiene considerando i minori d'ordine due e presi a segno alterno della matrice

$$A = \begin{pmatrix} 2, & 2, & 3 \\ 2, & 1, & -4 \end{pmatrix}$$

dei coefficienti. Quindi una base per lo spazio delle soluzioni è la terna (-11 , 14 , -2) . Poiché tale spazio ha dimensione 1 le sue soluzioni si ottengono considerando la terna (-11 , 14 , -2) e tutte quelle ad essa proporzionali.

C A P I T O L O IV

Prodotti scalari

1. Prodotti scalari.

In questo capitolo Vn denoterà sempre uno spazio vettoriale di dimensione n definito sul campo **reale**. Un prodotto scalare in V_n è una funzione

$$s : V_n \times V_n \rightarrow R$$

di $V_n \times V_n$ in R , *simmetrica e bilineare* , cioè avente le seguenti proprietà:

1. $s(v , w) = s(w , v)$

2. $s(v+v' ,w) = s(v, w) + s(v', w)$

3. $s(\alpha v , w) = \alpha s(v , w)$

per ogni coppia di vettori v e w e per ogni scalare α .

Il numero reale s(v,w) associato da s alla coppia ordinata (v,w) è chiamato *prodotto scalare di v e w* .

Dalle proprietà 1 , 2, 3 segue facilmente che s verifica inoltre le seguenti proprietà

2'. $s(v , w+w') = s(v, w) + s(v , w')$

3'. $s(v , \alpha w) = \alpha s(v , w)$

e ciò giustifica il termine *bilineare* dato ad **s**.

Inoltre dalla 3' segue che

$$s(v ,-w) = s(v , (-1)w) = - s (v ,w)$$

Facilmente possiamo provare ora la seguente

Proposizione 1.1 *Il prodotto scalare di s(v , w) di v con w è zero se uno dei due vettori è il vettore nullo.*

Dimostrazione. Supponiamo ad esempio sia $w = \underline{0}$. Si ha

$$s(v ,\underline{0}) = s(v , a + (-a))= s(v ,a) + s(v ,-a)= s(v,a) - s(v,a) = 0$$

Il prodotto scalare s si dice *definito positivo* se vale la seguente ulteriore proprietà

4. $s(v , v) \geq 0$ ed inoltre $s(v , v) = 0$ se e solo se è $v = \underline{0}$

cioè :
il prodotto scalare di un vettore v *con se stesso è un numero non negativo ed è zero se e solo se* v *è il vettore nullo.*

Ci sono due esempi importanti.

Esempio 1. Si consideri lo spazio vettoriale R^n delle n-ple ordinate di numeri reali. Come già detto, si chiama *prodotto scalare* dei vettori $\underline{a} = (a_1, a_2,\dots,a_n)$ e $\underline{b} = (b_1, b_2,\dots,b_n)$ il numero reale

$$< \underline{a} , \underline{b} > = a_1, b_1 + a_2 b_2 + \dots + a_n b_n$$

che si ottiene eseguendo *la somma dei prodotti degli elementi di egual posto*. Si verifica facilmente che la funzione

$$< \quad > : R^n \times R^n \rightarrow R$$

sopra definita è un prodotto scalare definito positivo in R^n . Tale prodotto scalare è chiamato *standard o euclideo.*

Esempio 2 . Si consideri un punto A dello spazio reale e sia V_A l'insieme di tutti i segmenti orientati AP di primo estremo A. Come già visto in precedenza V_A è uno spazio vettoriale sul campo reale di dimensione tre .In tale spazio vettoriale si definisce prodotto scalare dei vettori $v = AP$ e $w = AP'$ il seguente numero reale

$$s (v , w) = | AP | \; | AP'| \cos \vartheta$$

avendo indicato con $|AP|$ e $|AP'|$ le lunghezze dei segmenti AP ed AP' e con ϑ l'angolo formato dalle rette AP ed AP' .

Si prova che la funzione $s : V_A \times V_A \to R$ sopra definita è un prodotto scalare definito positivo dello spazio vettoriale V_A .

Riferendoci a tale esempio geometrico osserviamo quanto segue : se v = AP è un vettore si ha in base alla definizione **:**

$$s(v, v) = |AP|^2$$

Da cui segue quindi che

$$|AP| = \sqrt{s(v,v)} \tag{i}$$

Inoltre se v = AP e w = AP' sono due vettori non nulli si ha :

(ii) v è *ortogonale* a w se e soltanto se risulta s(v , w) = 0.

Le proprietà (i) e (ii) ora illustrate suggeriscono le seguenti definizioni .

Sia V_n uno spazio vettoriale reale dotato di un prodotto scalare s *definito positivo*.

Chiameremo *lunghezza* o *norma* di un vettore v il seguente numero reale | v | non negativo

$$\mathbf{|v| = \sqrt{s(v,v)}} \tag{1 . 1}$$

Inoltre se v e w sono due qualunque vettori di V_n diremo che essi sono *ortogonali* e scriveremo $v \perp w$ se risulta s(v , w) = 0. In simboli

$$v \perp w \quad \Leftrightarrow \quad s(v, w) = 0 \tag{1.2}$$

Ovviamente per quanto già provato nella proposizione 1.1 si ha :

Proposizione 1.2 *Il vettore nullo ha lunghezza zero ed è ortogonale ad ogni altro vettore.*

Se v e w sono due vettori tra loro ortogonali anche i vettori αv e βw , per ogni coppia di scalari α e β , sono tra loro ortogonali. Risulta infatti

$$s(\alpha v, \beta w) = \alpha\beta\, s(v, w) = 0.$$

Quindi " accorciando " o " allungando" due vettori v e w tra loro ortogonali essi restano ortogonali.

Quando si moltiplichi un vettore non **nullo** v per il numero $k = \frac{1}{|\mathbf{v}|}$ il vettore

$\mathbf{w} = \frac{1}{|\mathbf{v}|}\mathbf{v}$ che si ottiene ha lunghezza eguale ad 1 . Infatti risulta

$$|\mathbf{w}| = \sqrt{\mathbf{s(w,w)}} = \sqrt{\mathbf{s(kv,kv)}} = \sqrt{\mathbf{k^2 s(v,v)}} = k\sqrt{\mathbf{s(v,v)}} = \frac{1}{|\mathbf{v}|}|\mathbf{v}| = 1.$$

Di qui in avanti V_n è uno spazio vettoriale reale dotato di un prodotto scalare **s** ***definito positivo***.

Molto importante è la seguente

Proposizione 1.3 *Siano v_1 , v_2 ,……v_h h vettori , ciascuno non nullo, e a due a due ortogonali. Allora i vettori v_1 , v_2 ,……v_h risultano linearmente indipendenti.*

Dimostrazione. Siano $\alpha_1, \alpha_2, \ldots \alpha_h$ h scalari per i quali risulta

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h = \underline{0}$$

Se proviamo che è $\alpha_1 = \alpha_2 = \ldots = \alpha_h = 0$ allora i vettori v_1 , v_2 ,……v_h sono indipendenti e l'asserto sarà provato. Proviamo ad esempio che è $\alpha_1 = 0$. Si ha , tenendo conto della proposizione 1.1 e della bilinearità di s ,

(*) $\quad 0 = s(\underline{0}, v_1) = s(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_h v_h, v_1) =$

$= \alpha_1 s(v_1, v_1) + \alpha_2 s(v_2, v_1) + \dots \dots\dots + \alpha_h s(v_h, v_1)$

ora essendo per ipotesi v_1 non nullo ed ortogonale a tutti gli altri vettori si ha

$$s(v_1, v_1) > 0 \quad e \quad s(v_2, v_1) = \dots = s(v_h, v_1) = 0 \ .$$

Dalla relazione (*) trovata segue allora $\alpha_1 s(v_1, v_1) = 0$ e quindi $\alpha_1 = 0$ essendo

$s(v_1, v_1) > 0$.

Con lo stesso ragionamento , attraverso il calcolo di $s(\underline{0}, v_2)$, …, $s(\underline{0}, v_h)$ si prova che è anche $\alpha_2 = \dots\dots = \alpha_h = 0$ e quindi l'asserto.

La proposizione ora provata mostra che se siamo in grado di trovare nello spazio V_n , n vettori non nulli $e_1, e_2, \dots, e_n$ a due a due ortogonali allora tali vettori essendo indipendenti costiuiscono una base. Una siffatta base sarà detta ***ortogonale*** . Inoltre come abbiamo già osservato è anche possibile rendere i vettori $e_1, e_2, \dots, e_n$ ciascuno di lunghezza 1 senza alterare la reciproca ortogonalità . Una base $e_1, e_2, \dots, e_n$ che abbia i vettori a due a due ortogonali e ciascuno di lunghezza 1 è detta ***ortonormale.***

Vedremo in seguito che è molto utile disporre di una base ortonormale. Ma è sempre possibile costruire una base ortonormale ?

Viene incontro a questa nostra esigenza la seguente proposizione la quale offre un procedimento costruttivo per costruire una base ortogonale e quindi una base ortonormale .

Proposizione 1.4 *Sia V_n uno spazio vettoriale reale dotato di un prodotto scalare* ***s*** *definito positivo. Se $e_1, e_2, \dots, e_t$ sono t ($1 \le t < n$) vettori non nulli e a due a due ortogonali è possibile determinare un vettore non nullo e_{t+1} ortogonale ai vettori $e_1, e_2, \dots, e_t$.*

Dimostrazione . Sia w_{t+1} un vettore non appartenente al sottospazio $H = [e_1, e_2, \dots, e_t]$ generato dai vettori $e_1, e_2, \dots, e_t$ .Poichè il vettore w_{t+1} non appartiene al sottospazio $H = [e_1, e_2, \dots, e_t]$, per ogni scelta degli scalari $\alpha_1, \alpha_2, \dots\dots, \alpha_t$ il vettore

$$e_{t+1} = w_{t+1} + \alpha_1 e_1 + \alpha_2 e_2 + \dots\dots + \alpha_t e_t$$

non è nullo e non appartiene ad H . Vediamo se è possibile scegliere gli scalari $\alpha_1, \alpha_2, \dots\dots, \alpha_t$ in modo che il vettore e_{t+1} risulti ortogonale ai vettori $e_1, e_2, \dots, e_t$.

Tenendo conto della bilinearità del prodotto scalare e del fatto che i vettori $e_1, e_2, \ldots, e_t$ sono a due a due ortogonali si ha :

$s(e_{t+1}, e_1) = s(w_{t+1}, e_1) + \alpha_1 s(e_1, e_1)$

$s(e_{t+1}, e_2) = s(w_{t+1}, e_2) + \alpha_2 s(e_2, e_2)$

…..

……

$s(e_{t+1}, e_t) = s(w_{t+1}, e_t) + \alpha_t s(e_t, e_t)$

Pertanto il vettore e_{t+1} risulterà ortogonale ai vettori $e_1, e_2, \ldots, e_t$ cioè risulterà

$$s(e_{t+1}, e_1) = s(e_{t+1}, e_2) = \ldots = s(e_{t+1}, e_t) = 0$$

quando si scelga :

$$\alpha_1 = -\frac{s(w_{t+1}, e_1)}{s(e_1, e_1)}$$

$$\alpha_2 = -\frac{s(w_{t+1}, e_2)}{s(e_2, e_2)}$$

…..

…..

$$\alpha_t = -\frac{s(w_{t+1}, e_t)}{s(e_t, e_t)}$$

Ovviamente quando ai vettori $e_1, e_2, \ldots, e_t$ sia stato aggiunto il vettore e_{t+1} ora costruito si ottiene un insieme di t+1 vettori $e_1, e_2, \ldots, e_t, e_{t+1}$ ancora *a due a due ortogonali* sui quali si può quindi ripetere il procedimento ora illustrato così da costruire un altro vettore e_{t+2} ortogonale ad essi .

Dopo n-t passi si perverrà quindi a costruire n vettori $e_1, e_2, \ldots, e_n$ a due a due ortogonali e cioè, come si voleva, una *base ortogonale* .

Indaghiamo ora perché sia utile disporre nello spazio vettoriale V_n dotato di un prodotto scalare s definito positivo, di un riferimento ortonormale cioè di una sua base $B = (e_1, e_2, \ldots, e_n)$

ordinata e ortonormale.

Siano v e w due qualsiasi vettori di V_n . Poichè $B = (e_1, e_2, \ldots, e_n)$ è una base esistono opportuni scalari $(\alpha_1, \alpha_2, \ldots\ldots, \alpha_n)$ e $(\beta_1, \beta_2, \ldots\ldots, \beta_n)$ per cui si abbia

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n \ , \qquad w = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n$$

Tenendo conto che la base B è ortonormale e che quindi è :

$s(e_i, e_i) = 1$ per ogni $i = 1, 2, \ldots, n$

$s(e_i, e_j) = 0$ per $i \neq j$

si ha

$$s(v, w) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \ldots\ldots + \alpha_n \beta_n \ .$$

Concludendo, attraverso l'uso del riferimento ortonormale B, abbiamo reso semplice ed elementare il calcolo del numero $s(v, w)$ *risultando tale prodotto scalare di v e w eguale al prodotto scalare euclideo delle n-ple* $(\alpha_1, \alpha_2, \ldots\ldots, \alpha_n)$ *e* $(\beta_1, \beta_2, \ldots\ldots, \beta_n)$ *delle loro coordinate nel riferimento B.*

Sia sempre V_n uno spazio vettoriale reale dotato di un prodotto scalare s definito positivo .

Sia ora H un sottospazio di V_n .

Proviamo la seguente

Proposizione 1.5 . *Il sottoinsieme*

$$H^{\perp} = \{ w \in V_n \ : w \perp v \ \text{ per ogni } \ v \in H \}$$

costituito dai vettori w ortogonali ad ogni vettore di H è un sottospazio . I sottospazi H ed $H^{\perp}$ *sono tra loro supplementari .*

Dimostrazione . Il sottoinsieme $H^{\perp}$ è non vuoto perché di esso fa parte il vettore nullo. Siano w e w' due vettori di $H^{\perp}$. Per ogni vettore v di H si ha

$$s(w + w', v) = s(w, v) + s(w', v) = 0$$

e questa prova , per l'arbitrarietà di v , che anche $w + w'$ è un vettore di $H^{\perp}$.

