APPUNTI DI
ALGEBRA LINEARE

prof. MICHELE MININNI

a.a. 2014/15

Versione provvisoria.
Le segnalazioni di errori sono ben accette, anzi sono decisamente sollecitate.
e-mail: michele.mininni@uniba.it

INDICE

Capitolo V - Trasformazioni Lineari

Capitolo VI - Ortogonalitá

Capitolo VII - Autovalori - Autovettori

Capitolo VIII - Forme Quadratiche

Capitolo IX - Complementi

Capitolo I - CALCOLO VETTORIALE e MATRICIALE

1 . - Vettori e Calcolo Vettoriale

Si dice **vettore** o punto di $\mathbf{R}^n$ ogni n-pla ordinata di numeri reali, che diconsi le coordinate del vettore. Un vettore viene di solito rappresentato come una riga e piú spesso come una colonna:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \quad \text{vettore riga ,} \qquad\qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{vettore colonna.}$$

Se $n = 2$, (rispett. $n = 3$), il vettore $\mathbf{x} = (x_1, x_2)$, (rispett. $\mathbf{x} = (x_1, x_2, x_3)$), puó essere identificato con il punto P del piano (rispett. dello spazio tridimensionale) di coordinate (x_1, x_2), (rispett. (x_1, x_2, x_3)), rispetto ad un assegnato riferimento cartesiano ortogonale monometrico e quindi con il segmento orientato OP. Tale segmento orientato spiega la denominazione di "vettore", poiché esso determina una *direzione* nel piano o nello spazio tridimensionale, un *verso*, (quello da O verso P), ed una *intensitá*, data dalla lunghezza del segmento OP.
In particolare il vettore $(0, 0)$, (rispett. $(0, 0, 0)$), puó essere identificato con il punto O, origine del riferimento cartesiano.

Si noti che dal teorema di Pitagora segue che la lunghezza del segmento OP é data da

$$\sqrt{x_1^2 + x_2^2}, \qquad \text{(rispett. } \sqrt{x_1^2 + x_2^2 + x_3^2} \text{).}$$

In generale, se $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ é un vettore di $\mathbf{R}^n$, dicesi **lunghezza** o **intensitá** o **norma euclidea** del vettore $\mathbf{x}$ il numero reale

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \geq 0.$$

■

Nell'insieme $\mathbf{R}^n$ dei vettori si definiscono un' operazione di **addizione tra due vettori** e un' operazione di **moltiplicazione di un numero per un vettore**. Infatti, dati un numero reale α e due vettori $\mathbf{x}$ e $\mathbf{y}$, diconsi rispettivamente **vettore somma** di $\mathbf{x}$ e $\mathbf{y}$ e **vettore prodotto** di α e $\mathbf{x}$ i vettori $\mathbf{x}+\mathbf{y}$ ed $\alpha\mathbf{x}$ le cui generiche coordinate $i-$esime sono rispettivamente la somma delle coordinate i-esime di x e y e il prodotto di α con la coordinata $i-$esima di $\mathbf{x}$. In altri termini si pone

$$(1) \qquad \mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad \alpha\mathbf{x} = \alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Ad esempio

$$\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \begin{pmatrix} 0 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \qquad -2 \begin{pmatrix} 1 \\ -3 \\ 5/2 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ -5 \end{pmatrix}.$$

Il vettore che ha tutte le coordinate nulle si dice **vettore nullo** e viene denotato $\mathbf{0}$; inoltre, per ogni $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, il vettore $(-1) \cdot \mathbf{x} = (-x_1, -x_2, \ldots, -x_n)$ viene denotato $-\mathbf{x}$ e viene detto **vettore opposto** ad $\mathbf{x}$. Inoltre, dicesi **vettore differenza** tra $\mathbf{y}$ e $\mathbf{x}$ il vettore $\mathbf{y} - \mathbf{x} = \mathbf{y} + (-\mathbf{x})$.

Infine dati due vettori $\mathbf{x}$ e $\mathbf{y}$ si dice **prodotto scalare** di $\mathbf{x}$ e $\mathbf{y}$ il numero reale

$$(\mathbf{x}|\mathbf{y}) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Ad esempio se $\mathbf{x} = (1, 2, -1/2)$ ed $\mathbf{y} = (-2, 1/3, -3)$, allora

$$(\mathbf{x}|\mathbf{y}) = 1 \cdot (-2) + 2 \cdot 1/3 + (-1/2) \cdot (-3) = 1/6.$$

■

Osservazione 1.1 - Nel caso $n = 2$, possiamo dare una interpretazione geometrica interessante del vettore $-\mathbf{x}$, del vettore prodotto di uno scalare α per un vettore $\mathbf{x}$, del vettore somma $\mathbf{x} + \mathbf{y}$, del vettore differenza $\mathbf{y} - \mathbf{x}$ e del prodotto scalare dei vettori $\mathbf{x}$ e $\mathbf{y}$.

Infatti, detti P e Q i punti del piano di cui $\mathbf{x}$ e $\mathbf{y}$ sono i vettori delle coordinate rispetto ad un assegnato riferimento cartesiano ortogonale monometrico, allora si ha quanto segue.

(1) $-\mathbf{x}$ é il vettore delle coordinate del punto P' simmetrico di P rispetto all'origine O; in altri termini, $-\mathbf{x}$ ha la stessa direzione, la stessa lunghezza, ma verso opposto rispetto al vettore $\mathbf{x}$.

(2) Se $\alpha > 0$, allora il vettore $\alpha\mathbf{x}$ ha la stessa direzione e lo stesso verso del vettore $\mathbf{x}$ ed una lunghezza pari al prodotto di α per la lunghezza di $\mathbf{x}$.

Infatti, detto S il punto di coordinate $(\alpha x_1, \alpha x_2)$ e detti H e K le proiezioni di P ed S sull'asse x_1, cioé i punti di coordinate $(x_1, 0)$ e $(\alpha x_1, 0)$, si ha che i triangoli rettangoli OHP ed OKS sono simili, (poiché hanno i cateti proporzionali secondo il fattore α); ne segue che gli angoli di tali triangoli hanno la stessa ampiezza, e quindi che i punti P ed S si trovano sulla stessa semiretta uscente dall'origine O; inoltre la lunghezza del segmento OS é uguale al prodotto di α per la lunghezza del segmento OP.

(3) Se $\alpha < 0$, allora $\alpha\mathbf{x} = -|\alpha|\mathbf{x}$ é il vettore che ha la stessa direzione e verso opposto al vettore $\mathbf{x}$ ed una lunghezza pari al prodotto di $|\alpha|$ per la lunghezza di $\mathbf{x}$;

(4) $\mathbf{x} + \mathbf{y}$ rappresenta il vettore delle coordinate del punto R, quarto vertice del parallelogramma di cui due lati sono i segmenti OP ed OQ,

(5) $\mathbf{y} - \mathbf{x} = \mathbf{y} + (-\mathbf{x})$ é il vettore delle coordinate cartesiane del punto R', quarto vertice del parallelogramma di cui due lati sono i segmenti OP' ed OQ, ed é evidente che il segmento orientato OR' é parallelo ed ha la stessa lunghezza e lo stesso verso del segmento orientato PQ.

In altre parole, nel parallelogramma $OPRQ$ costruito sui vettori $\mathbf{x}$ e $\mathbf{y}$, il vettore somma $\mathbf{x} + \mathbf{y}$ rappresenta la *diagonale principale OR,* mentre il vettore differenza $\mathbf{y} - \mathbf{x}$ rappresenta la *diagonale secondaria* PQ, nel verso da P a Q, cioé dalla punta di $\mathbf{x}$ alla punta di $\mathbf{y}$.

Infine, per quanto riguarda il prodotto scalare di $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$, dette α e β le ampiezze degli angoli formati dal semiasse positivo delle ascisse rispettivamente con la semiretta OP e con la semiretta OQ, risulta: $x_1 = ||\mathbf{x}|| \cos\alpha$, $x_2 = ||\mathbf{x}|| \sin\alpha$, $y_1 = ||\mathbf{y}|| \cos\beta$, $y_2 = ||\mathbf{y}|| \sin\beta, ||$ e quindi:

$$(\mathbf{x}|\mathbf{y}) = x_1 y_1 + x_2 y_2 = ||\mathbf{x}|| \cdot ||\mathbf{y}|| \cdot (\cos\alpha\cos\beta + \sin\alpha\sin\beta) = ||\mathbf{x}|| \cdot ||\mathbf{y}|| \cdot \cos(\alpha - \beta).$$

In altri termini si ha che

(6) *il prodotto scalare di due vettori piani é uguale al prodotto delle loro norme per il coseno dell'angolo $\theta = \alpha - \beta$ compreso tra di essi, cioé tra le semirette orientate OP ed OQ .*

In particolare, le semirette OP ed OQ sono ortogonali se e solo se risulta $\cos\theta = 0$ e dunque se e solo se risulta $(\mathbf{x}|\mathbf{y}) = 0$.

■

Si puó dimostrare che le proprietá (1)-(6) continuano a valere nel caso di vettori dello spazio tridimensionale; in particolare, anche nel caso tridimesionale, le rette OP ed OQ sono ortogonali se e solo se risulta $(\mathbf{x}|\mathbf{y}) = 0$.

Per questo motivo, in generale due vettori di $\mathbf{R}^n$ si dicono **ortogonali** se risulta $(\mathbf{x}|\mathbf{y}) = 0$.

■

Osservazione 1.2 - Si verifica immediatamente che, per ogni $\alpha, \beta \in \mathbf{R}$ e per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^n$, sussistono le seguenti proprietá:

(A1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$,
(A2) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$,
(A3) $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$,
(A4) $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$,
(M1) $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$,
(M2) $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$,
(M3) $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$,
(M4) $1\mathbf{x} = \mathbf{x}$,

Infine, le principali proprietá del prodotto scalare sono le seguenti:

(S1) $(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})$,
(S2) $(\mathbf{x}' + \mathbf{x}''|\mathbf{y}) = (\mathbf{x}'|\mathbf{y}) + (\mathbf{x}''|\mathbf{y}), \quad (\mathbf{x}|\mathbf{y}' + \mathbf{y}'') = (\mathbf{x}|\mathbf{y}') + (\mathbf{x}|\mathbf{y}'')$,
(S3) $(\alpha\mathbf{x}|\mathbf{y}) = \alpha(\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\alpha\mathbf{y})$
(S4) $(\mathbf{x}|\mathbf{x}) = x_1^2 + x_2^2 + \ldots + x_n^2 = ||\mathbf{x}||^2 \geq 0$,
(S5) $(\mathbf{x}|\mathbf{x}) = 0$ se e solo se $\mathbf{x} = 0$.

■

Osservazione 1.3 - Dalla proprietá associativa dell'addizione si deduce che é possibile definire la somma di tre vettori, ponendo:

$$\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3) \qquad \text{per ogni } \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{R}^n.$$

In maniera analoga si puó definire il vettore somma di quattro vettori, e in generale di n vettori.

Inoltre, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ sono k vettori ed $\alpha_1, \alpha_2, \ldots, \alpha_k$ sono k numeri reali, allora il vettore

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k$$

dicesi **combinazione lineare** dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ secondo i coefficienti $\alpha_1, \alpha_2, \ldots, \alpha_k$.

■

Osservazione 1.4 - Tutto quanto detto sopra puó essere ripetuto quasi senza alcuna modifica sostituendo i numeri reali con i numeri complessi.

Ad esempio $(i, 1 + i, 2 - 3i, -4)$ é un vettore complesso quadridimensionale. In generale dicesi vettore complesso n-dimensionale una qualunque n-pla ordinata $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ di numeri complessi; l'insieme di tali vettori viene indicato con il simbolo $\mathbf{C}^n$.

La norma di un vettore $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ di numeri complessi é definita ponendo

$$||\mathbf{x}|| = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2} \geq 0,$$

dove $|\alpha|$ é il modulo di un qualunque numero complesso α.

La somma di due vettori di $\mathbf{C}^n$ e il prodotto di un numero complesso α per un vettore di $\mathbf{C}^n$ vengono definite esattamente come nella (1).

Invece il prodotto scalare di due vettori $\mathbf{x}$ e $\mathbf{y}$ di $\mathbf{C}^n$ viene definito ponendo

$$(\mathbf{x}|\mathbf{y}) = x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n},$$

dove $\overline{\alpha}$ é il numero complesso coniugato di α, per ogni numero complesso α.

Si vede facilmente che le proprietá (A1-A4) ed (M1-M4) continuano a valere anche nell'ambito dei vettori di $\mathbf{C}^n$. Le proprietá del prodotto scalare devono invece essere leggermente modificate come segue:

(S1)' $(\mathbf{x}|\mathbf{y}) = \overline{(\mathbf{y}|\mathbf{x})}$,
(S2)' $(\mathbf{x}' + \mathbf{x}''|\mathbf{y}) = (\mathbf{x}'|\mathbf{y}) + (\mathbf{x}''|\mathbf{y}), \quad (\mathbf{x}|\mathbf{y}' + \mathbf{y}'') = (\mathbf{x}|\mathbf{y}') + (\mathbf{x}|\mathbf{y}'')$,
(S3)' $(\alpha\mathbf{x}|\mathbf{y}) = \alpha(\mathbf{x}|\mathbf{y}), \quad (\mathbf{x}|\alpha\mathbf{y}) = \overline{\alpha}(\mathbf{x}|\mathbf{y})$,
(S4)' $(\mathbf{x}|\mathbf{x}) = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 = ||\mathbf{x}||^2 \geq 0$,
(S5)' $(\mathbf{x}|\mathbf{x}) = 0$ se e solo se $\mathbf{x} = 0$.

■

Osservazione 1.5. - Il concetto di spazio vettoriale puó essere esteso, nel senso che dicesi **spazio vettoriale reale** (rispett. complesso) un qualunque insieme V in cui si puó definire una operazione di addizione tra due suoi elementi ed una operazione esterna di moltiplicazione di un numero reale (o complesso) per un elemento di V, in modo che siano soddisfatte le proprietá (A1-A4) ed (M1-M4), (piú esattamente in cui esiste un elemento $\mathbf{0}$ (detto l'elemento nullo) che verifica la (A3) e per ogni elemento $\mathbf{x}$ esiste un elemento $-\mathbf{x}$ (detto l' elemento opposto ad $\mathbf{x}$) che verifica la (A4).

Ad esempio, l' insieme delle funzioni reali definite in un intervallo I o l' insieme delle funzioni continue in I, con le usuali operazioni di addizione di due funzioni e di moltiplicazione di una funzione costante per una funzione, sono spazi vettoriali reali. L' elemento nullo é la funzione costante di costante valore 0; l' opposto di una funzione f é la funzione $-f : x \to -f(x)$ ed é evidente che sono soddisfatte le proprietá (A1)-(A4) ed (M1)-(M4).

Un altro esempio di spazio vettoriale é l'insieme dei polinomi a coefficienti reali (o complessi) con le usuali operazioni di addizione tra due polinomi e di moltiplicazione di un numero per un polinomio.

Infine, dicesi **spazio euclideo** uno spazio vettoriale reale V in cui é definito un **prodotto scalare**, cioé una funzione $(\cdot|\cdot) : V \times V \mapsto \mathbf{R}$ soddisfacente le proprietá (S1)-(S5).
Lo spazio $\mathbf{R}^n$ con l'usuale prodotto scalare é quindi uno spazio euclideo; un altro esempio di spazio euclideo é fornito dallo spazio $V = \mathcal{C}([a,b],\mathbf{R})$ delle funzioni continue in un intervallo $[a,b]$, con il prodotto scalare

$$(u|v) = \int_a^b u(x)v(x)\,dx, \qquad \text{per ogni} \quad u,v \in V = \mathcal{C}([a,b],\mathbf{R}).$$

Nel seguito ci limiteremo a considerare gli spazi euclidei del tipo $\mathbf{R}^n$, e accenneremo poi nell'ultimo capitolo ad alcune estensioni degli argomenti trattati al caso di spazi vettoriali o euclidei astratti.

■

2 . - Matrici di tipo $k \times n$

Dicesi **matrice di tipo** $k \times n$ o **matrice con** k **righe ed** n **colonne** ogni tabella di $k \cdot n$ numeri reali disposti su k righe ed n colonne. Il generico elemento della matrice viene contrassegnato da una coppia ordinata di indici, che individuano rispettivamente la riga e la colonna a cui l' elemento appartiene. Ad esempio $a_{2,3}$ sta ad indicare l'elemento della seconda riga e terza colonna della matrice.

Una matrice verrá in generale indicata da una lettera latina maiuscola $A, B, \ldots$. Perció una matrice A verrá indicata con il simbolo

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k,1} & a_{k,2} & \ldots & a_{k,n} \end{pmatrix}.$$

■

Nel seguito adotteremo la seguente convenzione:

A_i rappresenta la riga $i-$esima della matrice A, per ogni $i = 1, 2, \ldots, k$,

A^j rappresenta la colonna $j-$esima della matrice A, per ogni $j = 1, 2, \ldots, n$,

Se $k = 1$ allora la matrice si riduce ad avere una sola riga; se $n = 1$ la matrice si riduce ad avere una sola colonna. Pertanto i vettori riga o i vettori colonna sono particolari matrici.

■

Se A é una matrice con k righe ed n colonne, allora la matrice con n righe e k colonne che si ottiene scambiando le righe con le colonne si dice **matrice trasposta** di A e si indica con il simbolo A^T.

Pertanto la prima riga di A^T coincide con la prima colonna di A, la seconda riga di A^T coincide con la seconda colonna di A, e cosí via. Ad esempio si ha

$$\begin{pmatrix} 1 & 2 & 4 \\ -2 & 0 & -3 \end{pmatrix}^T = \begin{pmatrix} 1 & -2 \\ 2 & 0 \\ 4 & -3 \end{pmatrix}.$$

In particolare un vettore colonna non é altro che la matrice trasposta del vettore riga corrispondente e viceversa un vettore riga é la matrice trasposta del vettore colonna corrispondente. Di conseguenza il simbolo $(x_1, x_2, \dots, x_n)^T$ potrá essere usato per indicare il vettore colonna di coordinate $x_1, x_2, \dots, x_n$.

E' del tutto evidente che **la trasposta della trasposta di una matrice A concide con A.**

■

Talvolta in una matrice A con k righe ed n colonne é utile distinguere la sottomatrice B formata dalle prime k_1 righe di A e la sottomatrice C formata dalle restanti $k_2 = k - k_1$ righe. In tal caso la matrice A si dice decomposta in blocchi e viene rappresentata con il simbolo

$$A = \begin{pmatrix} B \\ C \end{pmatrix}.$$

Analogamente il simbolo

$$A = (B \quad C)$$

serve a rappresentare la decomposizione di A nei due blocchi formati rispettivamente dalle prime n_1 colonne e dalle restanti $n_2 = n - n_1$ colonne.

Naturalmente é possibile decomporre una matrice in piú di due blocchi. Ad esempio si puó porre

$$A = \begin{pmatrix} B_1 & C_1 \\ B_2 & C_2 \end{pmatrix},$$

dove

B_1 rappresenta la sottomatrice di A formata dalle prime k_1 righe e dalle prime n_1 colonne,

C_1 rappresenta la sottomatrice di A formata dalle prime k_1 righe e dalle restanti $n_2 = n - n_1$ colonne,

B_2 rappresenta la sottomatrice di A formata dalle ultime $k_2 = k - k_1$ righe e dalle prime n_1 colonne,

C_2 rappresenta la sottomatrice di A formata dalle ultime k_2 righe e dalle ultime n_2 colonne.

■

3 . - Matrici quadrate

Se il numero k delle righe coincide con il numero n delle colonne si dice che la matrice é una **matrice quadrata di ordine** n.

Se A é una matrice quadrata, i numeri $a_{1,1}, a_{2,2}, \dots, a_{n,n}$ si dice che formano la **diagonale** di A.

Si dice che A é una matrice quadrata **simmetrica** se coincide con la sua trasposta A^T e quindi se gli elementi simmetrici rispetto alla diagonale $a_{i,j}$ ed $a_{j,i}$ sono uguali.

Ad esempio é simmetrica la matrice

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

Si dice che A é una **matrice triangolare superiore** se sono nulli tutti gli elementi di A al di sotto della diagonale, cioé se risulta $a_{ij} = 0$ per ogni i, j tali che $i > j$.

Si dice invece che A é una **matrice triangolare inferiore** se sono nulli tutti gli elementi di A al di sopra della diagonale, cioé se risulta $a_{ij} = 0$ per ogni i, j tali che $i < j$.

Infine si dice che A é una **matrice diagonale** se sono nulli tutti gli elementi di A tranne quelli della diagonale, cioé se risulta $a_{ij} = 0$ per ogni $i \neq j$.

Ad esempio delle matrici

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

si ha che A é triangolare superiore, B é triangolare inferiore e C é una matrice diagonale.

■

Tra le matrici diagonali assume una particolare importanza la matrice che ha uguali ad 1 tutti gli elementi della diagonale (ed uguali a 0 tutti gli altri elementi); tale matrice prende il nome di **matrice identica** e si denota con il simbolo I, (o piú esattamente I_n se é necessario precisare l'ordine n della matrice).

4 . - Operazioni tra matrici.

Nell' insieme delle matrici di tipo $k \times n$ si definiscono un'operazione di addizione tra due matrici e una operazione di moltiplicazione di un numero reale α per una matrice. Si pone infatti:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{k,1} & a_{k,2} & \dots & a_{k,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \dots & \dots & \dots & \dots \\ b_{k,1} & b_{k,2} & \dots & b_{k,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \dots & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \dots & a_{2,n}+b_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{k,1}+b_{k,1} & a_{k,2}+b_{k,2} & \dots & a_{k,n}+b_{k,n} \end{pmatrix}$$

$$\alpha \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & & & \\ a_{k,1} & a_{k,2} & \dots & a_{k,n} \end{pmatrix} = \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \dots & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \dots & \alpha a_{2,n} \\ \dots & & & \\ \alpha a_{k,1} & \alpha a_{k,2} & \dots & \alpha a_{k,n} \end{pmatrix}.$$

In altri termini, $A + B$ si ottiene sommando ordinatamente le righe (o le colonne) di A con le righe (o le colonne) di B, mentre αA si ottiene moltiplicando per α le righe (o le colonne) di A, cioé

$$(A+B)_i = A_i + B_i, \qquad (\alpha A)_i = \alpha A_i \quad \text{per ogni } i = 1, 2, \dots, k,$$
$$(A+B)^j = A^j + B^j, \qquad (\alpha A)^j = \alpha A^j \quad \text{per ogni } j = 1, 2, \dots, n.$$

Ad esempio si ha

$$\begin{pmatrix} 1 & 2 & 4 \\ -2 & 0 & -3 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 3 \\ -4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 7 \\ -6 & 1 & -1 \end{pmatrix},$$

$$-2 \begin{pmatrix} 1 & 2 & 0 \\ 2 & -3 & 5/2 \end{pmatrix} = \begin{pmatrix} -2 & -4 & 0 \\ -4 & 6 & -5 \end{pmatrix}.$$

■

La matrice che ha tutti gli elementi nulli si dice **matrice nulla** e viene denotata O; per ogni matrice A la matrice $(-1)A$ il cui generico elemento é $-a_{ij}$ viene denotata $-A$ e viene detta **matrice opposta** della matrice A. Inoltre si pone $A - B = A + (-B)$ per ogni coppia di matrici A, B di tipo $k \times n$.

Si verifica allora immediatamente che le operazioni di addizione tra due matrici e di moltiplicazione di uno scalare per una matrice godono delle seguenti proprietá:

(A1) $A + B = B + A$,
(A2) $A + (B + C) = (A + B) + C$,
(A3) $A + O = O + A = A$,
(A4) $A + (-A) = (-A) + A = O$,
(M1) $\alpha(A + B) = \alpha A + \alpha B$,
(M2) $(\alpha + \beta)A = \alpha A + \beta A$,
(M3) $\alpha(\beta A) = (\alpha\beta)A$,
(M4) $1A = A$.

■

5 . - Prodotto matriciale

Se ora A é una matrice di tipo $k \times n$ e B é una matrice di tipo $n \times p$ si dice **matrice prodotto** di A e B la matrice $C = A\,B$ con k righe e p colonne il cui elemento generico $c_{i,j}$ é il prodotto scalare della riga $i-$esima di A per la colonna $j-$esima di B, cioé

$$c_{ij} = (A_i|B^j) = \sum_{h=1}^{n} a_{ih}b_{hj}, \qquad \text{per ogni } \; i = 1, 2, \dots, k \quad \text{ e per ogni } \; j = 1, 2, \dots, p.$$

■

E' evidente dalla definizione che il prodotto di due matrici si puó calcolare se e solo se il numero delle colonne della prima matrice coincide con il numero delle righe della seconda.

Di conseguenza, se pure esiste la matrice $A\,B$, in generale non esiste la matrice $B\,A$, a meno che non sia $p = k$; in tal caso peró $A\,B$ é una matrice quadrata di ordine k mentre $B\,A$ é una matrice quadrata di ordine n, e quindi se $k \neq n$ sará $A\,B \neq B\,A$.

D' altra parte, anche se risulta $k = n = p$, (e quindi A, B, AB e BA sono tutte matrici quadrate di ordine n), in generale le matrici AB e BA sono diverse, come mostra il seguente esempio:

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Dunque il prodotto matriciale **non gode** della proprietá commutativa $AB = BA$.

■

Si puó facilmente verificare invece che (a condizione che i prodotti abbiano senso) sussistono le seguenti proprietá:

(P1) $A(BC) = (AB)C$,
(P2) $I_k A = A, \qquad AI_n = A$,
(P3) $A(B + C) = AB + AC, \quad (B + C)A = BA + CA$,
(P4) $(\alpha A)B = \alpha(AB) = A(\alpha B)$,
(P5) $(AB)^T = B^T A^T$.

■

6 . - Algoritmo di Gauss Jordan

Se A é una matrice non nulla di tipo $k \times n$, le seguenti operazioni:

a) scambiare di posto due righe o due colonne,
b) aggiungere ad una riga un multiplo di un' altra,
c) moltiplicare o dividere una riga per una costante $\alpha \neq 0$,

vengono dette **operazioni elementari** sulla matrice A.

Ora, se $a_{hj} \neq 0$, adoperando in maniera sistematica la b), é possibile trasformare la matrice A in una matrice in cui sono nulli tutti gli elementi della colonna $j-$esima ad eccezione di a_{hj}.