Siano w un vettore di $H^{\perp}$ e sia α un numero reale . Risulta , per ogni vettore v di H ,

$$s(\alpha w , v) = \alpha s(w , v) = 0$$

e questa prova , per l'arbitrarietà di v, che anche αw è un vettore di $H^{\perp}$.

Abbiamo così provato che $H^{\perp}$ è un sottospazio. Proviamo ora che H ed $H^{\perp}$ sono tra loro supplementari . L'asserto è ovvio se H è banale . Supponiamo quindi H non banale . Sia $t = \dim H$ e sia $e_1, e_2,\dots, e_t$ una base ortogonale di H. Completiamo la base $e_1, e_2,\dots, e_t$ ortogonale di H in una base $B = \{ e_1, e_2,\dots, e_t \quad e_t \ ,e_{t+1} \dots e_n \}$ ortogonale di V_n aggiungendo opportuni n-t vettori $e_{t+1} \dots e_n$. In forza della proposizione 5.2 del capitolo I l'asserto sarà provato se mostriamo che $H^{\perp} = [e_{t+1} \dots e_n]$.

Denotiamo con $L = [e_{t+1} \dots e_n]$ il sottospazio generato dai vettori $e_{t+1} \dots e_n$. Se w è un vettore di L si ha $w = \beta_{t+1}e_{t+1} + \dots\dots + \beta_n e_n$ e conseguentemente per ogni vettore $v = \alpha_1e_1 + \alpha_2e_2 + \dots + \alpha_n e_t$ di H si ha

$$s (w , v) = s(\beta_{t+1}e_{t+1} + \dots\dots + \beta_n e_n \ , \alpha_1e_1 + \alpha_2e_2 + \dots + \alpha_n e_t \ ) = 0$$

essendo $s (e_i , e_j) = 0$ per $i \neq j$. Pertanto per l'arbitrarietà di v, il vettore w è un vettore di $H^{\perp}$.Abbiamo così provato che è $L \subseteq H^{\perp}$. Viceversa sia w un vettore di $H^{\perp}$. Poiché B è una base si ha che esistono opportuni scalari $\alpha_1,\alpha_2, \dots \alpha_n$ per cui risulti

$$(*) \qquad w = \alpha_1e_1 + \alpha_2e_2 + \dots + \alpha_n e_t + \alpha_{t+1}e_{t+1} + \dots\dots + \alpha_n e_n$$

Poichè w è un vettore di $H^{\perp}$ esso risulta ortogonale ai vettori $e_1 ,e_2 \dots\dots e_t$ di H .

Si ha così $s (w , e_i) = 0$ per ogni $i = 1,2,\dots, t$.

Ma poiché è $\qquad 0 = s (w , e_i) = \alpha_i \ s (e_i , e_i)$

si ha $\alpha_i = 0$ per ogni $i = 1,2,\dots, t$ essendo $s (e_i , e_i) > 0$.

Ma se $\alpha_1=\alpha_2= \dots =\alpha_t = 0$ allora da (*) segue

$$w = \alpha_{t+1}e_{t+1} + \dots\dots + \alpha_n e_n$$

la quale mostra che w è un vettore di L . Avendo quindi provato che è anche $H^{\perp} \subseteq L$ si ha

$L = H^{\perp}$ e la proposizione è provata.

Il sottospazio $H^{\perp}$ ora illustrato è detto *complemento ortogonale di H* .

Osserviamo infine che è facile provare la seguente equivalenza

Proposizione 1.6 *Sia V_n uno spazio vettoriale reale dotato di un prodotto scalare s definito positivo . Siano $e_1, e_2, \dots, e_t$, t suoi vettori ed $H = [e_1, e_2, \dots, e_t]$ il sottospazio da essi generato. Un vettore w è ortogonale a ciascuno dei vettori $e_1, e_2, \dots, e_t$ se e solo se w appartiene al sottospazio $H^{\perp}$.*

Concludiamo con qualche utile applicazione delle cose provate in questo numero.

Si consideri un sistema omogeneo di **m** equazioni lineari ciascuna in **n** incognite.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 .. + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 .. + a_{2n}x_n = 0 \\ .. \\ a_{m1}x_1 + a_{m2}x_2 .. + a_{mn}x_n = 0 \end{cases}$$

e sia

$$A = \begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} \dots\dots\dots a_{2n} \\ \dots & \\ a_{m1} & a_{m2} \dots\dots\dots a_{mn} \end{pmatrix}$$

la matrice del sistema.

Una n-pla $(y_1, y_2, \dots, y_n)$ è soluzione del sistema S se e solo se essa è un vettore di R^n ortogonale ai vettori riga della matrice A. Tenendo conto della proposizione 1.6 , l'insieme $\mathcal{S}$ delle soluzioni di S coincide quindi col *sottospazio* ortogonale $H^{\perp}$ del sottospazio

$H = [\underline{a}_1, \underline{a}_2, \dots, \underline{a}_m]$ generato dalle righe di A.

Se la matrice A ha rango p allora $\dim H = p$ e così $\dim \mathcal{S} = \dim H^{\perp} = n-p$.

Abbiamo così ritrovato il risultato già stabilito per i sistemi omogenei nella proposizione 2.1 del capitolo III.

C A P I T O L O V

Triangolazione di una matrice quadrata e diagonalizzazione di un endomorfismo.

1. Triangolazione di una matrice quadrata.

Abbiamo visto nei capitoli precedenti quanto sia utile il calcolo del determinante di una matrice quadrata A d'ordine n ad elementi su un fissato campo K. Allo stesso tempo è apparso chiaro come tale calcolo sia poco agevole quando l'ordine n di A sia molto grande.
Vogliamo ora illustrare un procedimento, traducibile in un programma per il computer ,che consenta di semplificare tale calcolo.
Sia quindi

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots\ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots\ldots\ldots & a_{2n} \\ \ldots\ldots\ldots & & & \\ & & & \\ a_{n1} & a_{n2} & \ldots\ldots\ldots\ldots & a_{nn} \end{pmatrix}$$

una matrice quadrata d'ordine n ($n \geq 2$) ad elementi in un fissato campo K.
Supporremo ovviamente che **nessuna riga di A e nessuna colonna di A sia il vettore nullo** altrimenti è già detA = 0.
Per comodità di scrittura indicheremo con $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2, \ldots, \underline{\mathbf{a}}_n$ le n righe di A.
Abbiamo già osservato parlando delle proprietà dei determinanti , che quando nella matrice A una sua riga $\underline{\mathbf{a}}_i$ venga sostituita con la nuova riga $\underline{\mathbf{a}}_i + \alpha \underline{\mathbf{a}}_j$, con α scalare non nullo e $j \neq i$, si ottiene una nuova matrice A' che differisce da A solo nella riga di posto i ma che ha lo stesso determinante di A. Quando si agisce sulla matrice A nel modo ora indicato si dice che si è effettuata su A una *trasformazione elementare* sulle sue righe. Noi opereremo per semplicità ed omogeneità solo sulle righe di A pur consapevoli che il procedimento che illustreremo potrebbe essere effettuato anche sulle colonne di A.

Useremo questa simbologia

$$\begin{pmatrix} a_{11} & a_{12} & \ldots\ldots\ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots\ldots\ldots & a_{2n} \\ & \ldots\ldots\underline{a}_i\ldots\ldots & & \\ & & & \\ a_{n1} & a_{n2} & \ldots\ldots\ldots\ldots & a_{nn} \end{pmatrix} \xrightarrow{\mathbf{T}_i \;\; \underline{\mathbf{a}}_i + \alpha \underline{\mathbf{a}}_j} \begin{pmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \ldots\ldots\ldots\ldots & \mathbf{a}_{1n} \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \ldots\ldots\ldots\ldots & \mathbf{a}_{2n} \\ & \ldots\ldots\underline{\mathbf{a}}_i + \alpha\underline{\mathbf{a}}_j\ldots & & \\ & & & \\ \mathbf{a}_{n1} & \mathbf{a}_{n2} & \ldots\ldots\ldots\ldots & \mathbf{a}_{nn} \end{pmatrix}$$

per indicare che stiamo trasformando la riga $\underline{\mathbf{a}}_i$ di posto i con la nuova riga $\underline{\mathbf{a}}_i + \alpha\, \underline{\mathbf{a}}_j$.

Vogliamo ora provare che

Proposizione 1.1 *Con un numero finito di trasformazioni elementari sulle righe è possibile trasformare la matrice A in una matrice B che ha lo stesso determinante di A ma che ha l'ultima colonna eguale a (0, 0, ..., 1).*

In simboli

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} \end{pmatrix} \text{-- } \mathbf{T_1} \text{-----} \ \mathbf{T_2} \ldots \text{-----} \ \mathbf{T_k} \ \text{---}> \mathbf{B}=\begin{pmatrix} b_{11} & b_{12} \ldots\ldots\ldots\ldots 0 \\ b_{21} & b_{22} \ldots\ldots\ldots\ldots 0 \\ & \\ & \\ b_{n1} & b_{n2} \ldots\ldots\ldots\ldots 1 \end{pmatrix}$$

Una volta provata la proposizione 1.1 possiamo evidenziare i vantaggi che ne derivano.
Il vantaggio è che il determinante della matrice B è eguale al determinante della matrice H che si ottiene cancellando nella matrice B l'ultima sua riga e l'ultima sua colonna.
Quindi con questo primo passo abbiamo ricondotto il calcolo del determinante di A al calcolo del determinante di H che però **ha ordine n-1.**
Se la matrice H non possiede una riga o una colonna nulla possiamo ripetere sulla matrice H il procedimento usato per trasformare A in B e così il determinante di H sarà eguale al determinante di una matrice H' d'ordine n-2. Avremo quindi

$$\det A = \det B = \det H = \det H'$$

Possiamo così ,iterando il procedimento un certo numero di volte, ricondurre il calcolo del determinante di A , forse *molto grande* , al calcolo del determinante di una matrice A' , *abbastanza piccola.*

Proviamo quindi la proposizione 1.1

Dimostrazione . Iniziamo col caso più favorevole:

Caso 1 : $\mathbf{a_{nn} = 1}$

Se $a_{nn} = 1$ allora il risultato si raggiunge facilmente. Infatti se la riga $\underline{\mathbf{a}}_i$ di posto i ($i < n$) ha l'ultimo elemento diverso da zero ed eguale ad h allora sommando alla riga $\underline{\mathbf{a}}_i$ l'ultima riga moltiplicata per -h si ottiene una nuova matrice B che ha l'elemento $b_{in} = 0$,come desiderato.

In simboli :

$$\begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} \dots\dots\dots a_{2n} \\ \dots\dots \underline{a}_i \dots\dots & h \\ \\ a_{n1} & a_{n2} \dots\dots\dots 1 \end{pmatrix} \text{----}\ \mathbf{T_i}\ \text{-------}\ \underline{\mathbf{a}}_i + -h\,\underline{\mathbf{a}}_n\ \text{------>} \begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} \dots\dots\dots a_{2n} \\ \dots\dots\dots\dots\dots\dots & 0 \\ \\ a_{n1} & a_{n2} \dots\dots\dots 1 \end{pmatrix}$$

Quindi se nell'ultima colonna di A ci sono s elementi diversi da zero (incluso 1) con s-1 trasformazioni elementari sulle sue righe si può pervenire ad una matrice B con l'ultima colonna eguale a (0,0,…,1) come richiesto.

Caso 2 : $\mathbf{a_{nn} = 0}$

Se $\mathbf{a_{nn} = 0}$ allora poiché l'ultima colonna di A non è il vettore nullo esiste i , $i < n$ tale che sia $a_{in} \neq 0$. Posto $h = a_{in}$ allora sommando all'ultima riga di A la riga $\underline{\mathbf{a}}_i$ moltiplicata per $\frac{1}{h}$ si ottiene una matrice B che l'elemento $b_{nn} = 1$

In simboli :

$$\begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} \dots\dots\dots a_{2n} \\ \dots\dots \underline{a}_i \dots\dots & h \\ \\ a_{n1} & a_{n2} \dots\dots\dots 0 \end{pmatrix} \text{----}\ \mathbf{T_n}\ \text{-------}\ \underline{\mathbf{a}}_n + \frac{1}{h}\underline{\mathbf{a}}_i\ \text{------>} \begin{pmatrix} a_{11} & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} \dots\dots\dots a_{2n} \\ \dots\dots \underline{a}_i \dots\dots & h \\ \\ \dots\dots\dots\dots\dots\dots & 1 \end{pmatrix}$$

Dopo questo primo passo ci siamo quindi ricondotti al ***caso 1*** già esaminato.

Caso 3 : $\mathbf{a_{nn} \neq 0}$ e $\mathbf{a_{nn} \neq 1}$ ed esiste $i < n$ tale che sia $\mathbf{a_{in} \neq 0}$.

Posto $\mathbf{a = a_{nn}}$ e $\mathbf{h = a_{in}}$ allora sommando all'ultima riga di A la riga $\underline{\mathbf{a}}_i$ moltiplicata per $\dfrac{-(a-1)}{h}$ si ottiene una matrice B che l'elemento $b_{nn} = 1$.

In simboli :

$$\begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots \underline{a}_i \ldots\ldots & h \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a \end{pmatrix} \text{---- } \mathbf{T_n} \text{ ------- } \underline{\mathbf{a}}_n + \frac{-(a-1)}{h}\,\underline{\mathbf{a}}_i \text{ ------ >} \begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots \underline{a}_i \ldots\ldots & h \\ & \\ \ldots\ldots\ldots\ldots\ldots\ldots & 1 \end{pmatrix}$$

Dopo questo primo passo ci siamo quindi ricondotti al ***caso 1*** già esaminato.