Ad esempio, data la matrice

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

considerato che $a_{1,1} = 1 \neq 0$, sostituendo la riga A_2 con $A_2 + 2A_1$, la riga A_3 con $A_3 - 3A_1$ e la riga A_4 con $A_4 + 3A_1$ si ottiene la matrice

$$A' = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 7 & -2 & -3 \\ 0 & -8 & 2 & 2 \\ 0 & 8 & -3 & -2 \end{pmatrix},$$

in cui nella prima colonna solo il primo elemento $a_{1,1}$ é diverso da 0.

Analogamente sostituendo la riga A_1 con $A_1 + A_2$, la riga A_3 con $A_3 + 2A_2$ e la riga A_4 con $A_4 - (3/2)A_2$ si ottiene la matrice

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

in cui nella terza colonna solo l' elemento $a_{2,3} = 2$ é diverso da 0.

Se poi, oltre ad effettuare le operazioni precedenti, dividiamo per a_{hj} la riga h-sima di A, allora A sará stata trasformata in una matrice in cui la j-sima colonna ha tutti gli elementi nulli tranne l'h-simo che é uguale ad 1, e quindi coincide con la $h-$sima colonna della matrice identica I_k.

La procedura di sostituzione delle righe di A in modo da rendere nulli tutti gli elementi della colonna $j-$esima tranne a_{hj} e rendere uguale ad 1 tale elemento dicesi **passo di pivot** su a_{hj} e l'elemento non nullo a_{hj} dicesi **pivot** o cardine del processo.

Tale procedura consiste nel :

(1) dividere la riga h-esima per a_{hj},
(2) sostituire, (per ogni $i \neq h$), la riga i-esima A_i con $A_i - (a_{ij}/a_{hj})A_h$.

Se ci limitiamo ad eseguire l'operazione (2) solo sulle righe sotto il pivot, cioé solo per $i > h$, allora A verrá trasformata in una matrice in cui sono nulli tutti gli elementi della colonna j-esima sotto il pivot a_{hj}. Questa seconda procedura verrá detta **passo di pivot parziale** su a_{hj}.

■

Ció premesso descriviamo un procedimento che ci sará di grande utilitá nel seguito, in quanto consente di trasformare qualunque matrice non nulla in una matrice particolarmente semplice.

Proposizione 6.1 - (Algoritmo di Gauss Jordan). *Mediante passi di pivot parziali, scambi di riga e, (se necessario), scambi di colonna, é possibile trasformare una qualunque matrice non nulla A di tipo $k \times n$ in una matrice A' che puó essere decomposta a blocchi in una delle forme*

$$B, \qquad (\,B \quad C\,), \qquad \begin{pmatrix} B \\ O \end{pmatrix}, \qquad \begin{pmatrix} B & C \\ O & O \end{pmatrix},$$

dove B é una matrice triangonale superiore con elementi diagonali diversi da 0.

Dim. In primo luogo, scambiando eventualmente la prima riga con una delle successive o, (se necessario), scambiando la prima colonna con una delle successive, possiamo fare in modo che $a_{1,1}$ sia diverso da 0. Eseguendo un passo di pivot parziale su $a_{1,1}$, la matrice A viene trasformata in una matrice di tipo

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{k,2} & \dots & a_{k,n} \end{pmatrix}, \qquad \text{cioé} \qquad \begin{pmatrix} B & C \\ O & C' \end{pmatrix} \qquad \text{dove} \qquad B = (a_{11}).$$

A questo punto se C' é la matrice nulla, la tesi é stata dimostrata; in caso contrario, scambiando eventualmente la attuale seconda riga con una delle successive o, (se necessario), scambiando la attuale seconda colonna con una delle successive, possiamo fare in modo che $a_{2,2}$ sia diverso da 0. Eseguendo un passo di pivot parziale su $a_{2,2}$, la matrice A viene trasformata in una matrice di tipo

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{k,n} \end{pmatrix}, \qquad \text{cioé} \qquad \begin{pmatrix} B & C \\ O & C' \end{pmatrix} \qquad \text{dove} \qquad B = \begin{pmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix}.$$

A questo punto se C' é la matrice nulla, la tesi é stata dimostrata; in caso contrario si ripete lo stesso procedimento, finché é possibile.

Pertanto l'algoritmo si interrompe:

- perché non sono state esaurite né le righe, né le colonne, ma le righe residue sono nulle, e quindi A é stata trasformata in una matrice A' del tipo $\begin{pmatrix} B & C \\ O & O \end{pmatrix}$,
- perché sono state esaurite le righe e le colonne di A, e quindi A é stata trasformata in una matrice del tipo B,
- perché sono state esaurite le righe (ma non le colonne) di A, e quindi A é stata trasformata in una matrice del tipo $(\,B \quad C\,)$,
- perché sono state esaurite le colonne (ma non le righe) di A, e quindi A é stata trasformata in una matrice del tipo $\begin{pmatrix} B \\ O \end{pmatrix}$,

con B matrice triangolare superiore con elementi diagonali diversi da 0.

■

Osservazione 6.2 - Si noti che se modifichiamo l'algoritmo di Gauss Jordan, nel senso che effettuiamo passi di pivot completi anziché parziali, ed effettuiamo un passo di pivot completo anche sull'eventuale elemento diagonale non nullo della k-sima riga, allora A verrá trasformata in una matrice A' che si puó decomporre a blocchi in una delle forme

$$I_k, \qquad (\,I_k \quad C\,), \qquad \begin{pmatrix} I_n \\ O \end{pmatrix}, \qquad \begin{pmatrix} I_r & C \\ O & O \end{pmatrix}.$$

Tale forma dell'algoritmo di Gauss Jordan verrá detta **algoritmo di Gauss Jordan completo.**

■

Osservazione 6.3 - É evidente che se A é una matrice quadrata di ordine n, allora l'algoritmo di Gauss Jordan trasforma A in una matrice del tipo B o del tipo $\begin{pmatrix} B & C \\ O & O \end{pmatrix}$, con B matrice triangolare superiore con elementi diagonali diversi da 0 e l'algoritmo di Gauss Jordan completo trasforma A nella matrice identica I_n o in una matrice del tipo $\begin{pmatrix} I_r & C \\ O & O \end{pmatrix}$.

Ebbene, se durante l'esecuzione dell'algoritmo di Gauss Jordan si rende necessario uno scambio di colonne, vuol dire che si é trovato un indice $h < n$ tale che $a'_{i,h} = 0$, per $i = h, h+1, \dots, n$; nelle successive fasi dell'algoritmo, la colonna h-sima cambierá posizione, ma i suoi elementi non cambieranno, sicché alla fine A sará stata trasformata in una matrice in cui l'ultima riga é la riga nulla, e quindi in una matrice del tipo $\begin{pmatrix} B & C \\ O & O \end{pmatrix}$,

Pertanto se l'algoritmo di Gauss Jordan completo ha trasformato A nella matrice identica I_n, allora questo risultato é stato ottenuto senza scambi di colonne, ma solo con scambi di riga e passi di pivot.

■

Osservazione 6.4 - Possiamo renderci conto facilmente di cosa realmente facciamo quando su una matrice A di tipo $k \times n$ eseguiamo l'algoritmo di Gauss Jordan o l'algoritmo di Gauss Jordan completo.

Infatti si verifica facilmente che se la matrice (A, I_k) viene trasformata in (A', B) mediante una delle operazioni

a) scambiare di posto due righe,
b) aggiungere ad una riga un multiplo di un' altra riga,
c) moltiplicare gli elementi di una riga per una costante $\alpha \neq 0$,

allora risulta $A' = B \cdot A$.

Ad esempio, se (A', B) é ottenuta da (A, I_k) scambiando la prima con la seconda riga, allora si ha

$$B \cdot A = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \cdot A = \begin{pmatrix} a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \dots & \dots & \dots & \dots \\ a_{k,1} & a_{k,2} & \dots & a_{k,n} \end{pmatrix} = A'.$$

Analogamente se (A', B) é ottenuta da (A, I_k) sostituendo la seconda riga di (A, I_k) con la somma della seconda riga con il multiplo della prima riga secondo il fattore α, allora si ha

$$B \cdot A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \alpha & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \cdot A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \alpha a_{1,1} + a_{2,1} & \alpha a_{1,2} + a_{2,2} & \dots & \alpha a_{2,n} + a_{1,n} \\ \dots\dots & & & \\ a_{k,1} & a_{k,2} & \dots & a_{k,n} \end{pmatrix} = A'.$$

Infine se (A', B) é ottenuta da (A, I_k) moltiplicando per $\alpha \neq 0$ la prima riga di (A, I_k), allora si ha

$$B \cdot A = \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \cdot A = \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \dots & \alpha a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & & \\ a_{k,1} & a_{k,2} & \dots & a_{k,n} \end{pmatrix} = A'.$$

Pertanto effettuare su una matrice ciascuna delle operazioni a), b) o c) equivale a moltiplicare a sinistra tale matrice per una opportuna matrice quadrata di ordine k.

Invece effettuare su A uno scambio di colonne, (cioé di righe della trasposta di A), equivale a moltiplicare a destra A per una opportuna matrice quadrata di ordine n.

Ne segue, adoperando la proprietá associativa del prodotto matriciale, che eseguire su una matrice A un numero finito di scambi di riga, scambi di colonna e passi di pivot equivale a moltiplicarla a sinistra per una opportuna matrice quadrata B di ordine k e a destra per una opportuna matrice quadrata C di ordine n .

Pertanto se l'algoritmo di Gauss Jordan trasforma A in A', allora esistono opportune matrici quadrate B e C tali che $A' = B\,A\,C$

■

7 - Generalitá sui sistemi lineari

Capita spesso di dover cercare dei numeri $x_1, x_2, \ldots, x_n$ che soddisfano contemporaneamente un numero finito di equazioni di primo grado nelle incognite $x_1, x_2, \ldots, x_n$, o, come si dice in genere, un *sistema* di equazioni *lineari.* Un sistema di k equazioni lineari in n incognite si rappresenta dunque sotto la forma

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n = b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n = b_2, \\ \ldots\ldots\ldots\ldots \\ a_{k,1}x_1 + a_{k,2}x_2 + \ldots + a_{k,n}x_n = b_k. \end{cases}$$

A tale sistema resta associata la matrice con k righe ed n colonne

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k,1} & a_{k,2} & \ldots & a_{k,n} \end{pmatrix},$$

che viene detta la matrice dei **coefficienti** del sistema.
Invece il vettore colonna $\mathbf{b} = (b_1, b_2, \ldots, b_k)^T$ dicesi il **vettore dei termini noti** del sistema.

Infine dicesi **soluzione** del sistema ogni vettore $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ le cui coordinate soddisfano simultaneamente tutte le equazioni del sistema.

■

Interpretando $\mathbf{x}$ e $\mathbf{b}$ rispettivamente come matrici di tipo $n \times 1$ e $k \times 1$, allora $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ é soluzione del sistema se e solo se risulta $A\mathbf{x} = \mathbf{b}$. Pertanto il sistema puó essere rappresentato sotto forma di equazione matriciale

$$A\mathbf{x} = \mathbf{b}.$$

Definizione 7.1. - *Il sistema* $A\mathbf{x} = \mathbf{b}$ *si dice* **compatibile** *se ha almeno una soluzione, si dice* **incompatibile** *se non ha nessuna soluzione. Un sistema compatibile si dice* **determinato** *se ha una sola soluzione, si dice* **indeterminato** *se ha almeno due soluzioni.*

Il sistema si dice **omogeneo** *se il vettore* $\mathbf{b}$ *dei termini noti é il vettore nullo, si dice* **non omogeneo** *se é* $\mathbf{b} \neq \mathbf{0}$.
Infine due sistemi lineari di k equazioni in n incognite si dicono equivalenti se hanno le stesse soluzioni.

■

Osservazione 7.2 - Se il sistema $A\mathbf{x} = \mathbf{b}$ é indeterminato, allora esso ha infinite soluzioni.
Infatti, se $\mathbf{x}'$ e $\mathbf{x}''$ sono due distinte soluzioni del sistema $A\mathbf{x} = \mathbf{b}$, allora per ogni $\alpha \in \mathbf{R}$ si ha che anche $\alpha\mathbf{x}' + (1-\alpha)\mathbf{x}''$ é soluzione dello stesso sistema, poiché risulta:

$$A(\alpha\mathbf{x}' + (1-\alpha)\mathbf{x}'') = \alpha A\mathbf{x}' + (1-\alpha)A\mathbf{x}'' = \alpha\mathbf{b} + (1-\alpha)\mathbf{b} = \alpha\mathbf{b} + \mathbf{b} - \alpha\mathbf{b} = \mathbf{b}.$$

■

Osservazione 7.3 - E' evidente inoltre che un sistema omogeneo é sempre compatibile perché il vettore nullo é sempre una sua soluzione; per i sistemi omogenei il problema significativo é dunque cercare le eventuali soluzioni non nulle.

Pertanto i problemi che dobbiamo affrontare nello studio dei sistemi lineari sono i seguenti:

(1) sotto quali condizioni un sistema lineare omogeneo ha solo la soluzione banale $\mathbf{x} = \mathbf{0}$?
(2) sotto quali condizioni un sistema nonomogeneo é compatibile o incompatibile?
(3) se il sistema é compatibile, sotto quali condizioni é determinato o indeterminato?
(4) se il sistema é indeterminato, qual'é la struttura dell'insieme delle sue soluzioni?
(5) come si fa a trovare la soluzione o le soluzioni del sistema?

■

Osservazione 7.4 - Affronteremo piú avanti questi problemi; per il momento ci limitiamo ad osservare che: il sistema $A\mathbf{x} = \mathbf{b}$ viene trasformato in un sistema equivalente se

- si scambiano di posto due equazioni,
- si moltiplicano i coefficienti e il termine noto di un'equazione per una costante diversa da 0,
- si aggiunge ad una equazione il multiplo di un' altra equazione,

cioé se nella matrice $(A, \mathbf{b})$ dei coefficienti e termini noti del sistema

- si scambiano di posto due righe,
- si moltiplica o divide una riga per una costante diversa da 0,
- si aggiunge ad una riga il multiplo di un' altra,

e quindi se in $(A, \mathbf{b})$ si effettuano scambi di righe e passi di pivot parziali o completi.

D' altra parte possiamo scambiare di posto le colonne di A perché questo corrisponde a scambiare di posto le incognite $x_1, x_2, \ldots, x_n$, (purché naturalmente memorizziamo tali scambi).

Pertanto, a meno di uno scambio di posizione delle incognite, il sistema $A\mathbf{x} = \mathbf{b}$ viene trasformato in un sistema equivalente $A'\mathbf{x} = \mathbf{b}'$ se la matrice $(A', \mathbf{b}')$ é stata ottenuta da $(A, \mathbf{b})$ eseguendo l'algoritmo di Gauss Jordan descritto nella Proposizione 6.1 o l'algoritmo di Gauss Jordan completo descritto nella Osservazione 6.2, senza peró mai scambiare di posto la colonna dei termini noti.

In particolare, se il vettore dei termini noti é il vettore nullo, allora, durante l'esecuzione dell'algoritmo di Gauss Jordan sulla matrice $(A, \mathbf{0})$, l'ultima colonna rimarrá sempre il vettore nullo e quindi il sistema omogeneo $A\mathbf{x} = \mathbf{0}$ viene trasformato nel sistema omogeneo $A'\mathbf{x} = \mathbf{0}$ se A' é stata ottenuta da A mediante l'algoritmo di Gauss Jordan o l'algoritmo di Gauss Jordan completo.

■

E S E R C I Z I

n. 1 - Dati i vettori $\mathbf{u}_1 = (1, 1, -1, 0)$, $\mathbf{u}_2 = (-1, 3, 0, 2)$ e $\mathbf{u}_3 = (2, 1, 3, 3)$,

(1) trovare i vettori $\mathbf{u}_1 + \mathbf{u}_2$, $\mathbf{u}_1 + \mathbf{u}_3$, $\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$;
(2) trovare la combinazione lineare dei vettori $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ mediante i coefficienti $1/2, -2, -3/5$;
(3) trovare il prodotto scalare dei vettori $\mathbf{u}_1$ ed $\mathbf{u}_2$ e dei vettori $\mathbf{u}_2$ ed $\mathbf{u}_3$.

n. 2 - Date le matrici

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

trovare le matrici trasposte di A e di B, le matrici $-1/2 \cdot A$ e $3 \cdot B$, la matrice somma $A + B$, le matrici prodotto $C = AB^T$ e $D = A^T B$.

n. 3 - Date le matrici

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

trovare le matrici trasposte di A e di B, le matrici $2 \cdot A$ e $1/3 \cdot B$, la matrice somma $A + B$, le matrici prodotto $C = AB^T$ e $D = BA^T$.

n. 4 - Eseguire l'algoritmo di Gauss Jordan sulle matrici A, B, C e D dei precedenti esercizi n. 2 e 3.

Capitolo II - DETERMINANTI

In questo capitolo studieremo una classe particolarmente interessante di matrici: le matrici quadrate. Per tali matrici, introdurremo una delle nozioni principali dell' Algebra lineare, la nozione di **determinante**.

1. - Definizione e prime Proprietá

Definizione 1.1 -. *Dicesi* **determinante** *di una matrice quadrata A di ordine n, (in simboli* $\det(A)$ *), il numero reale ottenuto secondo la seguente procedura:*

(1) *si prende una matrice quadrata A' ottenuta da A scambiando di posto le colonne, si calcola il prodotto degli elementi diagonali di A' e si prende tale prodotto con il proprio segno o con il segno cambiato a seconda che A' sia stata ottenuta da A con un numero pari o dispari di scambi di colonne,*

(2) *si ripete il passo (1) per tutte le matrici quadrate A' che é possibile ottenere da A scambiando di posto le colonne, (compresa quella che é ottenuta con 0 scambi, cioé A stessa),*

(3) *si calcola la somma di tutti i numeri ottenuti con il procedimento descritto nel passo (1).*

In altri termini il determinante di A é la somma degli $n!$ numeri reali del tipo

$$(-1)^k a_{1,j_1} \cdot a_{2,j_2} \cdot \ldots \cdot a_{n,j_n},$$

dove $(j_1, j_2, \ldots, j_n)$ é una permutazione dell'insieme $\{1, 2, \ldots, n\}$ e k é il numero degli scambi di indici necessari per ottenere la permutazione $(j_1, j_2, \ldots, j_n)$ dalla permutazione naturale $(1, 2, \ldots, n)$.

Definizione 1.2 -. *Si dice che A é una matrice* **singolare** *o* **non singolare** *a seconda che risulti* $\det(A) = 0$ *o* $\det(A) \neq 0$.

■

Esempio 1 - Se $n = 1$, l'unica matrice che si puó ottenere da $A = (a_{1,1})$ scambiando le colonne é la stessa A e richiede 0 scambi; pertanto si ha

$$\det(a_{1,1}) = a_{1,1} \qquad \text{per ogni matrice quadrata } (a_{1,1}) \text{ di ordine 1.}$$

■

Esempio 2 - Se $n = 2$, le uniche due matrici che é possibile ottenere da A scambiando le colonne sono

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \quad \text{(con 0 scambi)} \qquad \text{e} \qquad \begin{pmatrix} a_{1,2} & a_{1,1} \\ a_{2,2} & a_{2,1} \end{pmatrix} \quad \text{(con 1 scambio).}$$

Ne segue che per ogni matrice A di ordine 2 risulta

$$\det A = \det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Ad esempio si ha

$$\det \begin{pmatrix} 3 & -5 \\ 2 & -4 \end{pmatrix} = 3(-4) - 2(-5) = -2, \qquad \det \begin{pmatrix} -2 & 4 \\ 1/2 & -3 \end{pmatrix} = (-2)(-3) - 4(1/2) = 4.$$

■

Esempio 3 - Se $n = 3$, allora le $3! = 6$ matrici che é possibile ottenere da A scambiando le colonne sono

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \qquad \text{(con 0 scambi)}$$

$$\begin{pmatrix} a_{1,2} & a_{1,1} & a_{1,3} \\ a_{2,2} & a_{2,1} & a_{2,3} \\ a_{3,2} & a_{3,1} & a_{3,3} \end{pmatrix}, \quad \begin{pmatrix} a_{1,3} & a_{1,2} & a_{1,1} \\ a_{2,3} & a_{2,2} & a_{2,1} \\ a_{3,3} & a_{3,2} & a_{3,1} \end{pmatrix}, \quad \begin{pmatrix} a_{1,1} & a_{1,3} & a_{1,2} \\ a_{2,1} & a_{2,3} & a_{2,2} \\ a_{3,1} & a_{3,3} & a_{3,2} \end{pmatrix}, \qquad \text{(con 1 scambio)}$$

$$\begin{pmatrix} a_{1,2} & a_{1,3} & a_{1,1} \\ a_{2,2} & a_{2,3} & a_{2,1} \\ a_{3,2} & a_{3,3} & a_{3,1} \end{pmatrix}, \quad \begin{pmatrix} a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,3} & a_{3,1} & a_{3,2} \end{pmatrix}, \qquad \text{(con 2 scambi)}\ .$$

Ne segue che per ogni matrice A di ordine 3 risulta

$$\begin{aligned} \det A = \det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} &= \\ = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2}&. \end{aligned}$$

Si é cosí ottenuta la cosiddetta **Regola di Sarrus** per calcolare il determinante di una matrice quadrata A di ordine 3:

(1) si aggiungono ad A le prime due colonne,

(2) il determinante di A é dato dalla somma dei prodotti degli elementi delle diagonali *discendenti* meno la somma dei prodotti degli elementi delle diagonali *ascendenti.*

$$\det A = \left(\begin{array}{cccccc} a_{1,1} & & a_{1,2} & & a_{1,3} & \\ & \searrow & & \times & & \times \\ a_{2,1} & & a_{2,2} & & a_{2,3} & \\ & \nearrow & & \times & & \times \\ a_{3,1} & & a_{3,2} & & a_{3,3} & \end{array}\right) \begin{array}{ccc} a_{1,1} & & a_{1,2} \\ & \nearrow & \\ a_{2,1} & & a_{2,2} \\ & \searrow & \\ a_{3,1} & & a_{3,2} \end{array} \quad =$$

$$= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{3,1}a_{2,2}a_{1,3} - a_{3,2}a_{2,3}a_{1,1} - a_{3,3}a_{2,1}a_{1,2}.$$

■

Ad esempio si ha

$$\det \begin{pmatrix} 2 & 3 & -1 \\ 4 & 2 & 2 \\ 6 & 1 & 3 \end{pmatrix} = \left(\begin{array}{cccccc} 2 & & 3 & & -1 & \\ & \searrow & & \times & & \times \\ 4 & & 2 & & 2 & \\ & \nearrow & & \times & & \times \\ 6 & & 1 & & 3 & \end{array}\right) \begin{array}{ccc} 2 & & 3 \\ & \nearrow & \\ 4 & & 2 \\ & \searrow & \\ 6 & & 1 \end{array}$$

$$= 2 \cdot 2 \cdot 3 + 3 \cdot 2 \cdot 6 + (-1) \cdot 4 \cdot 1 - 6 \cdot 2 \cdot (-1) - 1 \cdot 2 \cdot 2 - 3 \cdot 4 \cdot 3 = 16.$$

■

Per matrici di ordine $n > 3$, il calcolo del determinante mediante la definizione risulta estremamente complesso, poiché il numero $n!$ delle permutazioni delle colonne cresce rapidamente e diventa sempre piú laborioso contare il numero degli scambi di colonne effettuati in ognuna delle $n!$ matrici A'.

Fortunatamente, esistono altri metodi che consentono il calcolo del determinante senza dover ricorrere alla definizione. A tal fine é opportuno premettere il seguente

Teorema 1.3 - (Proprietá dei determinanti).

(D1) *Il determinante di A coincide con il determinante della trasposta A^T di A.*

(D2) *Se A ha una riga o una colonna nulla, allora si ha* $\det(A) = 0$.

(D3) *Se A é una matrice triangolare superiore o inferiore, allora il determinante di A é dato dal prodotto degli elementi diagonali. In particolare si ha che* $\det(I_n) = 1$.

(D4) *Se B é ottenuta da A scambiando di posto due righe o due colonne, allora si ha* $\det(B) = -\det(A)$.

(D5) *Se due righe o due colonne di A coincidono, allora si ha* $\det(A) = 0$.

(D6) (Omogeneitá in ogni riga o colonna) - *Se la matrice B é ottenuta da A moltiplicando una qualunque riga o colonna di A per una costante α, allora si ha:* $\det(B) = \alpha \cdot \det(A)$.

(D7) (Additivitá in ogni riga e in ogni colonna) - *Se A, B e C sono tre matrici che differiscono solo nella h-esima riga o colonna, e tale h-esima riga (o colonna) di A coincide con la somma delle h-esime righe (o colonne) di B e C, allora si ha che* $\det(A) = \det(B) + \det(C)$.

(D8) *Se la matrice B é ottenuta da A aggiungendo ad una riga (o colonna) un multiplo di un' altra riga, (rispett. colonna), allora si ha* $\det(B) = \det(A)$.

(D9) *Se la matrice B é ottenuta da A mediante un passo di pivot parziale, allora si ha* $\det(B) = \det(A)$.

■

Dim. di (D1). La tesi é ovvia perché gli addendi presenti nella definizione del determinante di A e di A^T sono esattamente gli stessi e con lo stesso segno; infatti, per ogni permutazione $(j_1, j_2, j_3, \ldots, j_n)$ dell'insieme $\{1, 2, \ldots, n\}$, detta $(i_1, i_2, i_3, \ldots, i_n)$ la permutazione inversa di $(j_1, j_2, j_3, \ldots, j_n)$, cioé la permutazione definita dalla relazione

$$j_r = s \Longleftrightarrow i_s = r \quad \text{per ogni } r, s = 1, 2, \ldots, n,$$

si ha che le due permutazioni si ottengono dalla permutazione naturale con lo stesso numero di scambi e risulta

$$a_{1,j_1} \cdot a_{2,j_2} \cdot \ldots \cdot a_{n,j_n} = a_{i_1,1} \cdot a_{i_2,2} \cdot \ldots \cdot a_{i_n,n}.$$

Dim. di (D2). La tesi discende dal fatto che tutti gli addendi presenti nella definizione del determinante di A sono nulli, poiché sono il prodotto di n fattori, uno dei quali é nullo perché é stato scelto nella stato scelto sulla riga (o colonna) nulla.