Infine resta il

Caso4 : $\mathbf{a_{nn} \neq 0}$ e $\mathbf{a_{nn} \neq 1}$ e per ogni $i < n$ è $\mathbf{a_{in} = 0}$ **.**

Sommando alla prima riga di A l'ultima riga di A si ricade nel ***caso 3.***

In simboli

$$\begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots 0 \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots 0 \\ & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} \end{pmatrix} \text{---- } \mathbf{T_1} \text{ ------- } \underline{\mathbf{a}}_1 + \underline{\mathbf{a}}_n \text{ ------ >} \begin{pmatrix} \ldots\ldots\ldots & \ldots\ldots\ldots\ldots a_{nn} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots 0 \\ & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} \end{pmatrix}$$

Riepilogando, il procedimento ora illustrato applicato alla matrice A conduce alla matrice B

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} \end{pmatrix} \text{-- } \mathbf{T_1} \text{ ----- } \mathbf{T_2} \ldots \text{ ----- } \mathbf{T_k} \text{ --- >} \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \ldots\ldots\ldots\ldots 0 \\ b_{21} & b_{22} \ldots\ldots\ldots\ldots 0 \\ & \\ & \\ b_{n1} & b_{n2} \ldots\ldots\ldots\ldots 1 \end{pmatrix}$$

Detta H la matrice formata dalle prime n-1 righe e n-1 colonne di B

$$H = \begin{pmatrix} b_{11} & b_{12} \ldots\ldots\ldots\ldots\ldots & b_{1,n-1} \\ b_{21} & b_{22} \ldots\ldots\ldots\ldots\ldots & b_{2,n-1} \\ & & \\ & & \\ b_{n-1,1} & b_{n-1,2} \ldots\ldots\ldots & b_{n-1,n-1} \end{pmatrix}$$

si ha det A = det B = det H.

Se H ha una riga o una colonna nulla è detH = 0. In caso contrario possiamo ripetere il procedimento sulla matrice H " *normalizzandola* " cioè rendendo l'ultima sua colonna eguale a (0,0,…,1).

Ora ogni trasformazione elementare fatta sulle righe di H può ottenersi eseguendo la stessa operazione sulle corrispondenti righe di B che presentano ormai l'ultimo numero eguale a zero. Questa osservazione è in sostanza la dimostrazione della seguente proposizione che è corollario della proposizione 1.1.

Supposto che tutte le matrici H che di volta in volta appaiono si possono normalizzare allora

Proposizione 1. 2 *Con un numero finito di trasformazioni elementari sulle righe è possibile trasformare la matrice A in una matrice B triangolare che ha lo stesso determinante di A che ha* b_{11}*= detA e bii=1* $i \neq 1$

In simboli

$$A=\begin{pmatrix} a_{11} & a_{12} \ldots\ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots\ldots\ldots\ldots a_{2n} \\ \ldots\ldots\ldots & \\ & \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots\ldots a_{nn} \end{pmatrix} \text{-- } \mathbf{T_1} \text{ ----- } \mathbf{T_2} \ldots \text{ ----- } \mathbf{T_k} \text{ ---> } B=\begin{pmatrix} b_{11} & 0 \ldots\ldots\ldots\ldots\ldots 0 \\ b_{21} & 1 \ldots\ldots\ldots\ldots\ldots 0 \\ & \\ & \\ b_{n1} & b_{n2} \ldots\ldots\ldots\ldots\ldots 1 \end{pmatrix}$$

2. ***Diagonalizzazione di un endomorfismo.***

Per ciò che segue è utile introdurre alcune semplici notazioni . Sia V_n uno spazio vettoriale di dimensione finita n sul campo K. Sia

$$A = \begin{pmatrix} \alpha_1 \ \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 \ \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots \\ \alpha_n \ \beta_n \ldots\ldots \gamma_n \end{pmatrix}$$

una matrice quadrata d'ordine n ad elementi nel campo K. Se $(\underline{\mathbf{v}}_1, \underline{\mathbf{v}}_2, \ldots, \underline{\mathbf{v}}_n)$ è una n-pla ordinata di vettori di V_n col simbolo

$$(\underline{\mathbf{v}}_1, \underline{\mathbf{v}}_2, \ldots, \underline{\mathbf{v}}_n) \begin{pmatrix} \alpha_1 \ \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 \ \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots \\ \alpha_n \ \beta_n \ldots\ldots \gamma_n \end{pmatrix} = (\underline{\mathbf{w}}_1, \underline{\mathbf{w}}_2, \ldots, \underline{\mathbf{w}}_n)$$

vogliamo rappresentare la n-pla di vettori $(\underline{\mathbf{w}}_1, \underline{\mathbf{w}}_2, \ldots, \underline{\mathbf{w}}_n)$ dove è :

$\underline{\mathbf{w}}_1 = \alpha_1 \ \underline{\mathbf{v}}_1 + \ \alpha_2 \ \underline{\mathbf{v}}_2 + \ldots + \ \alpha_n \ \underline{\mathbf{v}}_n$

$\underline{\mathbf{w}}_2 = \beta_1 \ \underline{\mathbf{v}}_1 + \ \beta_2 \ \underline{\mathbf{v}}_2 + \ldots + \ \beta_n \ \underline{\mathbf{v}}_n$

....

$\underline{\mathbf{w}}_n = \gamma_1 \ \underline{\mathbf{v}}_1 + \ \gamma_2 \underline{\mathbf{v}}_2 + \ldots\ldots + \gamma_n \ \underline{\mathbf{v}}_n$.

Sia ora M(n, K) l'insieme delle matrici quadrate d'ordine n ad elementi su un fissato campo K.

Le matrici quadrate *non degeneri* di M(n, K) cioè quelle che hanno il determinante diverso da zero formano gruppo rispetto al prodotto (righe per colonne) e tale gruppo , detto *gruppo lineare* , viene indicato col simbolo GL (n,K).

Due matrici A' e A di M(n, K) sono dette *simili* se esiste una matrice P non degenere tale che risulti :

$$A' = P^{-1} A P$$

Sia V_n uno spazio vettoriale sul campo K di dimensione finita n e sia

$$f : V_n \longrightarrow V_n$$

una funzione lineare di V_n in sé . Un vettore $\underline{v}$ **non nullo** è chiamato un ***autovettore*** per la funzione f se è trasformato da f in un vettore ad esso proporzionale , risulta cioè

$$(2.1) \qquad f(\underline{v}) = \lambda \underline{v}$$

Lo scalare λ che figura nella (2.1) è chiamato l'***autovalore*** dell' autovettore $\underline{v}$ ed in ogni caso esso è chiamato un ***autovalore della funzione f*** .

Se λ è un autovalore della funzione f possiamo considerare tutti gli autovettori di f che hanno lo scalare λ come loro autovalore.

Denotiamo quindi con V_λ il seguente sottoinsieme di V_n

$$V_\lambda = \{ \underline{v} \in V_n \ : \ f(\underline{v}) = \lambda \underline{v} \}$$

Il vettore nullo pur non essendo un autovettore per f verifica la (2.1) e quindi esso appartiene al sottoinsieme V_λ .

Il sottoinsieme V_λ, come è facile verificare , è un sottospazio di V_n . Esso è chiamato l' ***autospazio*** corrispondente all'autovalore λ . I vettori non nulli di V_λ sono quindi tutti gli autovettori di f che hanno lo scalare λ come autovalore.

La dimensione del sottospazio V_λ è chiamata la ***molteplicità geometrica dell'autovalore*** $\boldsymbol{\lambda}$.

In seguito si troveranno le giustificazioni del perché abbiamo chiamato la dimensione di V_λ in questo modo. Poiché gli autovettori sono vettori non nulli allora è sempre, per ogni autovalore λ,

$$\dim V_\lambda \geq 1$$

Ovviamente gli autovettori di autovalore zero sono tutti e soli i vettori non nulli del nucleo di f e quindi essi esistono solo quando la funzione non è iniettiva.

Sia

$$A = \begin{pmatrix} a_{11} \ \ a_{12} \ldots a_{1n} \\ a_{21} \ a_{22} \ldots a_{2n} \\ \ldots \\ a_{n1} \ a_{n2} \ldots a_{nn} \end{pmatrix}$$

una matrice quadrata d'ordine n ad elementi nel campo K . Denotiamo con K^n lo spazio vettoriale numerico di dimensione n su K . Se

$$X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

è un elemento di K^n eseguendo il prodotto (righe per colonne) di A per X si ottiene un altro vettore Y

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} \ \ a_{12} \ldots a_{1n} \\ a_{21} \ a_{22} \ldots a_{2n} \\ \ldots \\ a_{n1} \ a_{n2} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

di K^n , dove è esplicitamente :

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n} x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n} x_n \\ \qquad \ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn} x_n \end{cases}$$

La matrice A ci ha quindi consentito di costruire una funzione

$$A : K^n \longrightarrow K^n$$

di K^n in sé che indicheremo egualmente con la lettera A in quanto saremo sempre in grado di comprendere dal contesto se la lettera A rappresenta la matrice o la funzione che essa definisce.
La funzione A ora definita è **lineare** in quanto risulta , per ogni coppia X , X' di vettori di K^n e per ogni scalare α di K ,

1 . $A(X + X') = AX + AX'$
2. $A(\alpha X) = \alpha(AX)$

Indicheremo con $\underline{\mathbf{u}}_1, \underline{\mathbf{u}}_2, \dots, \underline{\mathbf{u}}_n$ i vettori della base canonica di K^n

$$\underline{\mathbf{u}}_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix}, \quad \underline{\mathbf{u}}_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix}, \quad \dots\dots.., \quad \underline{\mathbf{u}}_n = \begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$$

e con $\underline{\mathbf{a}}^1, \underline{\mathbf{a}}^2, \dots, \underline{\mathbf{a}}^n$ i vettori colonna della matrice A .

Poiché una funzione lineare trasforma un sistema di generatori di K^n in un sistema di generatori dello spazio immagine allora i vettori $\underline{\mathbf{a}}^1, \underline{\mathbf{a}}^2, \dots, \underline{\mathbf{a}}^n$ essendo i trasformati tramite la funzione lineare A dei vettori $\underline{\mathbf{u}}_1, \underline{\mathbf{u}}_2, \dots, \underline{\mathbf{u}}_n$ costituiscono un sistema di generatori per lo spazio immagine della funzione lineare A . Ne segue che se la matrice A ha rango p allora lo spazio immagine ha dimensione p e conseguentemente il nucleo della funzione A ha dimensione n-p. Se quindi A è una matrice non degenere e quindi di rango n , la funzione lineare A è biettiva ed è quindi un isomorfismo.
Proviamo che le uniche funzioni lineari di K^n in sé si ottengono come ora abbiamo descritto cioè in corrispondenza ad una matrice quadrata A d'ordine n .
Sia quindi $F : K^n \to K^n$ una funzione lineare di K^n in sé .

Siano

$$\underline{\mathbf{a}}^1 = \begin{pmatrix} a_{11} \\ a_{21} \\ . \\ a_{n1} \end{pmatrix} = F\begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix} \qquad \underline{\mathbf{a}}^2 = \begin{pmatrix} a_{12} \\ a_{22} \\ . \\ a_{n2} \end{pmatrix} = F\begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix}, \quad \dots\dots.., \quad \underline{\mathbf{a}}^n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ . \\ a_{nn} \end{pmatrix} = F\begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$$

i vettori che F associa ai vettori

$$\underline{\mathbf{u}}_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix}, \quad \underline{\mathbf{u}}_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix}, \quad \ldots\ldots.., \quad \underline{\mathbf{u}}_n = \begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$$ della base canonica di K^n .

Sia $A = \begin{pmatrix} a_{11} \ a_{12} \ldots . a_{1n} \\ a_{21} \ a_{22} \ldots .. a_{2n} \\ \ldots .. \\ a_{n1} \ a_{n2} \ldots .. a_{nn} \end{pmatrix}$ la matrice quadrata le cui colonne sono date dai vettori

$\underline{\mathbf{a}}^1 = F(\underline{\mathbf{u}}_1)$, $\underline{\mathbf{a}}^2 = F(\underline{\mathbf{u}}_2)$, .. , $\underline{\mathbf{a}}^n = F(\underline{\mathbf{u}}_n)$ e sia $A: K^n \to K^n$ la funzione lineare

indotta su K^n dalla matrice A.

Poichè le funzioni lineari A ed F assumono gli stessi valori sui vettori di una base risultando

$$\underline{\mathbf{a}}^1 = F(\underline{\mathbf{u}}_1) = A(\underline{\mathbf{u}}_1), \quad \underline{\mathbf{a}}^2 = F(\underline{\mathbf{u}}_2) = A(\underline{\mathbf{u}}_2), \ldots, \underline{\mathbf{a}}^n = F(\underline{\mathbf{u}}_n) = A(\underline{\mathbf{u}}_n)$$

allora esse coincidono .

Sia ora V_n uno spazio vettoriale sul campo K di dimensione finita n e sia

$$f : V_n \longrightarrow V_n$$

una funzione lineare di V_n in sé .

Sia $B = (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots., \underline{\mathbf{e}}_n)$ una base ordinata di V_n . La funzione lineare f è determinata quando si conoscano i vettori $(f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots., f(\underline{\mathbf{e}}_n))$ cioè i trasformati di f dei vettori della base B. Indichiamo con

$(\alpha_1, \alpha_2, \ldots., \alpha_n)$ le coordinate di $f(\mathbf{e}_1)$ nella base B

$(\beta_1, \beta_2, \ldots., \beta_n)$ le coordinate di $f(\mathbf{e}_2)$ nella base B

…..

….

$(\gamma_1, \gamma_2, \ldots., \gamma_n)$ le coordinate di $f(\mathbf{e}_n)$ nella base B e sia $\mathbf{A_f}$ la matrice avente per colonne tali vettori coordinati :

$$\mathbf{A_f} = \begin{pmatrix} \alpha_1 \ \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 \ \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots \\ \alpha_n \ \beta_n \ldots\ldots \gamma_n \end{pmatrix}$$

La conoscenza dei vettori $(f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))$ consente la conoscenza di $\mathbf{A_f}$ e viceversa nota la matrice $\mathbf{A_f}$ sono noti i vettori $(f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))$ e quindi f .
Per tale ragione si dice che la matrice $\mathbf{A_f}$ così costruita ***rappresenta f nella base B.***

Per definizione è quindi

$$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n) \begin{pmatrix} \alpha_1 \ \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 \ \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots \\ \alpha_n \ \beta_n \ldots\ldots \gamma_n \end{pmatrix} = (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))$$

Il teorema che segue evidenzia una notevole proprietà della matrice $\mathbf{A_f}$ che rappresenta f nella base B.