Dim. di (D3). Se A é triangolare superiore, (rispett. inferiore), ed A' é una qualunque matrice ottenuta da A con almeno uno scambio di colonna, allora A' avrá sulla diagonale almeno un elemento che prima era sotto (rispett. sopra) la diagonale; pertanto uno almeno degli elementi diagonali di A' é nullo e quindi il loro prodotto é nullo. Ne segue che il $\det(A)$ si riduce al prodotto degli elementi diagonali della matrice che si ottiene da A con 0 scambi, cioé al prodotto degli elementi diagonali di A.

Dim. di (D4). Se ad esempio B é ottenuta da A scambiando la prima con la seconda riga, allora ogni addendo del $\det(B)$ coincide con l'opposto di uno degli addendi del $\det(A)$ e viceversa.

Infatti, se $(j_1, j_2, j_3, \ldots, j_n)$ é una permutazione di $\{1, 2, \ldots, n\}$ ottenuta con k scambi, allora si ha che

$$b_{1,j_1} \cdot b_{2,j_2} \cdot \ldots \cdot b_{n,j_n} = a_{2,j_1} \cdot a_{1,j_2} \cdot \ldots \cdot a_{n,j_n} = a_{1,j_2} \cdot a_{2,j_1} \cdot \ldots \cdot a_{n,j_n},$$

$$a_{1,j_1} \cdot a_{2,j_2} \cdot \ldots \cdot a_{n,j_n} = b_{2,j_1} \cdot b_{1,j_2} \cdot \ldots \cdot b_{n,j_n} = b_{1,j_2} \cdot b_{2,j_1} \cdot \ldots \cdot b_{n,j_n},$$

e la permutazione $(j_2, j_1, j_3, \ldots, j_n)$ é stata ottenuta dalla permutazione $\{1, 2, \ldots, n\}$ con $k + 1$ scambi.

Dim. di (D5). Detta B la matrice ottenuta scambiando le due righe o colonne uguali, per la (D4) deve aversi $\det(B) = -\det(A)$; d' altra parte si ha che $B = A$ e quindi $\det(B) = \det(A)$. Ne segue che $\det(A) = -\det(A)$ e dunque che $\det(A) = 0$.

Dim. di (D6). La tesi é ovvia perché ogni addendo presente nella definizione di $\det(B)$ é uguale al prodotto di α per il corrispondente addendo presente nella definizione di $\det(A)$.

Dim. di (D7). La tesi é ovvia perché ogni addendo presente nella definizione di $\det(A)$ é uguale alla somma dei corrispondenti addendi presenti nella definizione del determinante di B e di C.

Dim. di (D8). Consideriamo ad esempio la matrice B ottenuta da A sostituendo la prima riga A_1 con $A_1 + \alpha A_2$; il ragionamento é simile nel caso generale.
Si ha allora $\det(B) = \det(A_1 + \alpha A_2, A_2, \ldots\ldots, A_n) = \det(A_1, A_2, \ldots, A_n) + \alpha \det(A_2, A_2, \ldots, A_n)$ e l'ultimo determinante é nullo perché le prime due righe coincidono.

Dim. di (D9). La tesi é ovvia, perché un passo di pivot parziale consiste in un numero finito di sostituzioni delle righe di A con le stesse righe piú un multiplo della riga che contiene il pivot, e queste operazioni trasformano una matrice quadrata in una matrice con lo stesso determinante.

■

2. - Algoritmo di Gauss Jordan per il calcolo del determinante di una matrice

Osservazione 2.1 - Le precedenti proprietá (D3), (D4) e (D9) suggeriscono una strategia molto comoda per calcolare il determinante di una matrice quadrata A.
Infatti, eseguendo su A l'algoritmo di Gauss Jordan, ad ogni passo di pivot parziale si passa da una matrice ad una matrice con lo stesso determinante e ad ogni scambio di riga e/o colonna il determinante cambia segno: ebbene, possono presentarsi le seguenti alternative:

a) durante l'esecuzione dell'algoritmo di Gauss Jordan si rende necessario uno scambio di colonne;
b) con un numero finito di passi di pivot parziali ed eventualmente un numero h di scambi di righe, la matrice A viene trasformata in una matrice triangolare superiore A'.

Nel caso a), sappiamo che, con un numero finito di passi di pivot parziali ed un numero finito h di scambi di righe e/o colonne, l'algoritmo di Gauss Jordan trasforma A in una matrice del tipo $A' = \begin{pmatrix} B & C \\ O & O \end{pmatrix}$.
Ne segue che $\det(A) = (-1)^h \cdot \det(A') = 0$, poiché A' ha almeno una riga nulla .

Nel caso b) il $\det(A)$ sará dato semplicemente da $(-1)^h$ per il determinante di A', e quindi da $(-1)^h$ per il prodotto degli elementi diagonali di A'.

■

Esempio 1 - Per calcolare il determinante della matrice

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & -1 & -2 & -1 \\ -1 & 8 & 0 & 3 \\ 0 & 5 & 4 & 0 \end{pmatrix},$$

eseguiamo un passo di pivot parziale sull' elemento $a_{1,1} = 1$, cioé sostituiamo la riga A_2 con $A_2 - 2A_1$ e la riga A_3 con $A_3 + A_1$. La matrice A viene trasformata in

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -5 & 0 & -1 \\ 0 & 10 & -1 & 3 \\ 0 & 5 & 4 & 0 \end{pmatrix}.$$

Eseguiamo un secondo passo di pivot parziale su $a_{2,2} = -5$, cioé sostituiamo la riga A_3 con $A_3 + 2A_2$ e la riga A_4 con $A_4 + A_2$; si ottiene la matrice

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -5 & 0 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 4 & -1 \end{pmatrix}.$$

Eseguiamo infine un terzo passo di pivot parziale su $a_{3,3} = -1$, cioé sostituiamo la riga A_4 con $A_4 + 4A_3$; si ottiene la matrice triangolare superiore

$$A' = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -5 & 0 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Dal momento che A' é stata ottenuta da A senza scambi di righe o colonne, si ha:

$$\det A = \det(A') = 1 \cdot (-5) \cdot (-1) \cdot 3 = 15.$$

■

Esempio 2 - Per calcolare il determinante della matrice

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 0 & 4 \\ -1 & -2 & 1 & -1 \\ 3 & 6 & 2 & 3 \end{pmatrix},$$

eseguiamo un passo di pivot parziale sull' elemento $a_{1,1} = 1$, cioé sostituiamo la riga A_2 con $A_2 - 2A_1$, la riga A_3 con $A_3 + A_1$ e la riga A_4 con $A_4 - 3A_1$. La matrice A viene trasformata allora in

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 5 & -6 \end{pmatrix}.$$

A questo punto, essendo $a_{2,2} = a_{3,2} = a_{4,2} = 0$, si renderebbe necessario uno scambio di colonne; pertanto possiamo interrompere l'esecuzione dell'algoritmo e concludere che risulta $\det(A) = 0$.

In effetti, l'algoritmo di Gauss Jordan procederebbe scambiando ad esempio la seconda con la quarta colonna ed effettuando successivamente i necessari passi di pivot. La matrice A viene quindi via via trasformata nelle matrici

$$\begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & -2 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & -6 & 5 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dunque, con uno scambio di colonne e tre passi di pivot parziali, A é stata trasformata in una matrice A' con una riga nulla; ne segue che $\det(A) = -\det(A') = 0$.

■

Osservazione 2.2 - L' algoritmo di Gauss-Jordan descritto sopra consente di calcolare in maniera semplice ed efficiente il determinante di una matrice quadrata di qualunque ordine, senza dover ricorrere alla definizione.

Inoltre, per quanto visto nella precedente Osservazione 2.1, una matrice quadrata A é non singolare se e solo se, (senza scambi di colonna, ma solo con un numero finito di scambi di riga e passi di pivot parziali), l'algoritmo di Gauss Jordan trasforma A in una matrice triangolare superiore non singolare B e quindi se e solo se l'algoritmo di Gauss Jordan completo trasforma A nella matrice identica I_n.

■

Segnaliamo infine la seguente proprietá dei determinanti di cui omettiamo per semplicitá la dimostrazione.

Teorema 2.3 (di Binet) -. *Se A e B sono matrici quadrate di ordine n, allora si ha:*

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

Pertanto il prodotto di due o piú matrici quadrate non singolari é una matrice quadrata non singolare.

■

3. - Regola di Laplace

In alternativa all'algoritmo di Gauss Jordan, il determinante di una matrice quadrata puó essere calcolato con un procedimento solitamente attribuito a Laplace. A tal fine é opportuno premettere la seguente

Definizione 3.1. *Sia A una matrice quadrata di ordine $n > 1$ e, per ogni $i, j = 1, 2, \dots, n$, indichiamo con il simbolo $A_{i,j}$ la matrice quadrata di ordine $n-1$ ottenuta eliminando da A la riga $i-$esima e la colonna $j-$esima. Allora*

$\det A_{i,j}$ *dicesi* **minore complementare** *dell' elemento $a_{i,j}$ di A*;
$(-1)^{i+j} \det A_{i,j}$ *dicesi* **complemento algebrico** *di $a_{i,j}$.*

Si ha allora il seguente fondamentale

Teorema 3.2 (di Laplace). *Se A é una matrice quadrata di ordine $n > 1$, allora il determinante di A é uguale alla somma dei prodotti degli elementi di una qualunque riga o colonna per i rispettivi complementi algebrici.*

In altri termini per ogni $h = 1, 2, \dots, n$ si ha

$$\det A = \sum_{j=1}^{n} (-1)^{h+j} a_{h,j} \cdot \det A_{h,j} = \sum_{i=1}^{n} (-1)^{i+h} a_{i,h} \cdot \det A_{i,h}.$$

Dim. Dimostreremo la tesi per $n = 3$ prendendo come base di calcolo la prima riga; il caso generale si tratta in modo simile.

Infatti, dette E_1, E_2, E_3 le righe della matrice identica, essendo $A_1 = a_{1,1}E_1 + a_{1,2}E_2 + a_{1,3}E_3$, per le proprietá (D6-D7) si ha

$$\det A = a_{1,1} \det \begin{pmatrix} 1 & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} + a_{1,2} \det \begin{pmatrix} 0 & 1 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} + a_{1,3} \det \begin{pmatrix} 0 & 0 & 1 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}.$$

Effettuando un passo di pivot parziale sull' elemento non nullo della prima riga di ciascuna matrice, si ha

$$\det A = a_{1,1} \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,2} & a_{2,3} \\ 0 & a_{3,2} & a_{3,3} \end{pmatrix} + a_{1,2} \det \begin{pmatrix} 0 & 1 & 0 \\ a_{2,1} & 0 & a_{2,3} \\ a_{3,1} & 0 & a_{3,3} \end{pmatrix} + a_{1,3} \det \begin{pmatrix} 0 & 0 & 1 \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & 0. \end{pmatrix}$$

Ora, scambiando la prima con la seconda colonna, si ha

$$\det \begin{pmatrix} 0 & 1 & 0 \\ a_{2,1} & 0 & a_{2,3} \\ a_{3,1} & 0 & a_{3,3} \end{pmatrix} = -\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,1} & a_{2,3} \\ 0 & a_{3,1} & a_{3,3} \end{pmatrix}.$$

Invece, con due successivi scambi di colonna, si ha:

$$\det \begin{pmatrix} 0 & 0 & 1 \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & 0 \end{pmatrix} = -\det \begin{pmatrix} 0 & 1 & 0 \\ a_{2,1} & 0 & a_{2,2} \\ a_{3,1} & 0 & a_{3,2} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,1} & a_{2,2} \\ 0 & a_{3,1} & a_{3,2} \end{pmatrix}$$

Infine, eseguire l' algoritmo di Gauss-Jordan sulle matrici

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,2} & a_{2,3} \\ 0 & a_{3,2} & a_{3,3} \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,1} & a_{2,3} \\ 0 & a_{3,1} & a_{3,3} \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{2,1} & a_{2,2} \\ 0 & a_{3,1} & a_{3,2} \end{pmatrix}$$

equivale ad eseguirlo rispettivamente sulle matrici $A_{1,1}$, $A_{1,2}$ ed $A_{1,3}$, e quindi il loro determinante é uguale rispettivamente a $\det(A_{1,1}),\ \det(A_{1,2}),\ \det(A_{1,3})$.
Pertanto si ha $\det(A) = a_{1,1}\det(A_{1,1}) - a_{1,2}\det(A_{1,2}) + a_{1,3}\det(A_{1,3})$, come volevasi dimostrare.

■

Osservazione 3.3 - La regola di Laplace é utile poiché riduce il calcolo del determinante di una matrice di ordine n al calcolo di n determinanti di matrici di ordine $n-1$. Ad esempio si ha:

$$\begin{aligned}
\det\begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -1 \\ -1 & 2 & 3 \end{pmatrix} &= 1\cdot\det\begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix} - 2\cdot\det\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} + (-3)\cdot\det\begin{pmatrix} 2 & 5 \\ -1 & 2 \end{pmatrix} = \\
&= 1(5\cdot 3 - (-1)\cdot 2) - 2\cdot(2\cdot 3 - (-1)\cdot(-1)) - 3\cdot(2\cdot 2 - (-1)\cdot 5) = \\
&= (15+2) - 2(6-1) - 3(4+5) = 17 - 10 - 27 = -20.
\end{aligned}$$

Con lo stesso procedimento si potranno calcolare i determinanti di matrici di ordine 4, calcolando 4 determinanti del terzo ordine, e ciascuno di questi ultimi calcolando tre determinanti del secondo ordine, ecc...
Naturalmente, conviene calcolare il determinante sviluppandolo secondo gli elementi della riga o della colonna che contiene il maggior numero di elementi nulli. A tale scopo puó essere utile, (prima di applicare la formula di Laplace), aumentare il numero degli elementi nulli di una riga o colonna aggiungendo ad essa un opportuno multiplo di un' altra riga o colonna con lo stesso spirito con cui si eseguono i passi di pivot.
Ad esempio per calcolare il determinante della matrice

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

possiamo sostituire la prima riga con la prima piú la terza riga e poi la terza colonna con la terza meno la prima colonna. Si ha dunque:

$$\begin{aligned}
\det(A) = \det\begin{pmatrix} 2 & 1 & 3 \\ 1 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix} &= \det\begin{pmatrix} 4 & 0 & 4 \\ 1 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix} = \det\begin{pmatrix} 4 & 0 & 0 \\ 1 & 4 & -2 \\ 2 & -1 & -1 \end{pmatrix} = \\
= 4\cdot\det\begin{pmatrix} 4 & -2 \\ -1 & -1 \end{pmatrix} &= 4(-4-2) = -24.
\end{aligned}$$

E' del tutto evidente che la convenienza della regola di Laplace rispetto all'algoritmo di Gauss Jordan diminuisce rapidamente con l'aumentare dell'ordine di A.

■

4. - Sistemi lineari non omogenei - Teorema di Cramer

L'algoritmo di Gauss Jordan oltre a consentirci di calcolare facilmente il determinante di una matrice quadrata ci consente anche di risolvere un sistema lineare non omogeneo di n equazioni in n incognite. Si ha infatti la seguente

Proposizione 4.1 -. *Se A é una matrice quadrata non singolare di ordine n, allora per ogni $\mathbf{b} \in \mathbf{R}^n$ il sistema lineare nonomogeneo $A\mathbf{x} = \mathbf{b}$ ha una ed una sola soluzione; tale soluzione é il vettore $\mathbf{b}'$ tale che l'algoritmo di Gauss Jordan completo trasforma la matrice $(A, \mathbf{b})$ nella matrice $(I_n, \mathbf{b}')$.*

Dim. Infatti, per quanto visto nella Osservazione 7.4 del Capitolo I, il sistema lineare non omogeneo $A\mathbf{x} = \mathbf{b}$ é equivalente al sistema $A'\mathbf{x} = \mathbf{b}'$ dove $(A', \mathbf{b}')$ é stata ottenuta da $(A, \mathbf{b})$ mediante l'algoritmo di Gauss Jordan o l'algoritmo di Gauss Jordan completo.

D'altra parte, se il determinante di A é diverso da 0, allora per quanto visto nella Osservazione 2.2, l'algoritmo di Gauss Jordan completo trasforma (senza scambi di colonna, ma solo con scambi di riga e passi di pivot) la matrice A nella matrice I_n e quindi trasforma $(A, \mathbf{b})$ in $(I_n, \mathbf{b}')$.

Pertanto il sistema $A\mathbf{x} = \mathbf{b}$ é equivalente al sistema $I_n\mathbf{x} = \mathbf{b}'$, che ha l'unica soluzione $\mathbf{x} = \mathbf{b}'$ e quindi $\mathbf{b}'$ é l'unica soluzione del sistema $A\mathbf{x} = \mathbf{b}$.

■

Esempio 1 - Consideriamo il sistema

$$\begin{cases} x_1 + 2x_2 - x_3 & = 2 \\ 2x_1 - x_2 - 2x_3 - x_4 & = -1 \\ -x_1 + 8x_2 + \quad 3x_4 & = 0 \\ \quad 5x_2 + 4x_3 & = 1 \end{cases} \qquad \text{cioé} \qquad \begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 2 & -1 & 0 & 2 \\ 2 & -1 & -2 & -1 & -1 \\ -1 & 8 & 0 & 3 & 0 \\ 0 & 5 & 4 & 0 & 1 \end{array}$$

Per quanto visto nell'esempio 1 del paragrafo 2, sappiamo che il determinante della matrice A dei coefficienti del sistema é uguale a $15 \neq 0$. Ebbene eseguendo l'algoritmo di Gauss Jordan completo sulla matrice dei coefficienti e termini noti si ha che tale sistema é equivalente ai sistemi

$$\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 2 & -1 & 0 & 2 \\ 0 & -5 & 0 & -1 & -5 \\ 0 & 10 & -1 & 3 & 2 \\ 0 & 5 & 4 & 0 & 1 \end{array} \qquad \begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 0 & -1 & -2/5 & 0 \\ 0 & 1 & 0 & 1/5 & 1 \\ 0 & 0 & -1 & 1 & -8 \\ 0 & 0 & 4 & -1 & -4 \end{array}$$

$$\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 0 & 0 & -7/5 & 8 \\ 0 & 1 & 0 & 1/5 & 1 \\ 0 & 0 & 1 & -1 & 8 \\ 0 & 0 & 0 & 3 & -36 \end{array} \qquad \begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 0 & 0 & 0 & -44/5 \\ 0 & 1 & 0 & 0 & 17/5 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 1 & -12 \end{array}$$

Il sistema ha dunque l'unica soluzione $\mathbf{x} = (-44/5, 17/5, -4, -12)^T$.

■

Il teorema seguente fornisce un metodo di risoluzione del sistema lineare $A\mathbf{x} = \mathbf{b}$ alternativo all'algoritmo di Gauss Jordan.

Teorema 4.2 - (Teorema di Cramer) -. *Se A é una matrice quadrata non singolare, allora per ogni $\mathbf{b} \in \mathbf{R}^n$ l' unica soluzione del sistema $A\mathbf{x} = \mathbf{b}$ é il vettore $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ tale che*

$$(*) \qquad x_j = \frac{\det(A^1, A^2, \ldots, A^{j-1},\ \mathbf{b},\ A^{j+1}, \ldots, A^n)}{\det A} \qquad \textit{per ogni} \qquad j = 1, 2, \ldots, n.$$

Dim. Per quanto visto nella Proposizione 4.1 sappiamo che il sistema $A\mathbf{x} = \mathbf{b}$ ha un'unica soluzione $\mathbf{x}$. D'altra parte, tale unica soluzione puó essere identificata osservando che $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ é soluzione del sistema se e solo se risulta

$$\mathbf{b} = x_1 A^1 + x_2 A^2 + \ldots + x_n A^n.$$

Ora, per le proprietá dei determinanti, si ha che

$$\begin{aligned}\det(\mathbf{b}, A^2, \ldots, A^n) &= \det(x_1 A^1 + x_2 A^2 + \ldots + x_n A^n, A^2, \ldots, A^n) = \\ &= x_1 \det(A^1, A^2, \ldots, A^n) + x_2 \det(A^2, A^2, \ldots, A^n) + \ldots + x_n \det(A^n, A^2, \ldots, A^n);\end{aligned}$$

dal momento che il determinante di una matrice in cui due colonne coincidono é nullo, se ne deduce che

$$\det(\mathbf{b}, A^2, \ldots, A^n) = x_1 \det A \qquad \text{e quindi che} \qquad x_1 = \det(\mathbf{b}, A^2, \ldots, A^n)/\det A.$$

Analogamente, per ogni $j = 2, 3, \ldots, n$, si ha

$$\det(A^1, A^2, \ldots, A^{j-1},\ \mathbf{b},\ A^{j+1}, \ldots, A^n) = x_j \det A$$

e quindi che

$$x_j = \det(A^1, \ldots A^{j-1},\ \mathbf{b},\ A^{j+1}, \ldots, A^n)/\det A,$$

e questo prova la tesi.

■

Esempio 2 - Risolvere il sistema

$$\begin{cases} x + y + z &= 1 \\ x - y + z &= 0 \\ -x - 3y + z &= 2. \end{cases}$$

Dobbiamo calcolare i determinanti

$$D = \det A = \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & -3 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & -2 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = -4;$$

$$D_x = \det(\mathbf{b}, A^2, A^3) = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 2 & -3 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & -5 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & -6 \end{pmatrix} = 6,$$

$$D_y = \det(A^1, \mathbf{b}, A^3) = \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 3 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = -2,$$

$$D_z = \det(A^1, A^2, \mathbf{b}) = \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & -3 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & -2 & 3 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 4 \end{pmatrix} = -8.$$

L'unica soluzione del sistema é dunque il vettore di coordinate

$$x = D_x/D = -3/2, \quad y = D_y/D = 1/2, \quad z = D_z/D = 2.$$

■

Osservazione 4.3 - Il metodo di Cramer richiede il calcolo di $n+1$ determinanti di ordine n e quindi é alquanto oneroso se l'ordine n della matrice A non é piccolissimo.

Giá l' esercizio precedente che richiede il calcolo di quattro determinanti del terzo ordine poteva essere svolto piú facilmente utilizzando l' algoritmo di Gauss Jordan.

Infatti, con tre successivi passi di pivot sulla matrice dei coefficienti e termini noti, il sistema verrebbe via via trasformato nei sistemi equivalenti

$$\begin{array}{rrr|r} x & y & z & \\ \hline 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \\ -1 & -3 & 1 & 2 \end{array} \qquad \begin{array}{rrr|r} x & y & z & \\ \hline 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & -2 & 2 & 3 \end{array}$$

$$\begin{array}{rrr|r} x & y & z & \\ \hline 1 & 0 & 1 & 1/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 2 & 4. \end{array} \qquad \begin{array}{rrr|r} x & y & z & \\ \hline 1 & 0 & 0 & -3/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 2. \end{array}$$

e quindi ha l' unica soluzione $(x, y, z) = (-3/2,\ 1/2,\ 2)$.

■

5. - Matrici Quadrate Invertibili - Inversa di una Matrice Invertibile

Per le matrici quadrate di ordine n é possibile introdurre una ulteriore nozione, la nozione di invertibilitá nel senso della seguente

Definizione 5.1 -. *Se A é una matrice quadrata di ordine n, si dice che A é* **invertibile** *se esiste una matrice quadrata B di ordine n tale che* $BA = AB = I_n$.

Una siffatta matrice é unica, dicesi la **matrice inversa** *di A e si denota con il simbolo A^{-1}.*

Nota bene: l'unicitá deriva dal fatto che se B', B'' sono due matrici tali che $B'A = I_n$ ed $AB'' = I_n$, allora si ha

$$B' = B'I_n = B'(AB'') = (B'A)B'' = I_nB'' = B''.$$

■

E' facile rendersi conto che esistono matrici quadrate non nulle che non sono invertibili; ad esempio la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

non puó essere invertibile; infatti per ogni matrice B le righe di $A \cdot B$ sono coincidenti e quindi non possono mai coincidere con le righe della matrice identica.

Sorgono allora spontanee le domande:

Sotto quali condizioni una matrice quadrata A di ordine n é invertibile?

Se A é invertibile come si fa a trovarne l'inversa ?

A queste domande risponde la seguente

Proposizione 5.2 -. *Se A é una matrice quadrata di ordine n, allora le seguenti proprietá sono equivalenti:*

a) *A é invertibile,*
b) *il determinante di A é diverso da 0,*
c) *il determinante di A^T é diverso da 0,*
d) *A^T é invertibile.*

Dim. di a) $\implies$ b). Infatti, dal teorema di Binet segue che

$$\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(I_n) = 1 \neq 0 \qquad \text{e quindi che} \qquad \det(A) \neq 0.$$

Dim. di b) $\implies$ a). Per la Osservazione 2.2 si ha $\det(A) \neq 0$ se e solo se l'algoritmo di Gauss Jordan completo trasforma (senza scambi di colonna) A in I_n e quindi trasforma la matrice (A, I_n) in una matrice del tipo (I_n, B). Vogliamo dimostrare che A é invertibile e B é la sua inversa, cioé che risulta

$$A \cdot B = I_n \qquad \text{e} \qquad B \cdot A = I_n.$$

Infatti, (dette $E^1, E^2, \ldots, E^n$ le colonne della matrice identica I_n), se l'algoritmo di Gauss Jordan completo trasforma (A, I_n) in (I_n, B), allora esso trasforma le matrici $(A, E^1), (A, E^2), \ldots, (A, E^n)$ rispettivamente nelle matrici $(I_n, B^1), (I_n, B^2), \ldots, (I_n, B^n)$. Per la Proposizione 4.1, questo vuol dire che le colonne $B^1, B^2, \ldots B^n$ di B sono le uniche soluzioni dei sistemi $A\mathbf{x} = E^1$, $A\mathbf{x} = E^2$, $\ldots$ $A\mathbf{x} = E^n$.
Si ha dunque che $AB^1 = E^1$, $AB^2 = E^2$, $\ldots AB^n = E^n$ e quindi che $A \cdot B = I_n$.

D'altra parte, per quanto visto nella Osservazione 6.4 del Capitolo I, se (A, I_n) viene trasformata in (I_n, B) mediante un numero finito di scambi di riga e passi di pivot, allora esiste una opportuna matrice quadrata C tale che $(I_n, B) = C \cdot (A,\ I_n) = (C \cdot A,\ C \cdot I_n) = (C \cdot A, C)$.
Se ne deduce che $C \cdot A = I_n$ e $C = B$ e dunque che $B \cdot A = C \cdot A = I_n$ e questo prova completamente la tesi.