Proposizione 2.1 *Siano* $\mathbf{v}$ *un vettore di V e sia* $f(\mathbf{v})$ *il suo trasformato.*

Siano $(x_1, x_2, \ldots, x_n)$ *le coordinate di* $\mathbf{v}$ *nella base B e siano* $(y_1, y_2, \ldots, y_n)$ *le coordinate di* $f(\mathbf{v})$ *nella base B . La funzione lineare*

$$\mathbf{A_f} : K^n \to K^n$$

trasforma le coordinate di $\mathbf{v}$ *nelle coordinate di* $f(\mathbf{v})$ *. Risulta cioè :*

$$\begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \ \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 \ \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots \\ \alpha_n \ \beta_n \ldots\ldots \gamma_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

Dimostrazione . Con le notazioni adottate si ha :

(i) $$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)\begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = \mathbf{v}$$

e per la linearità di f si ha :

(ii) $$(f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))\begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = f(\mathbf{v})$$

Inoltre è :

$$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)\begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix} = f(\mathbf{v})$$

$$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)\begin{pmatrix} \alpha_1 & \beta_1 & \ldots & \gamma_1 \\ \alpha_2 & \beta_2 & \ldots & \gamma_2 \\ \ldots \\ \alpha_n & \beta_n & \ldots & \gamma_n \end{pmatrix} = (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))$$

Sostituendo tali espressioni nella (ii) si ha :

$$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)\begin{pmatrix} \alpha_1 & \beta_1 & \ldots & \gamma_1 \\ \alpha_2 & \beta_2 & \ldots & \gamma_2 \\ \ldots \\ \alpha_n & \beta_n & \ldots & \gamma_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)\begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix}$$

da cui segue per l'indipendenza dei vettori $(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)$ che è :

$$\begin{pmatrix} \alpha_1 & \beta_1 \ldots\ldots \gamma_1 \\ \alpha_2 & \beta_2 \ldots\ldots \gamma_2 \\ \ldots\ldots & \\ \alpha_n & \beta_n \ldots\ldots \gamma_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix}$$

e cioè l'asserto.

Ricordiamo che la funzione (*coordinazione di V_n nella base B*) che associa ad ogni vettore **v** di V_n il vettore $(x_1, x_2 ,\ldots, x_n)$ delle sue coordinate nella base B è un isomorfismo di V_n in K^n . Indicato con $\mathbf{c} : V_n \to K^n$ tale isomorfismo la proposizione 2.1 ora provata mostra che è :

(2 . 2) $$\mathbf{f} = \mathbf{c}^{-1} \circ \mathbf{A_f} \circ \mathbf{c}$$

La (2. 2) mostra che per calcolare f (**v**) si può utilizzare la matrice A_f che d'ora in poi indichiamo con

$$A_f = \begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots & \\ a_{n1} & a_{n2} \ldots\ldots a_{nn} \end{pmatrix}$$

la quale consente di calcolare le coordinate $(y_1, y_2 ,\ldots, y_n)$ di f (**v**) note che siano le coordinate $(x_1, x_2 ,\ldots, x_n)$ di **v** risultando :

(2 . 3)
$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}\, x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}\, x_n \\ \quad \ldots\ldots\ldots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nn}\, x_n \end{cases}$$

Quando n è "molto grande" le (2. 3) pur consentendo il calcolo di f(**v**) impongono molti calcoli forse " laboriosi " a meno che la matrice A_f non presenti "molti zeri" . Una situazione favorevole si avrebbe se essa avesse la forma diagonale :

$$A_f = \begin{pmatrix} a_{11} & 0 \ldots\ldots & 0 \\ 0 & a_{22} & 0 \ldots\ldots 0 \\ \ldots\ldots & & \\ 0 & 0 \ldots\ldots & a_{nn} \end{pmatrix}$$

perché in tal caso le (2. 3) diventano

$$\begin{cases} y_1 = a_{11}x_1 \\ y_2 = a_{22}x_2 \\ \ldots\ldots \\ y_n = a_{nn}x_n \end{cases}$$

e cioè " estremamente semplici " .

Si pongono allora le seguenti due domande :

a) come cambia la matrice A_f quando si cambia la base ?

b) come deve essere la base B affinchè A_f risulti di forma diagonale ?

I teoremi che seguono danno risposta ai quesiti posti .

Proposizione 2.2 . *Sia* $f : V_n \rightarrow V_n$ *una funzione lineare di V_n in sé . Siano*

$B = (\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n)$ *e* $B' = (\underline{e'}_1, \underline{e'}_2, \ldots, \underline{e'}_n)$

due basi ordinate di V_n . Le due matrici A_f ed A_f' che rappresentano f nelle basi B e B' rispettivamente risultano tra loro simili , risulta cioè :

$$A'_f = P^{-1} A_f P$$

con P matrice quadrata d'ordine n non degenere.

Dimostrazione. Denotiamo con

$$A_f = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \qquad \text{ed} \qquad A'_f = \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \dots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix}$$

le matrici che rappresentano f nelle basi B e B' ed indichiamo con P la matrice quadrata d'ordine n avente per colonne le coordinate di e'_1 , e'_2, .. , e'_n nella base B . Poiché il passaggio alle coordinate è un isomorfismo le colonne di P al pari dei vettori e'_1 , e'_2, .. , e'_n sono linearmente indipendenti e quindi la matrice P è non degenere e quindi invertibile. Si ha allora :

$$\text{(i)} \qquad (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \dots, \underline{\mathbf{e}}_n) \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} = (\underline{\mathbf{e'}}_1, \underline{\mathbf{e'}}_2, \dots, \underline{\mathbf{e'}}_n)$$

dalla quale per la linearità di f segue

$$\text{(ii)} \qquad (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \dots, f(\underline{\mathbf{e}}_n)) \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} = (f(\underline{\mathbf{e'}}_1), f(\underline{\mathbf{e'}}_2), \dots, f(\underline{\mathbf{e'}}_n))$$

D'altra parte per definizione è

$$\text{(iii)} \qquad (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \dots, \underline{\mathbf{e}}_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \dots, f(\underline{\mathbf{e}}_n))$$

$$\text{(iv)} \qquad (\underline{\mathbf{e'}}_1, \underline{\mathbf{e'}}_2, \dots, \underline{\mathbf{e'}}_n) \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \dots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix} = (f(\underline{\mathbf{e'}}_1), f(\underline{\mathbf{e'}}_2), \dots, f(\underline{\mathbf{e'}}_n))$$

Sostituendo nella (iv) le espressioni di (i) , (ii), (iii) si ha :

:

$$(\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n) \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \dots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix} =$$

$$= (\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}$$

Tale eguaglianza comporta , per l’indipendenza dei vettori $\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n$, che è :

$$P\, A'_f = A_f P$$

da cui segue :

$$A'_f = P^{-1} A_f\, P$$

e cioè l’asserto.

Proposizione 2.3 . *Sia* $f : V_n \to V_n$ *una funzione lineare di* V_n *in sé . Sia* $B = (\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n)$ *una base ordinata di* V_n . *La matrice* A_f *che rappresenta* f *nella base B ha la forma diagonale se e soltanto se i vettori* $\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n$ *della base B sono autovettori per* f.

Dimostrazione. La matrice A_f collega i vettori $\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n$ ai loro trasformati $f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \dots, f(\underline{\mathbf{e}}_n)$ risultando :

$$(\underline{\mathbf{e}}_1 , \underline{\mathbf{e}}_2 ,\dots,\underline{\mathbf{e}}_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \dots, f(\underline{\mathbf{e}}_n))$$

Pertanto se la matrice A_f è diagonale per esempio del tipo $\begin{pmatrix} \lambda_1 & 0 \ldots\ldots 0 \\ 0 & \lambda_2 \ldots\ldots 0 \\ \ldots\ldots & \\ 0 & 0 \ldots\ldots \lambda_n \end{pmatrix}$

allora da

$$(\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n) \begin{pmatrix} \lambda_1 & 0 \ldots\ldots 0 \\ 0 & \lambda_2 \ldots\ldots 0 \\ \ldots\ldots & \\ 0 & 0 \ldots\ldots \lambda_n \end{pmatrix} = (f(\underline{\mathbf{e}}_1), f(\underline{\mathbf{e}}_2), \ldots, f(\underline{\mathbf{e}}_n))$$

segue che è :

(j) $\qquad f(\underline{\mathbf{e}}_1) = \lambda_1 \underline{\mathbf{e}}_1, \quad f(\underline{\mathbf{e}}_2) = \lambda_2 \underline{\mathbf{e}}_2, \quad \ldots\ldots, \quad f(\underline{\mathbf{e}}_n) = \lambda_n \underline{\mathbf{e}}_n$

e ciò mostra che i vettori $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n$ sono autovettori per f .
Viceversa se i vettori $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n$ sono autovettori per f e valgono le relazioni (j) allora la matrice A_f avrà la forma

$$A_f = \begin{pmatrix} \lambda_1 & 0 \ldots\ldots 0 \\ 0 & \lambda_2 \ldots\ldots 0 \\ \ldots\ldots & \\ 0 & 0 \ldots\ldots \lambda_n \end{pmatrix}$$

e cioè è diagonale.

La funzione lineare f si dice che è ***diagonalizzabile*** se esiste una base B tale che la matrice A_f che rappresenta f nella base B ha la forma diagonale .

Abbiamo quindi provato nella proposizione precedente il seguente **fondamentale**

Teorema I . Una funzione lineare $f: V_n \rightarrow V_n$ è diagonalizzabile se e soltanto se lo spazio vettoriale V_n ha una base costituita da autovettori di f .

Questo teorema ci invita quindi a ricercare gli autovettori della funzione f e sperare che ne esistano n indipendenti.

Ma come si ricercano gli autovettori di f ? E' chiaro che la ricerca degli autovettori equivale alla ricerca degli autovalori di f . Quindi vediamo come si ricercano gli autovalori di f .

Sia quindi $f : V_n \to V_n$ una funzione lineare di V_n in sé .
$B = (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \ldots, \underline{\mathbf{e}}_n)$ sia una base ordinata di V_n ed $A_f = (a_{ij})$ sia la matrice che rappresenta f nella base B . Abbiamo già detto che la matrice A_f induce una applicazione lineare che indichiamo sempre con A_f di K^n in sé

$$A_f : K^n \to K^n$$

la quale, come già visto , ha la proprietà di trasformare le coordinate di un vettore **v** di V_n nelle coordinate del vettore f(**v**) .

Ora se

$$\mathbf{v} = x_1 \underline{\mathbf{e}}_1 + x_2 \underline{\mathbf{e}}_2 + \ldots + x_n \underline{\mathbf{e}}_n$$

è un vettore di V_n ed è **autovettore per f di autovalore** λ allora è :

$$f(\mathbf{v}) = \lambda \mathbf{v} = \lambda (x_1 \underline{\mathbf{e}}_1 + x_2 \underline{\mathbf{e}}_2 + \ldots + x_n \underline{\mathbf{e}}_n) = \lambda x_1 \underline{\mathbf{e}}_1 + \lambda x_2 \underline{\mathbf{e}}_2 + \ldots + \lambda x_n \underline{\mathbf{e}}_n$$

e quindi f(**v**) ha coordinate $(\lambda x_1, \lambda x_2, \ldots \lambda x_n)$. Poiché A_f trasforma le coordinate di **v** nelle coordinate di f(**v**) si ha

(*)
$$\begin{pmatrix} a_{11}\ a_{12} \ldots a_{1n} \\ a_{21}\ a_{22} \ldots a_{2n} \\ \ldots \\ a_{n1}\ a_{n2} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ .. \\ \lambda x_n \end{pmatrix}$$

e questa mostra che il vettore non nullo $(x_1, x_2, .., x_n)$ di K^n è autovettore per la funzione lineare A_f anch'esso con autovalore λ .
Viceversa se il vettore non nullo $(x_1, x_2, .., x_n)$ di K^n è autovettore per la funzione lineare A_f con autovalore λ allora vale la relazione (*) e quindi il vettore **v** non nullo

$$\mathbf{v} = x_1 \underline{\mathbf{e}}_1 + x_2 \underline{\mathbf{e}}_2 + \ldots + x_n \underline{\mathbf{e}}_n$$

determinato da (x_1 , x_2 , .., x_n) è trasformato nel vettore

$$f(\mathbf{v}) = \lambda x_1 \underline{\mathbf{e}}_1 + \lambda x_2 \underline{\mathbf{e}}_2 + \ldots + \lambda x_n \underline{\mathbf{e}}_n = \lambda \mathbf{v}$$

ed quindi **v** è un autovettore per f con autovalore λ .

Abbiamo così mostrato che l'isomorfismo

$$c: \quad V_n \to K^n$$

trasforma con la sua inversa gli autovettori di f negli autovettori di A_f ed inoltre che le due ***funzioni f ed A_f hanno gli stessi autovalori***. Se λ è un autovalore per f e quindi anche per A_f possiamo considerare l'autospazio $V\lambda$ di V_n corrispondente all'autovalore λ e l'autospazio $K\lambda$ di K^n corrispondente all'autovalore λ. Per quanto mostrato si ha :

$$c(V\lambda) = K\lambda \qquad \text{e} \qquad c^{-1}(K\lambda) = V\lambda$$

In particolare osserviamo che essendo c un isomorfismo i due sottospazi $V\lambda$ e $K\lambda$ hanno la stessa dimensione e quindi la molteplicità geometrica di λ non cambia sia che esso venga considerato autovalore di f e sia che esso venga pensato come autovalore di A_f.

Poiché le funzioni lineari f ed A_f hanno gli stessi autovalori , per trovare gli autovalori di f basterà determinare quelli di A_f . Come si può stabilire se uno scalare λ_o assegnato è autovalore per A_f ?

Lo scalare λ_o è autovalore per A_f se esiste un vettore non nullo $\underline{\mathbf{X}}$ di K^n che sia autovettore di A_f con autovalore λ_o, cioè che verifichi la relazione

$$\text{(a)} \qquad A_f \underline{\mathbf{X}} = \lambda_o \underline{\mathbf{X}}$$

Posto $\underline{\mathbf{X}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$ ed $A_f = \begin{pmatrix} a_{11} \; a_{12} \ldots a_{1n} \\ a_{21} \; a_{22} \ldots a_{2n} \\ \ldots\ldots \\ a_{n1} \; a_{n2} \ldots a_{nn} \end{pmatrix}$

la (a) esplicitamente è il seguente sistema di equazioni lineari nelle incognite $(x_1,x_2,\dots,x_n)$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots\dots + a_{1n}\,x_n = \lambda_o x_1 \\ a_{21}x_1 + a_{22}x_2 + \dots\dots + a_{2n}\,x_n = \lambda_o x_2 \\ \quad\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots\dots + a_{nn}\,x_n = \lambda_o x_n \end{cases}$$

che può essere anche scritto al seguente modo

$$\begin{cases} (a_{11} - \lambda_o)\,x_1 + a_{12}x_2 + \;..\; + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda_o)\,x_2 + \;..\; + a_{2n}x_n = 0 \\ \quad .. \\ a_{n1}x_1 + a_{n2}x_2 + \;..\; + (a_{nn} - \lambda_o)\,x_n = 0 \end{cases}$$

Ma tale sistema omogeneo ha una soluzione non banale se e soltanto se risulta

$$\det \begin{pmatrix} a_{11} - \lambda_o & a_{12} \dots\dots\dots\dots a_{1n} \\ a_{21} & a_{22} - \lambda_o \dots\dots a_{2n} \\ \dots\dots & \\ a_{n1} & a_{n2} \dots\dots\; a_{nn} - \lambda_o \end{pmatrix} = 0$$

Abbiamo quindi determinato la condizione cui deve soddisfare λ_o per essere autovalore di A_f:

λ_o è autovalore di A_f se sottraendo nella matrice A_f agli elementi a_{11}, a_{22}, ….,a_{nn} della sua diagonale lo scalare λ_o si ottiene una matrice con determinante zero .

Detta t una variabile possiamo allora considerare il seguente determinante

$$p(t) = \det(A_f - It) = \det \begin{pmatrix} a_{11} - t & a_{12} \dots\dots\dots a_{1n} \\ a_{21} & a_{22} - t \dots\dots a_{2n} \\ \dots\dots & \\ a_{n1} & a_{n2} \dots\dots\; a_{nn} - t \end{pmatrix}$$

che sviluppato fornisce un polinomio di grado n *e le cui radici* , ***ma solo quelle appartenenti a K***, *forniscono gli autovalori della funzione A_f e quindi di f* .

Il polinomio

$$p(t) = \det(A - I\,t) = (-1)^n\, t^n + \ldots\ldots$$

è chiamato il ***polinomio caratteristico*** della funzione f .

Tale polinomio è di nostro interesse avendo stabilito che sussiste la seguente

Proposizione 2.4. *Uno scalare λ_o di K è autovalore della funzione f se e solo se è $p(\lambda_o) = 0$.*

Osserviamo infine che se λ_o è un autovalore di A_f allora è

$$\det\begin{pmatrix} a_{11}-\lambda_o & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22}-\lambda_o \ldots\ldots a_{2n} \\ \ldots\ldots & \\ a_{n1} & a_{n2} \ldots\ldots a_{nn}-\lambda_o \end{pmatrix} = 0$$

e quindi la matrice $\begin{pmatrix} a_{11}-\lambda_o & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22}-\lambda_o \ldots\ldots a_{2n} \\ \ldots\ldots & \\ a_{n1} & a_{n2} \ldots\ldots a_{nn}-\lambda_o \end{pmatrix}$ ha rango $r < n$. Lo spazio delle soluzioni del sistema

$$\begin{cases} (a_{11}-\lambda_o)\,x_1 + a_{12}x_2 + \ \ldots + \ a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22}-\lambda_o)\,x_2 + \ \ldots + \ a_{2n}x_n = 0 \\ \quad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ \ldots + (a_{nn}-\lambda_o)\,x_n = 0 \end{cases}$$

ha allora , come sappiamo , dimensione n-r . Ma lo spazio delle soluzioni del sistema è l'autospazio K_{λ_o} corrispondente all'autovalore λ_o e quindi

se il rango della matrice $(A - I\,\lambda_o)$ è r l'autovalore λ_o ha molteplicità geometrica n-r .

Sia ora sempre λ_o un autovalore della funzione A_f. Poiché lo scalare λ_o è radice del polinomio caratteristico p(t) allora per un il teorema di Ruffini risulta che tale polinomio è divisibile per $(t - \lambda_o)$. Si ha cioè

$$p(t) = (t - \lambda_o)\ q(t)$$

con q(t) polinomio di grado n-1 .

Se λ_o non è radice di q(t) cioè si ha $q(\lambda_o) \neq 0$ allora λ_o è detta una radice ***semplice*** del polinomio p(t) .

Se invece risulta $q(\lambda_o) = 0$ allora , riapplicando il teorema di Ruffini , si ha

$$q(t) = (t - \lambda_o)\ g(t)$$

con g(t) polinomio di grado n-2 e quindi

$$p(t) = (t - \lambda_o)^2\ g(t)$$

Se risulta $g(\lambda_o) \neq 0$ allora λ_o è detta una radice ***doppia*** del polinomio p(t) .

Iterando il ragionamento si perviene alla seguente definizione .

La radice λ_o del polinomio p(t) si dice che ha ***molteplicità algebrica s*** se risulta

$$p(t) = (t - \lambda_o)^s\ \xi(t) \qquad e \qquad \xi(\lambda_o) \neq 0$$

Ad un autovalore λ_o della funzione lineare A abbiamo quindi assegnato due interi :

il primo $\chi(\lambda_o) = \dim \mathbf{V}_{\lambda_0}$ è la sua ***molteplicità geometrica*** (e questo intero ci informa su quanti autovettori indipendenti di autovalore λ_o possiamo trovare)

il secondo $\mu(\lambda_o)$ è la sua ***molteplicità algebrica*** (e questo intero ci informa su quante volte λ_o appare tra le radici del polinomio caratteristico)

Noi proveremo ora che la molteplicità geometrica dell'autovalore λ_o non supera mai quella algebrica . Prima di far ciò è essenziale provare la seguente

Proposizione 2.5 *Il polinomio caratteristico della funzione lineare f è indipendente dalla base scelta per rappresentare f .*

Dimostrazione. Siano quindi B e B' due basi ordinate di V_n e siano A_f ed A'_f le matrici che rappresentano f nelle due basi B e B' . Per quanto visto nella proposizione 2.2 le due matrici A_f ed A'_f sono simili tra loro e quindi esiste una matrice P non degenere per cui si abbia

$A'_f = P^{-1} A_f P$. Si ha allora

$$\det (A'_f - I t) = \det (P^{-1} A_f P - I t) = \det (P^{-1} A_f P - P^{-1} I t P) = \det [P^{-1} (A_f - I t) P] =$$
$$= \det P^{-1} \det (A_f - I t) \det P = \det (A_f - I t)$$

essendo $\det P^{-1} \det P = 1$.

Proviamo ora che :

Proposizione 2.6 *Sia λ_o un autovalore di f. La molteplicità geometrica di λ_o non supera la molteplicità algebrica di λ_o.*

Dimostrazione **.**Supponiamo che l'autovalore λ_o abbia molteplicità geometrica s . Nel-l' autospazio $\mathbf{V}_{\lambda_0}$ corrispondente all' autovalore λ_o si possono quindi scegliere s autovettori ($e_1, e_2, \ldots, e_s$) di autovalore λ_o e tra loro indipendenti. Per la proposizione 3.8 del capitolo I, si possono aggiungere ai vettori ($e_1, e_2, \ldots, e_s$) altri n-s vettori $w_{s+1}, w_{s+2}, \ldots, w_n$ in modo che i vettori ($e_1, e_2, \ldots, e_t, w_{s+1}, w_{s+2}, \ldots, w_n$) siano una base di V_n . Tenendo conto che è

$$f(e_1) = \lambda_o e_1 \ , \quad f(e_2) = \lambda_o e_2 , \ \ldots \ , \ f(e_s) = \lambda_o e_s$$

la matrice A_f che rappresenta f nella base B ha la seguente forma

$$A_f = \begin{pmatrix} \lambda_o & 0 \ldots\ldots & 0 \ldots\ldots & & \\ 0 & \lambda_o \ldots\ldots & 0 \ldots\ldots & & \\ .. & & & & \\ 0 & 0 \ldots\ldots & \lambda_o \ldots\ldots & & \\ 0 & 0 & 0 & a_{s+1,s+1} \ .. & a_{s+1\,n} \\ \ldots\ldots & & & & \\ 0 & 0 & 0 & a_{n,s+1} \ \ldots.. & a_{nn} \end{pmatrix}$$

Si ha quindi

$$p(t) = \det (A_f - I t) = (\lambda_o - t)^s g(t)$$

e questa mostra che la molteplicità algebrica di λ_o è almeno s . L' asserto è così provato.

Osserviamo esplicitamente che la proposizione 2.6 ora provata ha la seguente conseguenza. Sia λ_o un autovalore della funzione lineare f e supponiamo che sia radice del polinomio caratteristico p(t) di f con molteplicità algebrica 1 , cioè sia una *radice semplice* di p(t). Per la proposizione 2.6 ora provata risulta allora

$$\dim \mathbf{V}_{\lambda_0} \leq 1$$

ma poiché è anche $\dim \mathbf{V}_{\lambda_0} \geq 1$ si ha

$$\dim \mathbf{V}_{\lambda_0} = 1.$$

Riassumendo : ***ogni radice semplice del polinomio caratteristico è tale che la sua molteplicità geometrica eguaglia quella algebrica .***

Per i nostri scopi sono molto utili le proposizioni che seguono.

Proposizione 2.7. *Siano* $v_1, v_2, \dots, v_h$, *h autovettori per la funzione lineare f e siano* λ_1 , λ_2 , ... , λ_h *i loro corrispondenti autovalori . Se gli autovalori* λ_1 , λ_2 , ... , λ_h *sono distinti gli autovettori* $v_1, v_2, \dots, v_h$ *sono linearmente indipendenti .*

Dimostrazione. Ragioniamo per induzione sul numero h di autovettori . Se h=1 l'asserto è vero in quanto un autovettore , essendo un vettore non nullo , risulta da solo linearmente indipendente. Supponiamo quindi $h > 1$ e vero l'asserto per h-1 .
Supponiamo per assurdo che i vettori $v_1, v_2, \dots, v_h$ siano linearmente dipendenti . Poiché per l'ipotesi d'induzione $v_1, v_2, \dots, v_{h-1}$ sono indipendenti , il vettore v_h è combinazione lineare di $v_1, v_2, \dots, v_{h-1}$. Si ha quindi che esistono opportuni scalari $\alpha_1, \alpha_2, \dots\dots, \alpha_{h-1}$ per cui risulti

$$(\text{j.}) \qquad v_h = \alpha_1 v_1 + \alpha_2 v_2 + \dots\dots + \alpha_{h-1} v_{h-1}$$

Applicando ad ambo i membri della (j.) la funzione lineare f si ha , tenendo conto che i vettori $v_1, v_2, \dots, v_h$ sono autovettori di autovalori λ_1 , λ_2 , ... , λ_h

$$(\text{jj}) \qquad \lambda_h \, v_h = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots\dots + \alpha_{h-1} \lambda_{h-1} \, v_{h-1}$$

Moltiplicando ambo i membri della (j.) per lo scalare λ_h si ha :

$$(\text{jjj}) \qquad \lambda_h \, v_h = \alpha_1 \lambda_h v_1 + \alpha_2 \lambda_h v_2 + \dots\dots + \alpha_{h-1} \lambda_h v_{h-1}$$

Sottraendo tra loro la (jj) e la (jjj) si ha :

(jv) $\underline{0} = \alpha_1 (\lambda_h - \lambda_1)v_1 + \alpha_2 (\lambda_h - \lambda_2) v_2 + \ldots\ldots + \alpha_{h-1} (\lambda_h - \lambda_{h-1}) v_{h-1}$

Poiché , per l'ipotesi d'induzione i vettori $v_1, v_2, \ldots, v_{h-1}$ sono linearmente indipendenti , la (jv) sussiste solo con scalari tutti nulli, e pertanto si ha :

(**) $\alpha_1 (\lambda_h - \lambda_1) = 0$, $\alpha_2 (\lambda_h - \lambda_2) = 0$, ……. , $\alpha_{h-1} (\lambda_h - \lambda_{h-1}) = 0$

Essendo gli autovalori λ_1 , λ_2 , …. , λ_{h-1}, λ_h tra loro distinti risulta

$$\lambda_h - \lambda_1 \neq 0 \ , \ \lambda_h - \lambda_2 \neq 0 \ , \ldots\ldots., \ \lambda_h - \lambda_{h-1} \neq 0$$

e qundi dalla (**) segue che è $\alpha_1 = 0$, $\alpha_2 = 0$, ….., $\alpha_{h-1} = 0$. Ma allora dalla (j) segue che è $v_h = \underline{0}$ e ciò è assurdo , essendo v_h un autovettore.

Proposizione 2.8. Se *λ_1 , λ_2 , …. , λ_h sono h autovalori distinti per la funzione lineare f allora gli autospazi V_{λ_1} , V_{λ_2} , ….., V_{λ_h} ad essi corrispondenti sono tali che ciascuno di essi interseca nel solo vettore nullo lo spazio generato dai rimanenti.*

Dimostrazione. Per semplicità proviamo che è

$$V_{\lambda_1} \cap (V_{\lambda_2} + \ldots..+ V_{\lambda_h}) = \{ \underline{0} \}$$

Sia v un vettore di $V_{\lambda_1} \cap (V_{\lambda_2} + \ldots..+ V_{\lambda_h})$ e supponiamo per assurdo che v sia non nullo e quindi un autovettore . Poiché v appartiene al sottospazio $V_{\lambda_2} + \ldots..+ V_{\lambda_h}$ allora è

(i) $v = v_2 + \ldots. + v_h$

con $v_2 \in V_{\lambda_2}$, … , $v_h \in V_{\lambda_h}$

Applicando ad ambo i membri della (i) la funzione lineare f si ottiene , ricordando che $v \in V_{\lambda_1}$ e che $v_2 \in V_{\lambda_2}$, … , $v_h \in V_{\lambda_h}$

(ii) $\lambda_1 v = \lambda_2 v_2 + \ldots\ldots + \lambda_h v_h$

Moltiplicando ambo i membri della (i) per lo scalare λ_1 si ha

(iii) $$\lambda_1 v = \lambda_1 v_2 + \ldots\ldots + \lambda_1 v_h$$

Sottraendo tra loro la (ii) e la (iii) si ha :

(iv) $$\underline{0} = (\lambda_2 - \lambda_1)v_2 + \ldots\ldots + (\lambda_h - \lambda_1)\ v_h$$

Poiché gli autovalori λ_2 , …. , λ_h sono distinti , per la proposizione (2.7), gli autovettori v_2 ,…. ,v_h sono linearmente indipendenti e quindi la (iv) può sussistere solo con scalari tutti nulli. Si ha quindi $\lambda_1 = \lambda_2 = \ldots = \lambda_h$ e ciò contraddice l'ipotesi che gli autovalori siano distinti tra loro.