Dim. di b) $\iff$ c) $\iff$ d). L'equivalenza di b) e c) deriva dal fatto che $\det(A^T) = \det(A)$; infine l'equivalenza di c) e d) deriva dall'equivalenza di a) e b).

■

Osservazione 5.3 - La dimostrazione della proposizione precedente contiene un procedimento molto semplice ed efficiente per vedere se una matrice quadrata é invertibile o non invertibile e per costruirne l'eventuale inversa.

Si affianca ad A la matrice identica I_n e sulla matrice (A, I_n) si esegue l'algoritmo di Gauss Jordan completo:

- *se si rende necessario uno scambio di colonna o se A viene trasformata in una matrice con una riga nulla, allora si ha $\det(A) = 0$ e quindi A non é invertibile;*
- *se l'algoritmo di Gauss Jordan completo trasforma (A, I_n) in (I_n, B), allora A é invertibile e B é la sua inversa.*

■

Esempio 1 - Dire se é invertibile e trovare l' eventuale inversa della matrice

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

Consideriamo la matrice

$$\hat{A} = (A, I_n) = \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & -2 & 0 & 1 & 0 \\ 4 & -1 & -3 & 0 & 0 & 1 \end{array}\right)$$

ed effettuiamo un passo di pivot sull' elemento $\hat{a}_{1,1} = 1$, cioé dividiamo per 1 la prima riga A'_1 e sostituiamo la riga $\hat{A}_2$ con $\hat{A}_2 + \hat{A}_1$ e la riga $\hat{A}_3$ con $\hat{A}_3 - 4\hat{A}_1$: otteniamo la matrice

$$\left(\begin{array}{rrrcrrr} 1 & 2 & -1 & | & 1 & 0 & 0 \\ 0 & 5 & -3 & | & 1 & 1 & 0 \\ 0 & -9 & 1 & | & -4 & 0 & 1 \end{array}\right).$$

Effettuiamo ora un passo di pivot sull' elemento $\hat{a}_{2,2} = 5$, cioé dividiamo per 5 la seconda riga e sostituiamo la riga $\hat{A}_1$ con $\hat{A}_1 - (2/5)\hat{A}_2$ e la riga $\hat{A}_3$ con $\hat{A}_3 + (9/5)\hat{A}_2$: otteniamo la matrice

$$\left(\begin{array}{rrrcrrr} 1 & 0 & 1/5 & | & 3/5 & -2/5 & 0 \\ 0 & 1 & -3/5 & | & 1/5 & 1/5 & 0 \\ 0 & 0 & -22/5 & | & -11/5 & 9/5 & 1 \end{array}\right).$$

Infine effettuiamo un passo di pivot sull' elemento $\hat{a}_{3,3} = -22/5$, cioé sostituiamo la riga $\hat{A}_1$ con $\hat{A}_1+(1/22)\hat{A}_3$ e la riga $\hat{A}_2$ con $\hat{A}_2 - (15/22)\hat{A}_3$ e dividiamo per $-22/5$ la terza riga: otteniamo la matrice

$$\left(\begin{array}{rrrcrrr} 1 & 0 & 0 & | & 1/2 & -7/22 & 1/22 \\ 0 & 1 & 0 & | & 1/2 & -1/22 & -3/22 \\ 0 & 0 & 1 & | & 1/2 & -9/22 & -5/22 \end{array}\right).$$

La matrice A é dunque invertibile e la sua inversa é la matrice

$$B = \begin{pmatrix} 1/2 & -7/22 & 1/22 \\ 1/2 & -1/22 & -3/22 \\ 1/2 & -9/22 & -5/22 \end{pmatrix}.$$

■

Un metodo alternativo per calcolare l'inversa di una matrice invertibile A di ordine n é fornito dal teorema di Cramer. Si ha infatti la seguente

Proposizione 5.4. *La matrice inversa di A coincide con la matrice ottenuta dividendo per* $\det(A)$ *la matrice trasposta della matrice dei complementi algebrici degli elementi di A.*

[. Dim] Sia infatti x_{ij} il generico elemento della matrice inversa A^{-1} e siano $E^1, E^2, \dots, E^n$ le colonne della matrice identica. Allora la $j-$esima colonna della matrice inversa A^{-1} di A é la soluzione del sistema $A\mathbf{x} = E^j$; per il teorema di Cramer si ha quindi

$$x_{ij} = \frac{\det(A^1, A^2, \dots, A^{i-1}, E^j, A^{i+1}, \dots, A^n)}{\det A} \qquad \text{per ogni } \; i = 1, 2, \dots, n.$$

Ebbene, sviluppando il determinante a numeratore secondo gli elementi della $i-$esima colonna, si ha

$$x_{ij} = \frac{\text{complemento algebrico di } a_{ji}}{\det A} \qquad \text{per ogni } \; i = 1, 2, \dots, n \text{ e per ogni } j = 1, 2, \dots, n.$$

Questo dimostra la tesi.

■

Osservazione 5.5 - La matrice trasposta della matrice dei complementi algebrici degli elementi di una matrice quadrata A dicesi la **matrice aggiunta** di A.

Pertanto la Proposizione 5.4 puó essere riformulata dicendo che *la matrice inversa di una matrice quadrata invertibile é la matrice ottenuta dividendo per* $\det(A)$ *la matrice aggiunta di A.*

∎

Ad esempio nel caso della matrice

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

si ha:

$$\begin{aligned}
&\det A_{11} = \begin{pmatrix} 3 & -2 \\ -1 & -3 \end{pmatrix} = -11, \quad \det A_{12} = \begin{pmatrix} -1 & -2 \\ 4 & -3 \end{pmatrix} = 11, \quad \det A_{13} = \begin{pmatrix} -1 & 3 \\ 4 & -1 \end{pmatrix} = -11, \\
&\det A_{21} = \begin{pmatrix} 2 & -1 \\ -1 & -3 \end{pmatrix} = -7, \quad \det A_{22} = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix} = 1, \quad \det A_{23} = \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} = -9, \\
&\det A_{31} = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} = -1, \quad \det A_{32} = \begin{pmatrix} 1 & -1 \\ -1 & -2 \end{pmatrix} = -3, \quad \det A_{33} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = 5.
\end{aligned}$$

La matrice dei complementi algebrici degli elementi di A é la matrice

$$\begin{pmatrix} -11 & -11 & -11 \\ 7 & 1 & 9 \\ -1 & 3 & 5 \end{pmatrix},$$

la sua trasposta, (cioé la matrice aggiunta di A), é la matrice

$$\begin{pmatrix} -11 & 7 & -1 \\ -11 & 1 & 3 \\ -11 & 9 & 5 \end{pmatrix}.$$

Inoltre il determinante di A é

$$\det A = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} \det A_{1j} = -11 - 2 \cdot 11 - (-11) = -22.$$

L' inversa di A si ottiene allora dividendo per -22 la matrice aggiunta di A; si ha dunque

$$A^{-1} = \begin{pmatrix} 1/2 & -7/22 & 1/22 \\ 1/2 & -1/22 & -3/22 \\ 1/2 & -9/22 & -5/22 \end{pmatrix}.$$

E S E R C I Z I

n. 1 - Date le matrici

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 4 & -2 \\ -2 & -6 & 7 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 4 & -1 \\ -3 & -2 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 4 & -1 \\ -3 & -2 & -3 \end{pmatrix},$$

a) calcolare il determinante di A , di B, di C, di AB, di BC e di AC con la regola di Sarrus, con la regola di Laplace e con il metodo di Gauss-Jordan;
b) se la matrice A é non singolare, trovare l'unica soluzione del sistema $A\mathbf{x} = \mathbf{b}$ con $\mathbf{b} = (1, 3, 1)^T$, usando sia l'algoritmo di Gauss Jordan che il metodo di Cramer; ripetere l'esercizio per le matrici B, C, AB ed AC;
c) dire se le matrici A, B, C, AB, BC ed AC sono invertibili; in caso affermativo calcolarne la matrice inversa con il metodo di Gauss Jordan e con il metodo della matrice aggiunta.

n. 2 - Date le matrici

$$A = \begin{pmatrix} 1 & 2 & 0 & -3 \\ 0 & 1 & 4 & -2 \\ -2 & 0 & -6 & 7 \\ 1 & -1 & 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 2 & 3 & 1 \\ 1 & 0 & 4 & -1 \\ -3 & -2 & 0 & 1 \\ 0 & 1 & 2 & -3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix},$$

a) calcolare il determinante di A e di B usando l'algoritmo di Gauss Jordan e le formule di Laplace;
b) se tali matrici sono non singolari, trovare l'unica soluzione dei sistemi $A\mathbf{x} = \mathbf{b}$ e $B\mathbf{x} = \mathbf{b}$, usando sia l'algoritmo di Gauss Jordan che il metodo di Cramer;
c) se tali matrici sono invertibili, trovarne la matrice inversa.

n. 3 - Ripetere l'esercizio n. 2 per le matrici

$$A = \begin{pmatrix} 0 & 2 & 3 & 4 & 0 \\ 2 & 0 & 2 & -6 & 0 \\ 1 & 3 & 2 & 0 & -2 \\ 0 & -2 & 0 & 0 & 3 \\ 3 & 0 & 0 & 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & -2 & 0 & -3 & 0 \\ -1 & 2 & 0 & 0 & -6 \\ 1 & -3 & 2 & -4 & 0 \\ 0 & 2 & 0 & 0 & 3 \\ 3 & 0 & 0 & 2 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

n. 4 - Usando il teorema di Cramer, risolvere i sistemi lineari:

$$\begin{cases} x+2y-z = 0 \\ 2x+y+2z = 1 \\ 3x+5y-5z = -1 \end{cases}, \qquad \begin{cases} x-2y+z = 0 \\ 2x-2y+2z = 1 \\ -2x+4y-z = -1 \end{cases}, \qquad \begin{cases} x-2y+z+t = 0 \\ 2x-2y+2z = 1 \\ -2x+4y+t = -1 \\ y+3z = 2 \end{cases}.$$

Capitolo III - SOTTOSPAZI - BASI - DIMENSIONE

1. - Sottospazi vettoriali

Definizione 1.1 -. *Dicesi sottospazio vettoriale dello spazio vettoriale* $\mathbf{R}^n$ *ogni insieme non vuoto* $H \subseteq \mathbf{R}^n$ *tale che:*

(1) *per ogni* $\mathbf{u}, \mathbf{v} \in H$ *risulta* $\mathbf{u} + \mathbf{v} \in H$,
(2) *per ogni* $\alpha \in \mathbf{R}$ *e per ogni* $\mathbf{v} \in H$ *risulta* $\alpha \cdot \mathbf{v} \in H$.

■

Esempio 1 - E' evidente che gli insiemi $\{\mathbf{0}_n\}$ ed $\mathbf{R}^n$ sono sottospazi di $\mathbf{R}^n$.

■

Esempio 2 - Gli insiemi $\{(x_1, x_2) \in \mathbf{R}^2 | x_2 = 0\}$ e $\{(x_1, x_2) \in \mathbf{R}^2 | x_1 = 0\}$ sono sottospazi di $\mathbf{R}^2$.
Infatti, la somma di due vettori di $\mathbf{R}^2$ aventi la seconda (o la prima) coordinata nulla é un vettore che ha la seconda (o la prima) coordinata nulla; analogamente il multiplo di un vettore avente la seconda (o la prima) coordinata nulla é un vettore che ha la seconda (o la prima) coordinata nulla.

■

Osservazione 1.2 - E' del tutto evidente che se H é un sottospazio vettoriale di $\mathbf{R}^n$, allora:

(1) $\mathbf{0} = 0 \cdot \mathbf{v} \in H$,
(2) qualunque combinazione lineare finita di elementi di H é un elemento di H.

Ne segue che una parte di $\mathbf{R}^n$ che non contiene il vettore nullo non é un sottospazio vettoriale di $\mathbf{R}^n$.
Ad esempio un insieme del tipo $H = \{(x_1, x_2, x_3) \in \mathbf{R}^3 | x_1 - 2x_2 + x_3 = 2\}$ non contiene il vettore nullo $(0, 0, 0)$ e quindi non é un sottospazio di $\mathbf{R}^3$.

■

Esempio 3 - L'insieme H delle soluzioni dell'equazione $x_1 - 2x_2 + x_3 = 0$ é un sottospazio vettoriale di $\mathbf{R}^3$.
Infatti se $\alpha \in \mathbf{R}$ ed $\mathbf{x} = (x_1, x_2, x_3)$ ed $\mathbf{y} = (y_1, y_2, y_3)$ appartengono ad H, (cioé sono due punti di $\mathbf{R}^3$ tali che $x_1 - 2x_2 + x_3 = 0$ e $y_1 - 2y_2 + y_3 = 0$), allora si ha che $\mathbf{x} + \mathbf{y}$ ed $\alpha\mathbf{x}$ appartengono ad H.
Infatti si ha:

$$(x_1 + y_1) - 2(x_2 + y_2) + (x_3 + y_3) = (x_1 - 2x_2 + x_3) + (y_1 - 2y_2 + y_3) = 0 + 0 = 0,$$

$$(\alpha x_1) - 2(\alpha x_2) + (\alpha x_3) = \alpha(x_1 - 2x_2 + x_3) = \alpha \cdot 0 = 0.$$

Invece l'insieme H delle soluzioni dell'equazione $x_1 - 2x_2x_3 = 0$ non é un sottospazio vettoriale di $\mathbf{R}^3$.
Infatti se $\mathbf{x} = (x_1, x_2, x_3)$ ed $\mathbf{y} = (y_1, y_2, y_3)$ appartengono ad H, (cioé sono due punti di $\mathbf{R}^3$ tali che $x_1 - 2x_2x_3 = 0$ e $y_1 - 2y_2y_3 = 0$), allora si ha:

$$(x_1 + y_1) - 2(x_2 + y_2)(x_3 + y_3) = (x_1 - 2x_2x_3) + (y_1 - 2y_2y_3) - 2x_2y_3 - 2x_3y_2 = -2(x_2y_3 + x_3y_2),$$

e quindi non é vero che $\mathbf{x} + \mathbf{y}$ appartiene ad H.