Le due proposizioni ora provate hanno importanti conseguenze come ora faremo vedere. Ricordiamo che il problema che abbiamo posto e che stiamo studiando è il seguente :

Una funzione lineare è sempre diagonalizzabile ?

Una risposta al problema è data dal Teorema I il quale afferma che f è diagonalizzabile se e solo se esiste una base B di V_n costituita da autovettori per la funzione lineare f .

Ma tale base esiste sempre ? Vediamo.

Un caso favorevole è espresso dalla seguente :

Proposizione 2.9 *Sia $f : V_n \rightarrow V_n$ una funzione lineare . Se il polinomio caratteristico p(t) di f ha n radici λ_1 , λ_2 , …. , λ_n in K distinte tra loro allora f è diagonalizzabile.*

Dimostrazione. Siano $v_1,v_2,\ldots,v_n$ n autovettori corrispondenti agli autovalori λ_1 , λ_2 , …. , λ_n . Per la proposizione 2.7 gli autovettori $v_1,v_2,\ldots,v_n$ sono linearmente indipendenti e quindi costituiscono una base per lo spazio V_n . L'asserto segue quindi dal Teorema I .

Cosa si può dire se invece il polinomio p(t) ha radici multiple sempre nel campo K ?

La proposizione che segue risponde a tale quesito ed inoltre fornisce una risposta completa al problema della diagonalizzabilità della funzione lineare f .

Proposizione 2.10 *Sia* $f : V_n \rightarrow V_n$ *una funzione lineare . La funzione* f *è diagonalizzabile se e solo se il polinomio caratteristico* $p(t)$ *di* f *ha* n *radici* $\lambda_1, \lambda_2, \dots, \lambda_n$ *in K ciascuna con molteplicità geometrica eguale a quella algebrica.*

Dimostrazione. Supponiamo che f sia diagonalizzabile. Per il Teorema I esiste allora una base $B = \{ e_1,e_2,\dots,e_n \}$ costituita da autovettori per f . Raggruppiamo tali autovettori mettendo insieme quelli che hanno lo stesso autovalore. Siano $\lambda_1, \lambda_2, \dots, \lambda_t$ gli autovalori distinti forniti dagli autovettori $e_1,e_2,\dots,e_n$. Sia

T_1 il sottoinsieme di B costituito da tutti gli autovettori di autovalore λ_1 .

T_2 il sottoinsieme di B costituito da tutti gli autovettori di autovalore λ_2 .

…….

T_t il sottoinsieme di B costituito da tutti gli autovettori di autovalore λ_t .

Siano $n_1, n_2, \dots, n_t$ le cardinalità di $T_1, T_2, \dots, T_t$ e siano $s_1, s_2, \dots, s_t$ le molteplicità algebriche degli autovalori $\lambda_1, \lambda_2, \dots, \lambda_t$.

Poiché ogni autospazio V_{λ_i} $i = 1,2,\dots,t$ contiene i vettori di T_i , $i = 1,2,\dots,t$ si ha :

$$n_1 \leq \dim V_{\lambda_1} \leq s_1$$

$$n_2 \leq \dim V_{\lambda_2} \leq s_2$$

……………

$$n_t \leq \dim V_{\lambda_t} \leq s_t$$

Si ha allora

$$n = n_1 + n_2 + \dots + n_t \leq \dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_t} \leq s_1 + s_2 + \dots + s_t \leq n$$

Da tale diseguaglianza segue :

$$s_1 + s_2 + \dots + s_t = n$$

e questa mostra che le radici $\lambda_1, \lambda_2, \dots, \lambda_t$ contate con la loro molteplicità sono in numero di n. Inoltre da

$n = n_1 + n_2 + \ldots + n_t \le \dim V_{\lambda 1} + \dim V_{\lambda 2} + \ldots + \dim V_{\lambda t} \le s_1 + s_2 + \ldots + s_t = n$

segue che per nessun $i = 1,2,\ldots,t$ può essere $\dim V_{\lambda i} < s_i$ altrimenti è $n < n$.

Quindi è :

$$\dim V_{\lambda 1} = s_1, \quad \dim V_{\lambda 2} = s_2 , \ldots, \quad \dim V_{\lambda t} = s_t$$

e ciò mostra che le radici $\lambda_1 , \lambda_2 , \ldots , \lambda_t$ hanno ciascuna molteplicità geometrica eguale a quella algebrica.

Viceversa supponiamo ora che il polinomio caratteristico di f abbia n radici $\zeta_1, \zeta_2,\ldots, \zeta_n$ nel campo K e che ognuna di esse abbia molteplicità algebrica eguale a quella geometrica . Raggruppiamo le radici mettendo insieme quelle eguali tra loro. Siano $\lambda_1 , \lambda_2 , \ldots , \lambda_t$ le radici distinte che figurano tra le n radici $\zeta_1, \zeta_2,\ldots, \zeta_n$.

Supponiamo che

λ_1 appaia s_1 volte tra le radici $\zeta_1, \zeta_2,\ldots, \zeta_n$.

λ_2 appaia s_2 volte tra le radici $\zeta_1, \zeta_2,\ldots, \zeta_n$.

….

λ_t appaia s_t volte tra le radici $\zeta_1, \zeta_2,\ldots, \zeta_n$.

Quindi

λ_1 ha molteplicità algebrica s_1 .

λ_2 ha molteplicità algebrica s_2 .

…..

λ_t ha molteplicità algebrica s_t .

ed inoltre è ovviamente

$$s_1 + s_2 + \ldots + s_t = n.$$

Poiché per ipotesi la molteplicità algebrica è eguale a quella geometrica si ha :

$$\dim V_{\lambda 1} = s_1, \quad \dim V_{\lambda 2} = s_2 , \ldots, \quad \dim V_{\lambda t} = s_t$$

Siano ora

B_1 una base di $V_{\lambda 1}$, B_2 una base di V_{λ_2} , ……, B_t una base di $V_{\lambda t}$.

Essendo gli autovalori λ_1 , λ_2 , …. , λ_t distinti gli autospazi V_{λ_1} , V_{λ_2} , ….., $V_{\lambda t}$ ad essi corrispondenti sono tali che ciascuno di essi interseca nel solo vettore nullo lo spazio generato dai rimanenti ed allora per la proposizione 5.1 del capitolo I , unendo le basi B_1 , B_2,….. , B_t si ottiene un insieme B costituito da $s_1 + s_2 +….. + s_t = n$ autovettori indipendenti di V_n, cioè una base di autovettori.

Per il Teorema I allora f è diagonalizzabile e l'asserto è completamente provato.

Concludiamo con qualche esempio.

Esercizio 1.

Sia assegnata la seguente matrice reale

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

Si consideri la funzione lineare $A : R^3 \rightarrow R^3$ indotta dalla matrice A .

Si determini il nucleo di A e si stabilisca se essa ha una base di autovettori cioè se è diagonalizzabile .

Svolgimento.

I vettori $\underline{X} = (x,y,z)$ del nucleo di A sono quelli per cui risulti $A\underline{X} = \underline{0}$.

Tali vettori sono quindi le soluzioni del seguente sistema omogeneo :

$$\begin{cases} x + z = 0 \\ x + y + z = 0 \\ x + z = 0 \end{cases}$$

Tale sistema è ovviamente equivalente a

$$\begin{cases} x + z = 0 \\ x + y + z = 0 \end{cases}$$

le cui soluzioni sono il sottospazio di dimensione 1 generato da (-1 , 0 , 1). Il nucleo di A non è ridotto al solo vettore nullo e quindi (-1 , 0 , 1) è un autovettore di autovalore 0.
La matrice A rappresenta la funzione lineare A nella base canonica di R^3 e pertanto
il polinomio caratteristico di A è :

$$\det(A - I\,t) = \begin{pmatrix} 1-t & 0 & 1 \\ 1 & 1-t & 1 \\ 1 & 0 & 1-t \end{pmatrix} = (1-t)\,[\,(1-t)^2 - 1] = t\,(1-t)\,(t-2)$$

Gli autovalori di A sono quindi i numeri 0 , 1, 2 . Tali valori sono distinti e quindi A è diagonalizzabile per la proposizione 2.9.

Esercizio 2. Si consideri la seguente matrice reale

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

e sia $A : R^3 \to R^3$ la funzione lineare indotta da A .

$$A(x, y, z) = (x+2z,\ y+2z,\ 2z)$$

Si stabilisca se tale funzione ha una base di autovettori e cioè se è diagonalizzabile .

Svolgimento.
La matrice A rappresenta la funzione lineare A nella base canonica di R^3. Pertanto il polinomio caratteristico di A è dato da :

$$\det(A - I\,t) = \begin{pmatrix} 1-t & 0 & 2 \\ 0 & 1-t & 2 \\ 0 & 0 & 2-t \end{pmatrix} \ (1-t)^2 \ (2-t)$$

Le radici di tale polinomio sono i numeri 1 e 2 .

Il numero 2 è una radice semplice mentre 1 è una radice doppia.

Gli autovettori di autovalore 1 si ottengono attraverso le soluzioni de seguente sistema omogeneo.

$$0x + 0y + z = 0$$

Esse sono quindi un sottospazio S di dimensione due. Due soluzioni indipendenti di tale sistema sono (1,0,0) e (0,1,0) e quindi è S = [(1,0,0) , (0,1,0)] .

L'autospazio corrispondente all'autovalore 1 è quindi S = [(1,0,0) , (0,1,0)].

Avendo controllato che la molteplicità geometrica di 1 è eguale a quella algebrica la funzione lineare A è diagonalizzabile per la proposizione 2.10.

Esercizio 3.

Una funzione lineare $f : R^3 \rightarrow R^3$ assume i seguenti valori

$$f(1,0,0) = (2,0,0)$$
$$f(0,1,0) = (0,0,0)$$
$$f(0,1,1) = (2,2,0)$$

Si determini f e si dica se è diagonalizzabile.

Svolgimento. Poiché i vettori (1,0,0) , (0,1,0) , (0,1,1) sono indipendenti in quanto è

$$\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \neq 0$$

essi costituiscono una base di R^3 .La conoscenza dei valori che f assume su tali vettori determina quindi f . Vediamo.

Poiché $\{$ (1,0,0) , (0,1,0) , (0,1,1) $\}$ è una base un qualunque vettore (x, y, z) di R^3 risulta una loro combinazione lineare si ha cioè per opportuni scalari α , β , γ

(i) $$(x, y, z) = \alpha(1,0,0) + \beta\ (0,1,0) + \gamma\ (0,1,1) = (\alpha\ ,\ \beta + \gamma\ ,\ \gamma)$$

da (i) segue

$\alpha = x$

$\beta + \gamma = y$

$\gamma = z$

e cioè per realizzare l'eguaglianza (i) , occorre scegliere :

$$\alpha = x , \quad \beta = y - z , \quad \gamma = z .$$

Si ha quindi

$$(x, y, z) = x(1,0,0) + y\text{-}z\ (0,1,0) + z\ (0,1,1)$$

Essendo f lineare si ha allora

$$f(x, y, z) = x\ f(1,0,0) + (y\text{-}z)\ f(0,1,0) + z\ f(0,1,1)$$

ma è per ipotesi $f(1,0,0) = (2,0,0)$, $f(0,1,0) = (0,0,0)$, $f(0,1,1) = (2,2,0)$

e quindi è :

$$f(x, y, z) = x\ (2,0,0) + (y\text{-}z)\ (0,0,0) + z\ (2,2,0) = (2x + 2z , 2z , 0)$$

La funzione f è così determinata.

Determiniamo ora la matrice A_f che rappresenta f nella base assegnata . Si ha

$f(1,0,0) = (2,0,0) = 2(1,0,0) + 0\ (0,1,0) + 0\ (0,1,1)$

$f(0,1,0) = (0,0,0) = 0(1,0,0) + 0\ (0,1,0) + 0\ (0,1,1)$

$f(0,1,1) = (2,2,0) = 2(1,0,0) + 2\ (0,1,0) + 0\ (0,1,1)$

e quindi è :

$$A_f = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Il polinomio caratteristico di f è quindi

$$p(t) = \det \begin{pmatrix} 2-t & 0 & 2 \\ 0 & -t & 2 \\ 0 & 0 & -t \end{pmatrix} = t^2\,(2-t)$$

Gli autovalori di f sono quindi **0** (radice **doppia**) e **2** radice **semplice**.

L'autospazio corrispondente all'autovalore 0 si ottiene in corrispondenza alle soluzioni del sistema omogeneo

$$\begin{cases} 2x + 2z = 0 \\ 2z = 0 \end{cases}$$

che ha rango due essendo le due equazioni indipendenti .Le soluzioni di tale sistema costituiscono quindi un sottospazio di dimensione 1, con base (0 , -4 , 0) , base ottenuta in corrispondenza ai minori presi a segni alterni della matrice dei coefficienti del sistema

$$\begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

Pertanto la molteplicità geometrica di **0** è uno e quindi non coincide con la sua molteplicità algebrica che è due.

La funzione f non è quindi diagonalizzabile.

Capitolo VI

Funzioni lineari simmetriche

1.- Il gruppo lineare ed il gruppo ortogonale.

In questo numero useremo le seguenti notazioni.

Se $\underline{\mathbf{x}}=(x_1, x_2,\dots,x_n)$ ed $\underline{\mathbf{y}}=(y_1, y_2,\dots,y_n)$ sono due n-ple ordinate di numeri reali col simbolo

$$< \underline{\mathbf{x}} , \underline{\mathbf{y}} > = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

rappresenteremo il loro ***prodotto scalare euclideo***.

Inoltre se $A = (a_{ij})$ è una matrice quadrata d'ordine n ad elementi reali indicheremo coi simboli $\underline{a}_1, \underline{a}_2,\dots,\underline{a}_n$ le sue righe e con $\underline{a}^1$, $\underline{a}^2$,…, $\underline{a}^n$ le sue colonne.

Mantenendo queste notazioni ricordiamo che se $A = (a_{ij})$ e $B = (b_{ij})$ sono due matrici quadrate d'ordine n il loro prodotto AB (*eseguito righe per colonne*) è la matrice quadrata $C = (c_{ij})$ in cui è :

$$c_{ij} = < \underline{a}_i , \underline{b}^j >$$

Sia n un intero maggiore o eguale a due e sia GL(n,R) l'insieme di tutte le matrici quadrate d'ordine n ad elementi reali e che siano non degeneri cioè quelle dotate di inversa. Tali matrici come sappiamo sono tutte e sole quelle che hanno il determinante diverso da zero.

Il prodotto di due matrici A e B non degeneri è non degenere in quanto è $\det(AB)= \det A \det B \neq 0$.

Inoltre il prodotto tra matrici è associativo , la matrice identica **I** (d'ordine n) è non degenere ed ogni matrice non degenere ha l'inversa. Per queste ragioni l'insieme GL(n,R) è un gruppo detto ***gruppo lineare.***

Nel seguito denoteremo per ogni matrice A non degenere con A_t la sua trasposta (*cioè la matrice avente per colonne le righe di A e per righe le colonne di A*) e con A^{-1} la sua inversa.

Per ciò che segue sono utili le seguenti relazioni di facile dimostrazione.

Siano A e B due matrici non degeneri . Risulta :

(1.1) $$(AB)^{-1} = B^{-1} A^{-1} \qquad\qquad (AB)_t = B_t A_t$$

Abbiamo già visto che una matrice A col determinante diverso da zero è invertibile e però ,quando n è grande, la costruzione della sua inversa è faticosa in quanto coinvolge il calcolo dei complementi algebrici degli elementi di A.

Per questa ragione sono di un certo interesse le matrici non degeneri per le quali sia facile il calcolo della loro inversa. Fanno parte di questa famiglia le matrici ***ortogonali*** di cui ora ci occuperemo.

Una matrice A non degenere è detta ***ortogonale*** se risulta :

(1.2) $$A^{-1} = A_t$$

Per riconoscere se una matrice è ortogonale sono utili le seguenti proposizioni.

Proposizione 1.1 *Se A è una matrice ortogonale allora si ha* $\det A = \pm 1$.

Dimostrazione. Si ha

$$\det I = \det(A\,A^{-1}) = \det A \det A^{-1} = \det A \det A_t = (\det A)^2 = 1$$

da cui segue l'asserto.

La proposizione ora provata fornisce una condizione solo necessaria per la ortogonalità di una assegnata matrice . In forza di tale proposizione una matrice che abbia determinante diverso da -1 e 1 non è ortogonale.

La condizione non è però sufficiente come mostra il seguente esempio. Si consideri la matrice

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

il cui determinante è eguale ad 1 . La matrice A non è ortogonale in quanto risulta $A\,A_t \neq I$.

Una condizione necessaria e sufficiente è invece fornita dalla seguente:

Proposizione 1.2 *Una matrice A è ortogonale se e solo se risulta* $AA_t = I$.

Dimostrazione . Se A è ortogonale è $A^{-1} = A_t$ e quindi risulta $AA_t = I$.

Viceversa se $AA_t = I$ allora è $AA_t = A\,A^{-1}$ e quindi, per la legge di cancellazione valida in un gruppo , si ha $A_t = A^{-1}$.

La proposizione mostra quindi che per controllare se A_t coincide con l'inversa di A non è necessario controllare che essa è permutabile con A.

Con analoga dimostrazione si prova la seguente :

Proposizione 1.3 Una matrice A è ortogonale se e solo se risulta $A_t A = I$.

La Proposizione 1.2 ora provata è equivalente alla seguente :

Proposizione 1.4 . *Una matrice A è ortogonale se e solo se le sue righe sono una base ortonormale di R^n.*

Dimostrazione . Se A è ortogonale risulta $A A_t = I$ e quindi ricordando che le colonne di A_t sono le righe di A quando si esegue il prodotto AA_t righe per colonne, di fatto si moltiplicano le righe di A tra di loro e quindi è, essendo
$A A_t = I$:

per ogni i = 1,2,…,n $\quad < \underline{a}_i , \underline{a}_i > = 1$
e per $i \neq j$ $\quad < \underline{a}_i , \underline{a}_j > = 0$

e queste mostrano che le righe di A sono una base ortonormale di R^n .Viceversa se le righe di A sono una base ortonormale di R^n cioè si ha :

per ogni i = 1,2,…,n $\quad < \underline{a}_i , \underline{a}_i > = 1$
e per $i \neq j$ $\quad < \underline{a}_i , \underline{a}_j > = 0$

allora risulta $A A_t = I$ e quindi per la proposizione 1.2 la matrice A è ortogonale.

La proposizione ora provata mostra quindi quale sia il modo di costruire una matrice ortogonale .

Con argomentazioni del tutto analoghe quando si tenga conto della Proposizione 1.3 e della circostanza che righe di A_t sono le colonne di A si perviene
alla seguente :

Proposizione 1.5 . *Una matrice A è ortogonale se e solo se le sue colonne sono una base ortonormale di R^n.*

Abbiamo già visto che una matrice quadrata d'ordine n ad elementi reali ,

$$A = \begin{pmatrix} a_{11} \ \ a_{12} \ldots\ldots a_{1n} \\ a_{21} \ a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots \\ a_{n1} \ a_{n2} \ldots\ldots a_{nn} \end{pmatrix}$$

induce su R^n , spazio vettoriale numerico reale di dimensione n , un'applicazione lineare che denotiamo sempre con A ,

$$A : \ R^n \ \rightarrow \ R^n$$

quando si faccia corrispondere al vettore

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

il vettore

$$\underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} \ \ a_{12} \ldots\ldots a_{1n} \\ a_{21} \ a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots \\ a_{n1} \ a_{n2} \ldots\ldots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

ottenuto moltiplicando, righe per colonne , la matrice A per il vettore $\underline{\mathbf{x}}$ .Tale applicazione è un isomorfismo se la matrice A è non degenere.
Nel seguito denoteremo con

$$\underline{\mathbf{u_1}} = \begin{pmatrix} 1 \\ 0 \\ . \\ 0 \end{pmatrix}, \ \underline{\mathbf{u_2}} = \begin{pmatrix} 0 \\ 1 \\ . \\ 0 \end{pmatrix}, \ldots\ldots\ldots, \ \underline{\mathbf{u_n}} = \begin{pmatrix} 0 \\ 0 \\ . \\ 1 \end{pmatrix}$$

i vettori della base canonica di R^n . L' isomorfismo A trasforma i vettori $\underline{\mathbf{u_1}}$, $\underline{\mathbf{u_2}}$,…, $\underline{\mathbf{u_n}}$ della base canonica di R^n nei vettori colonna $\underline{a}^1$, $\underline{a}^2$,…, $\underline{a}^n$ della matrice A.

I teoremi che seguono caratterizzano le matrici ortogonali attraverso le proprietà di tale isomorfismo.

Proviamo che :

Proposizione 1.6 *Una matrice non degenere A è ortogonale se e solo se l'isomorfismo A conserva il prodotto scalare cioè risulta , per ogni coppia $\underline{x}$, $\underline{y}$ di R^n*

(1.3) $$< \underline{x} , \underline{y} > = < A\underline{x} , A\underline{y} > \quad .$$

Dimostrazione. Se A è ortogonale risulta $A^{-1} = A_t$ e quindi si ha :

$$< A\underline{x} , A\underline{y} > \; = \; (A\underline{x})_t A\underline{y} \; = \; \underline{x}_t A_t A\underline{y} \; = \; \underline{x}_t I \underline{y} = \; \underline{x}_t \; \underline{y} \; = \; < \underline{x} , \underline{y} >$$

e cioè la (1.3).

Viceversa se vale la (1.3) si ha :

$$< \underline{u_i} , \underline{u_j} > = < A \underline{u_i} , A \underline{u_j} > \; = < \underline{a}^i , \underline{a}^j >$$

e quindi è :

$$< \underline{a}^i , \underline{a}^j > = 1 \text{ per } i = j \quad \text{e} \quad < \underline{a}^i , \underline{a}^j > = 0 \text{ per } i \neq j$$

e ciò prova che A è ortogonale per la proposizione 1.3.

Proviamo ora la seguente

Proposizione 1.7 *Una matrice non degenere A è ortogonale se e solo se l'isomorfismo A conserva la lunghezza dei vettori, cioè per ogni vettore $\underline{x}$ di R^n risulta :*

(1.4) $$|\underline{x}| = |A\underline{x}|.$$

Dimostrazione. Se la matrice A è ortogonale allora per la proposizione precedente l'isomorfismo A conserva il prodotto scalare e quindi la lunghezza dei vettori. Viceversa supponiamo valga la (1.4). Si ha , per i e j diversi tra loro :

$$|u_i + u_j| = |A(u_i + u_j)| = |A u_i + A u_j| = |a^i + a^j|$$

da cui segue :

$$|u_i + u_j|^2 = |a^i + a^j|^2$$

Si ha allora :

$$|u_i + u_j|^2 = < u_i + u_j , u_i + u_j > = < u_i, u_i > + < u_j , u_j > + 2 < u_i, u_j > = 2$$

essendo $< u_i, u_i > = < u_j , u_j > = 1$ e $< u_i, u_j > = 0$.

$$|a^i + a^j|^2 = < a^i + a^j, a^i + a^j > = < a^i, a^i > + < a^j , a^j > + 2 < a^i, a^j > = 2.$$

Poiché per ipotesi è $< a^i, a^i > = < u_i, u_i > = 1$ e $< a^j , a^j > = < u_j , u_j > = 1$ si ha allora $< a^i, a^j > = 0$.

Da $< a^i, a^i > = 1$ ed $< a^i, a^j > = 0$ segue che A è ortogonale per la proposizione 1.5.

Concludiamo provando la seguente:

Proposizione 1.8 *Una matrice non degenere A è ortogonale se e solo se l'isomorfismo A trasforma una base ortonormale di R^n in una base ortonormale di R^n .*

Dimostrazione . Se A è ortogonale allora l'isomorfismo A per quanto provato conserva le lunghezze e l'ortogonalità ed inoltre essendo un isomorfismo trasforma una base in una base . Così una base ortonormale di R^n è trasformata in una base ortonormale di R^n. Viceversa se l'isomorfismo A trasforma una base ortonormale di R^n in una base ortonormale di R^n allora A è ortogonale, per la proposizione 1.5, in quanto le sue colonne, essendo i trasformati della base canonica di R^n , sono anch'esse una base ortonormale di R^n .

2. *Applicazioni lineari simmetriche.*

In questo numero V_n denota uno spazio vettoriale reale di dimensione finita n e dotato di un prodotto scalare

$$s : V_n \times V_n \rightarrow R$$

definito positivo . Abbiamo già visto che se

$$f : V_n \rightarrow V_n$$

è un'applicazione lineare di V_n in sé allora quando si scelga una base ordinata $B = (e_1 , e_2 ,\dots, e_n)$ di V_n si può considerare la matrice A_f che *rappresenta f* nella base scelta. Tale matrice , quadrata d'ordine n , ha per colonne le coordinate nella base B dei vettori

$$f(e_1) , f(e_2) ,\dots, f(e_n)$$

ed è utile per il calcolo della funzione f in quanto, come abbiamo già visto , l'endomorfismo , che indichiamo con A , che A_f induce su R^n

$$A : R^n \rightarrow R^n$$

fa corrispondere al vettore

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

delle coordinate di un vettore v di V_n il vettore

$$\underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} \ a_{12} \ .... a_{1n} \\ a_{21} \ a_{22} \ .....a_{2n} \\ ..... \\ a_{n1} \ a_{n2} \ .....a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$$

delle coordinate di f(v) rispetto alla base B.

Ovviamente tale calcolo diviene semplice , anche con n grande, quando tale matrice ha la forma diagonale. Tale matrice ha però la forma diagonale se e solo se la base B scelta è una base di autovettori per f . Abbiamo altresì visto che non sempre esiste una base di autovettori per f . Quando tale base ”speciale” esiste la funzione lineare viene detta *diagonalizzabile.*

Scelta la base ordinata B = (e_1 , e_2 ,…, e_n) di V_n si può considerare oltre alla matrice A_f che *rappresenta f* nella base B, anche la matrice quadrata

$$S= (s_{ij}) \qquad \text{con} \qquad s_{ij} = s(e_i , e_j)$$

d'ordine n che *rappresenta il prodotto scalare s nella base B*. Così come la matrice A_f può essere usata per il calcolo di f , anche la matrice S può essere usata per il calcolo di s. Vediamo. Siano v e w due vettori di V_n .

Sia $\underline{\mathbf{x}}$= (x_1 , x_2 ,…, x_n) il vettore delle coordinate di v nella base B e sia $\underline{\mathbf{y}}$= (y_1 , y_2 ,…, y_n) il vettore delle coordinate di w nella base B. Si ha :

$$v = x_1 e_1+ x_2 e_2 +\ldots+x_n e_n \quad , \quad w = y_1 e_1+ y_2 e_2 +\ldots+y_n e_n$$

Tenendo conto della bilinearità del prodotto scalare s si ha :

$$s(v , w) = s (x_1 e_1+ x_2 e_2 +\ldots+x_n e_n \quad , \quad y_1 e_1+ y_2 e_2 +\ldots+y_n e_n) = \sum_{i,j} s_{ij}\, x_i\, y_j$$

Quando la base scelta è ortogonale allora è

$$s_{ij} = s(e_i , e_j) = 0 \ \text{ per } \ i \neq j$$

e così la matrice S ha la forma diagonale ed allora il calcolo del numero s(v , w) è più semplice risultando

$$s(v , w) = s_{11}\, x_1\, y_1 + s_{22}\, x_2\, y_2 + \ldots. + s_{nn}\, x_n\, y_n$$

Se la base B è ortonormale allora è

$$\text{per ogni } \ i= 1,2,..,n \qquad s_{ii} = s(e_i , e_i) = 1$$

$$\text{e per } i \neq j \qquad s_{ij} = s(e_i , e_j) = 0$$

e così risulta S = I . In tal caso, come già visto, è :

$$s(v , w) = x_1\, y_1 + x_2\, y_2 + \ldots. + x_n\, y_n = < \underline{x} , \underline{y} >.$$

Da quanto detto segue che per rendere contemporaneamente diagonali le due matrici A_f ed S che rappresentano nella base B scelta , rispettivamente f ed s, la base B deve essere una base di autovettori per f a due a due ortogonali.Ci sono alcune funzioni lineari speciali che sono dotate di una base di autovettori a due a due ortogonali, come noi vogliamo, e che sono quindi diagonalizzabili. Tali funzioni lineari sono le *funzioni lineari simmetriche* di cui ora ci occuperemo.

Sia quindi V_n uno spazio vettoriale reale di dimensione finita n , dotato di un prodotto scalare

$$s : V_n \text{ x } V_n \ \rightarrow \ R$$

definito positivo e sia

$$f : V_n \ \rightarrow \ V_n$$

un'applicazione lineare di V_n in sé.

L'applicazione lineare f è detta ***simmetrica*** se per ogni coppia di vettori v e w di V_n risulta :

$$\text{(2.1)} \qquad s (v , f(w)) = s (f (v) , w).$$

Ci sono funzioni lineari simmetriche ? La risposta è affermativa.

Infatti sia A una matrice quadrata d'ordine n reale e simmetrica.
L' applicazione lineare

$$A : R^n \rightarrow R^n$$

che la matrice A induce su R^n è simmetrica . Infatti essendo $A = A_t$ risulta :

$$< \underline{x} , A\underline{y} > = \underline{\mathbf{x}}_t A \underline{\mathbf{y}} = \underline{\mathbf{x}}_t A_t \underline{\mathbf{y}} = < A\underline{x} , \underline{y} >$$

Sulla base della definizione è difficile controllare se una funzione lineare assegnata risulti simmetrica. Risulta pertanto utile la seguente :

Proposizione 2.1. *La funzione lineare f è simmetrica se e solo se la matrice che rappresenta f in una base ortonormale di V_n risulta simmetrica.*

Dimostrazione. Supponiamo che la funzione lineare f sia simmetrica e sia $B = (e_1 , e_2 ,\dots, e_n)$ una base ortonormale di V_n. Indichiamo con A la matrice che rappresenta f nella base B. Bisogna controllare che la matrice A risulta simmetrica.
Poiché la matrice A ha per colonne le coordinate di $f(e_1) , f(e_2) , \dots, f(e_n)$ si ha fissati due indici i e j distinti tra loro :

$$f(e_i) = a_{1i}\, e_1 + a_{2i}\, e_2 + \dots + a_{ni}\, e_n$$

$$f(e_j) = a_{1j}\, e_1 + a_{2j}\, e_2 + \dots + a_{nj}\, e_n$$

Si ha allora , essendo la base B ortonormale , :

$$s\,(f(e_i) , e_j) = a_{ji}$$

$$s\,(f(e_i) , e_j) = a_{ij}$$

e poiché f è simmetrica risulta $s\,(f(e_i) , e_j) = s\,(f(e_i) , e_j)$ e cioè $a_{ji} = a_{ij}$.

Viceversa supponiamo che la matrice A che rappresenta f nella base ortonormale B risulti simmetrica e proviamo che f è simmetrica.

Siano quindi v e w due vettori di V_n e siano

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} \qquad \text{ed} \qquad \underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix}$$

le loro coordinate nella base B. Si ha :

$$v = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \quad , \qquad w = y_1 e_1 + y_2 e_2 + \dots + y_n e_n$$

Ora ,come abbiamo già ricordato, la matrice A che rappresenta f nella base B , trasforma le coordinate di v nelle coordinate di f(v) e le coordinate di w nelle coordinate di f(w). Inoltre poiché la base è ortonormale il prodotto scalare di due vettori coincide col prodotto scalare euclideo delle loro coordinate. Si ha allora :

$$s(f(v), w) = \langle A\underline{x}, \underline{y} \rangle = \underline{\mathbf{x}}_t A_t \underline{\mathbf{y}}$$
$$s(v, f(w)) = \langle \underline{x}, A\underline{y} \rangle = \underline{\mathbf{x}}_t A \underline{\mathbf{y}}$$

Poiché A è per ipotesi simmetrica si ha $A = A_t$ e quindi

$$s(f(v), w) = s(v, f(w))$$

il che prova che f è simmetrica.

Come conseguenza della proposizione ora provata si ha facilmente che le uniche applicazioni lineari simmetriche di R^n in sé sono quelle indotte su R^n da matrici simmetriche.

Per ciò che segue è importante la seguente :

Proposizione 2.2 *Le radici del polinomio caratteristico di una funzione lineare f simmetrica sono tutte reali.*

Dimostrazione. Come già sappiamo il polinomio caratteristico di una funzione lineare f è indipendente dalla base scelta per rappresentare f. Scegliamo quindi in V_n una base B ortonormale e sia A la matrice simmetrica reale che rappresenta f nella base B. Sia λ_o una radice del polinomio p(t) = det(A –It) caratteristico di f e sia $\underline{\mathbf{y}} = (y_1, y_2, \dots, y_n)$ una soluzione non nulla del sistema omogeneo $(A - I\lambda_o)\,\underline{\mathbf{x}} = \underline{\mathbf{0}}$. Si ha quindi $A\,\underline{\mathbf{y}} = \lambda_o \underline{\mathbf{y}}$ **,** esplicitamente è :

$$(*) \qquad \begin{array}{l} a_{11} y_1 + a_{12} y_2 + \dots + a_{1n} y_n = \lambda_o y_1 \\ a_{21} y_1 + a_{22} y_2 + \dots + a_{2n} y_n = \lambda_o y_2 \\ \dots\dots \\ a_{n1} y_1 + a_{n2} y_2 + \dots + a_{nn} y_n = \lambda_o y_n \end{array}$$

indichiamo con :

$\bar{y}_1$ il numero complesso coniugato del numero y_1

$\bar{y}_2$ il numero complesso coniugato del numero y_2

……

$\bar{y}_n$ il numero complesso coniugato del numero y_n

Ora moltiplichiamo la prima delle eguaglianze (*) per $\bar{y}_1$, la seconda per $\bar{y}_2$ e l'ultima per $\bar{y}_n$ e poi sommiamo membro a membro. Si ha allora :

$$a_{11} y_1 \bar{y}_1 + a_{22} y_2 \bar{y}_2 + \ldots + a_{nn} y_n \bar{y}_n + a_{12} (y_2 \bar{y}_1 + y_1 \bar{y}_2) + \ldots +$$

$$+ a_{1n} (y_n \bar{y}_1 + y_1 \bar{y}_n) = \lambda_o (y_1 \bar{y}_1 + y_2 \bar{y}_2 + \ldots + y_n \bar{y}_n)$$

Ora essendo la soluzione $(y_1,y_2,\ldots,y_n)$ non nulla si ha che il coefficiente $(y_1 \bar{y}_1 + y_2 \bar{y}_2 + \ldots + y_n \bar{y}_n)$ di λ_o è un numero reale positivo ed è reale anche il termine a primo membro in quanto i numeri a_{ij} sono reali , i termini $y_i \bar{y}_i$ sono reali ed altrettanto risultano reali i termini $(y_i \bar{y}_j + y_j \bar{y}_i)$ in quanto ciascuno di essi coincide col suo complesso coniugato. Pertanto λ_o risultando rapporto di due numeri reali è reale.

Siamo ora in grado di provare il seguente teorema fondamentale.

Proposizione 2.3. *Sia* $f : V_n \rightarrow V_n$ *una funzione lineare di* V_n *in sé. Esiste una base ortonormale di autovettori per f se e solo se f è simmetrica.*

Dimostrazione. Se esiste una base B ortonormale di autovettori per f la matrice A che rappresenta f nella base B è diagonale e quindi simmetrica. Ma se A è simmetrica la funzione f è simmetrica per la Proposizione 2.1.

Supponiamo ora che $f : V_n \rightarrow V_n$ sia una funzione lineare simmetrica e proviamo che esiste una base ortonormale di autovettori. Ragioniamo per induzione sulla dimensione n dello spazio vettoriale V_n .

Se $n = 1$ allora è $f : V_1 \rightarrow V_1$. Sia w un vettore non nullo di V_1. Poiché la dimensione di V_1 è uno allora w costituisce una base di V_1 e quindi risulta $f(w) =\lambda w$. Pertanto w è un autovettore per f e così $\frac{1}{|w|} w$ è un autovettore di lunghezza 1 e costituisce una base ortonormale per V_1 .

Supponiamo quindi $n > 1$ e vero il teorema per $n-1$. Sia λ_1 un autovalore di f e sia e_1 un autovettore di autovalore λ_1 con lunghezza eguale ad 1. Poiché e_1 è non nullo lo spazio $H_1= [e_1]$ da esso generato ha dimensione 1. Sia W il sottospazio complemento ortogonale del sottospazio H_1 . Come sappiamo il sosttospazio W è supplementare di H_1 e quindi esso ha dimensione $n-1$. Consideriamo la restrizione di f al sottospazio W e proviamo che essa è una funzione di W in W. Infatti sia w un vettore di W e proviamo che anche $f(w)$ appartiene a W . Infatti , essendo f simmetrica , risulta :

$$s (f(w) , e_1) = s (w , f(e_1)) = s (w , \lambda_1 e_1) = \lambda_1 s (w , e_1) = 0$$

Pertanto è $f : W \rightarrow W$ ma poiché è $\dim W = n -1$ allora per l'ipotesi di induzione esiste in W una base $(e_2 , e_3, \ldots, e_n)$ ortonormale di autovettori per f . Aggiungendo ai vettori $(e_2 , e_3, \ldots, e_n)$ il vettore e_1 si ottiene una base $(e_1, e_2 , e_3, \ldots, e_n)$ ortonormale di V_n costituita da autovettori per f .

Il teorema ora provato pur mostrando che una funzione lineare simmetrica è dotata di una base ortonormale di autovettori non fornisce un metodo per costruire tale base .

Il teorema che segue ci aiuta a risolvere questo problema .

Proposizione 2.4 *Sia $f : V_n \rightarrow V_n$ una funzione lineare simmetrica di V_n in sé. Due autovettori v e w di f che abbiano autovalori distinti sono tra loro ortogonali.*

Dimostrazione. Siano v e w due autovettori per f con autovalori λ e η distinti tra loro. Vogliamo provare che v e w risultano tra loro ortogonali e cioè che è $s (v , w) = 0$. Infatti si ha per la simmetria di f :

$$(**) \qquad s (f(v) , w) = s (v , f(w))$$

ed essendo v e w autovettori di autovalori λ e η la (**) diventa :

$$s(\lambda v, w) = s(v, \eta w)$$

da questa segue ancora :

$$\lambda\ s(v, w) = \eta\ s(v, w)\ .$$

Si ha quindi $(\lambda - \eta)\ s(v, w) = 0$ ed essendo $\lambda - \eta \neq 0$ in quanto i due autovalori λ e η sono distinti si ha $s(v, w) = 0$ come si voleva provare.

Per la costruzione della base ortonormale di autovettori per f ora possiamo procedere al seguente modo.

Attraverso il polinomio caratteristico determiniamo gli autovalori di f . Tali autovalori sono n e tutti reali e ciascuno ha molteplicità algebrica eguale a quella geometrica in quanto f essendo simmetrica è diagonalizzabile (cfr. proposizione 2.4).

Siano λ_1 , λ_2 …. λ_t gli autovalori distinti di f e siano H_1 , H_2 ,…, H_t i corrispondenti autospazi. Si costruisca in ogni sottospazio H_i una base B_i ortonormale . Unendo tra loro le basi B_1 , B_2 , …, B_t si ottiene una base ortonormale di autovettori per f .

Un altro modo di procedere , più adatto alle applicazioni , è il seguente. Si fissi in V_n una base **B**=(e_1, e_2 , e_3, …, e_n) ortonormale e si consideri la matrice A simmetrica che rappresenta f nella base fissata. L' applicazione lineare

$$A : R^n \rightarrow R^n$$

indotta da A è simmetrica ed è quindi dotata di una base ortonormale di autovettori .

Come sappiamo f ed A hanno gli stessi autovalori ed inoltre ogni autovettore di A è il vettore delle coordinate nella base **B** di un autovettore per f .

Siano quindi

$$\underline{p}_1 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ . \\ \alpha_n \end{pmatrix}, \quad \underline{p}_2 = \begin{pmatrix} \beta_1 \\ \beta_2 \\ . \\ \beta_n \end{pmatrix}, \ldots\ldots\ldots, \quad \underline{p}_n = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ . \\ \gamma_n \end{pmatrix}$$

n autovettori di A a due a due ortogonali e ciascuno di lunghezza 1.

I vettori

$$v_1 = \alpha_1\ e_1 + \alpha_2\ e_2 + .\ ....+ \alpha_n\ e_n$$

$$v_2 = \beta_1\ e_1 + \beta_2\ e_2 + .\ ....+ \beta_n\ e_n$$

.....

$$v_n = \gamma_1\ e_1 + \gamma_2\ e_2 + .\ ....+ \gamma_n\ e_n$$

sono quindi autovettori per f . Inoltre essendo la base **B** ortonormale risulta

$$s\ (\ v_i\ ,\ v_j) = <\ \underline{p}_i,\ \underline{p}_j >$$

e quindi $(v_1, v_2, \ldots, v_n)$ è una base ortonormale di autovettori per f.

Facciamo qualche esempio.

Esercizio.

Si stabilisca se la seguente funzione lineare

$$f(\ x\ ,\ y\ )\ = (\ 2x + y\ ,\ \ x + 2y)$$

di R^2 in sé è simmetrica . In caso affermativo si determini una base ortogonale di autovettori per f .

Svolgimento. Si consideri la base canonica B = {(1,0) , (0,1) } di R^2 . Tale base è ortonormale e la matrice che rappresenta f in tale base è la seguente :

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

Poiché A è simmetrica la funzione f è simmetrica.

Gli autovalori di f sono le radici del polinomio caratteristico

$$p(t) = \begin{pmatrix} 2-t & 1 \\ 1 & 2-t \end{pmatrix} = (\ 2-t\ )^2 - 1 = t^2 - 4t + 3 = 0$$

e sono quindi i numeri 1 e 3 .Gli autovettori di A sono anche autovettori per f in quanto la base B è la base canonica. Gli autovettori di autovalore 1 si ottengono attraverso le soluzioni del sistema omogeneo

$$x + y = 0$$

e sono quindi (1 , -1) e tutti i suoi multipli.

Gli autovettori di autovalore 3 si ottengono attraverso le soluzioni del sistema omogeneo

$$-x + y = 0$$

e sono quindi (1 , 1) e tutti i suoi multipli. Pertanto {(-1,1) , (1, 1) } è una base ortogonale di autovettori ognuno dei quali ha lunghezza $\sqrt{2}$. Moltiplicando ciascuno di essi per $\frac{1}{\sqrt{2}}$ si ottiene una base ortonormale di autovettori.

Indice