■

Esempio 4 - Per ogni $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbf{R}^n$ l'insieme $H = \{\mathbf{x} \in \mathbf{R}^n | (\mathbf{a}|\mathbf{x}) = 0\}$, cioé l'insieme delle soluzioni dell'equazione lineare omogenea $a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$, é un sottospazio di $\mathbf{R}^n$.

Infatti, per ogni $\alpha \in \mathbf{R}$ e per ogni $\mathbf{x}, \mathbf{y} \in H$, si ha che $\alpha\mathbf{x}$ ed $\mathbf{x} + \mathbf{y}$ appartengono ad H, dal momento che risulta $(\mathbf{a}|\mathbf{x}) = 0$, $(\mathbf{a}|\mathbf{y}) = 0$, e quindi $(\mathbf{a}|\alpha\mathbf{x}) = \alpha(\mathbf{a}|\mathbf{x}) = \alpha \cdot 0 = 0$, ed $(\mathbf{a}|\mathbf{x}+\mathbf{y}) = (\mathbf{a}|\mathbf{x})+(\mathbf{a}|\mathbf{y}) = 0+0 = 0$.

■

Esempio 5 - Se A é una matrice di tipo $k \times n$, allora l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ é un sottospazio di $\mathbf{R}^n$.

Infatti, per ogni $\alpha \in \mathbf{R}$ e per ogni $\mathbf{x}, \mathbf{y} \in H$, si ha che $\alpha\mathbf{x}$ ed $\mathbf{x} + \mathbf{y}$ appartengono ad H, dal momento che risulta $A\mathbf{x} = \mathbf{0}_k$, $A\mathbf{y} = \mathbf{0}_k$, e quindi $A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha \cdot \mathbf{0}_k = \mathbf{0}_k$, ed $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = 0_k + 0_k = 0_k$.

■

Esempio 6 - Se $\mathbf{v} \in \mathbf{R}^n$, allora l'insieme

$$\mathcal{L}(\mathbf{v}) = \{\mathbf{x} \in \mathbf{R}^n \quad | \quad \exists \alpha \in \mathbf{R} \text{ tale che } \mathbf{x} = \alpha\mathbf{v}\}$$

é un sottospazio vettoriale di $\mathbf{R}^n$; esso contiene $\mathbf{v}$ ed é contenuto in ogni sottospazio che contiene $\mathbf{v}$.

Infatti, se $\mathbf{x}, \mathbf{y} \in \mathcal{L}(\mathbf{v})$ ed $\alpha \in \mathbf{R}$, allora esistono $\beta, \gamma \in \mathbf{R}$ tali che $\mathbf{x} = \beta\mathbf{v}$, $\mathbf{y} = \gamma\mathbf{v}$, e quindi:

$$\mathbf{x} + \mathbf{y} = \beta\mathbf{v} + \gamma\mathbf{v} = (\beta + \gamma)\mathbf{v} \in \mathcal{L}(\mathbf{v}), \qquad \alpha\mathbf{x} = \alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v} \in \mathcal{L}(\mathbf{v}).$$

Questo prova che $\mathcal{L}(\mathbf{v})$ é un sottospazio di $\mathbf{R}^n$.

D'altra parte $\mathcal{L}(\mathbf{v})$ contiene $\mathbf{v}$, essendo $\mathbf{v} = 1\mathbf{v} \in \mathcal{L}(\mathbf{v})$. Infine, se H é un qualunque sottospazio contenente $\mathbf{v}$, allora H contiene tutti i multipli di $\mathbf{v}$ e quindi contiene $\mathcal{L}(\mathbf{v})$.

■

Esempio 7 - Se $\mathbf{v}_1, \mathbf{v}_2$ sono due vettori di $\mathbf{R}^n$, allora l'insieme

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2) = \{\mathbf{x} \in \mathbf{R}^n | \exists \alpha_1, \alpha_2 \in \mathbf{R} \text{ tali che } \mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2\}$$

é un sottospazio vettoriale di $\mathbf{R}^n$, che contiene $\{\mathbf{v}_1, \mathbf{v}_2\}$ ed é contenuto in ogni sottospazio contenente $\{\mathbf{v}_1\mathbf{v}_2\}$.

In generale, se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono r vettori di $\mathbf{R}^n$, allora l'insieme di tutte le combinazioni lineari dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, cioé l'insieme

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r) = \{\mathbf{x} \in \mathbf{R}^n \quad | \quad \exists \alpha_1, \alpha_2, \ldots \alpha_r \in \mathbf{R} \text{ tali che } \mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r\}$$

é un sottospazio vettoriale di $\mathbf{R}^n$. Esso contiene l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ ed é contenuto in ogni sottospazio che contiene $\{\mathbf{v}_1\mathbf{v}_2, \ldots, \mathbf{v}_r\}$.

Infatti, se $\mathbf{x}, \mathbf{y} \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ ed $\alpha \in \mathbf{R}$, allora esistono $\beta_1, \beta_2, \ldots, \beta_r$, $\gamma_1, \gamma_2, \ldots, \gamma_r \in \mathbf{R}$ tali che

$$\mathbf{x} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_r\mathbf{v}_r \qquad \text{e} \qquad \mathbf{y} = \gamma_1\mathbf{v}_1 + \gamma_2\mathbf{v}_2 + \ldots + \gamma_r\mathbf{v}_r,$$

e quindi

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_r\mathbf{v}_r) + (\gamma_1\mathbf{v}_1 + \gamma_2\mathbf{v}_2 + \ldots + \gamma_r\mathbf{v}_r) = \\ &= (\beta_1 + \gamma_1)\mathbf{v}_1 + (\beta_2 + \gamma_2)\mathbf{v}_2 + \ldots + (\beta_r + \gamma_r)\mathbf{v}_r \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r), \end{aligned}$$

$$\alpha\mathbf{x} = \alpha(\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_r\mathbf{v}_r) = (\alpha\beta_1)\mathbf{v}_1 + (\alpha\beta_2)\mathbf{v}_2 + \ldots + (\alpha\beta_r)\mathbf{v}_r \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r).$$

Questo prova che $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ é un sottospazio vettoriale di $\mathbf{R}^n$.

Tale sottospazio contiene $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, poiché risulta:

$$\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_r, \quad \mathbf{v}_2 = 0\mathbf{v}_1 + 1\mathbf{v}_2 + \ldots + 0\mathbf{v}_r, \quad \ldots, \quad \mathbf{v}_r = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 1\mathbf{v}_r.$$

Infine, se H é un qualunque sottospazio di $\mathbf{R}^n$ ed H contiene i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, allora H contiene tutte le combinazioni lineari di questi vettori, e quindi contiene $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$.

■

Definizione 1.3 -. *Si dice che* $\mathcal{L}(\mathbf{v})$ *é il* **sottospazio generato** *da* $\mathbf{v}$ *o che* $\mathbf{v}$ *é un* **generatore** *di* $\mathcal{L}(\mathbf{v})$. *In generale, si dice che l'insieme* $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$ *é il* **sottospazio generato** *dai vettori* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *o che* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *sono i* **generatori** *di* $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$.

■

Osservazione 1.4 - Se $n = 2$ o $n = 3$ e $\mathbf{v} \neq \mathbf{0}$, allora lo spazio generato dal vettore $\mathbf{v}$ é l'insieme dei vettori delle coordinate dei punti di una retta, la retta congiungente l'origine con il punto P di coordinate $\mathbf{v}$.
Analogamente, se $\mathbf{v}_1$, $\mathbf{v}_2$ sono due vettori di $\mathbf{R}^3$ e P e Q sono i punti dello spazio tridimensionale di cui $\mathbf{v}_1$, $\mathbf{v}_2$ sono i vettori delle coordinate rispetto ad un assegnato riferimento cartesiano ortogonale monometrico, allora lo spazio generato dai vettori $\mathbf{v}_1, \mathbf{v}_2$ é l'insieme dei vettori delle coordinate dei punti

a) del piano passante per l'origine e per i punti P e Q se tali punti non sono allineati con l'origine,
b) della retta passante per O, P e Q se tali punti sono allineati.

■

Osservazione 1.5 -. *Se* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r,\ \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s$ *sono* $r + s$ *vettori di* $\mathbf{R}^n$ *tali che*

$$\mathbf{w}_i \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r) \qquad \textit{per ogni } i = 1, 2, \ldots, s,$$

allora si ha:

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r,\ \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s) = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r).$$

Infatti, $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r,\ \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s)$ contiene $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, e quindi contiene $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$.
Viceversa, $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$ contiene sia i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ che i vettori $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s$, e quindi contiene $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r,\ \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s)$.

■

2. Vettori linearmente dipendenti o indipendenti

Definizione 2.1 -. *Si dice che i vettori* $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ *sono linearmente dipendenti se esiste una combinazione lineare nulla di tali vettori con coefficienti non tutti nulli, cioé se esistono k coefficienti non tutti nulli,* $\alpha_1, \alpha_2, \ldots \alpha_r$, *tali che* $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r = \mathbf{0}$.
Si dice invece che i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ *sono linearmente indipendenti se non sono linearmente dipendenti e quindi se* **NON ESISTE** *una combinazione lineare nulla di tali vettori con coefficienti non tutti nulli, e quindi se e solo se:*

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r = \mathbf{0} \qquad \Longrightarrow \qquad \alpha_1 = \alpha_2 = \ldots = \alpha_r = 0.$$

Si dice infine che un vettore $\mathbf{u}$ *dipende linearmente dai vettori* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *se appartiene allo spazio generato da* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, *cioé se esistono* $\alpha_1, \alpha_2, \ldots, \alpha_r$ *tali che*

$$\mathbf{u} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r.$$

■

Osservazione 2.2 - Se $r = 1$, allora $\mathbf{v}_1$ é linearmente dipendente se e solo se esiste $\alpha_1 \neq 0$ tale che $\alpha_1 \mathbf{v}_1 = \mathbf{0}$, e quindi se e solo se risulta $\mathbf{v}_1 = \mathbf{0}$. Conseguentemente un singolo vettore é linearmente indipendente se e solo se é diverso dal vettore nullo.

■

La seguente Proposizione caratterizza invece la nozione di vettori linearmente dipendenti o indipendenti nel caso $r > 1$

Proposizione 2.3 -. *Se* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *sono* r *vettori di* $\mathbf{R}^n$ *con* $r > 1$*, allora tali vettori sono linearmente dipendenti se e solo se uno di essi dipende linearmente dagli altri, cioé appartiene al sottospazio generato dagli altri.*

Viceversa i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$*, (con* $r > 1$*), sono linearmente indipendenti se nessuno di essi dipende dagli altri.*

Dim. Infatti, se i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono linearmente dipendenti ed $\alpha_1, \alpha_2, \ldots \alpha_r$ sono r coefficienti tali che $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}$ ed uno dei coefficienti (ad esempio α_1) é diverso da 0, allora si ha che

$$\mathbf{v}_1 = (-\alpha_2/\alpha_1) \cdot \mathbf{v}_2 + (-\alpha_3/\alpha_1) \cdot \mathbf{v}_3 + \ldots + (-\alpha_r/\alpha_1) \cdot \mathbf{v}_r,$$

e quindi $\mathbf{v}_1$ dipende da $\mathbf{v}_2, \mathbf{v}_3, \ldots \mathbf{v}_r$.

Viceversa, se uno dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, (ad esempio $\mathbf{v}_1$) appartiene al sottospazio generato dagli altri, allora esistono $\alpha_2, \alpha_3, \ldots, \alpha_r$ tali che $\mathbf{v}_1 = \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \ldots + \alpha_r\mathbf{v}_r$. Ne segue che risulta:

$$(-1)\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}, \qquad \text{con} \qquad (-1, \alpha_2, \alpha_3, \ldots \alpha_r) \neq (0, 0, \ldots 0),$$

e quindi i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono linearmente dipendenti.

■

Osservazione 2.4 - Se i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ sono linearmente dipendenti, allora uno di essi, (ad esempio $\mathbf{v}_1$), appartiene allo spazio generato dagli altri; ne segue, per la Osserv. 1.5, che

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r) = \mathcal{L}(\mathbf{v}_2, \ldots \mathbf{v}_r).$$

Pertanto, il fatto che i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ sono linearmente dipendenti significa che é possibile eliminare uno dei generatori senza modificare il sottospazio da essi generato; invece i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ sono linearmente indipendenti, se, eliminando uno dei generatori, il sottospazio generato cambia. In altri termini, i vettori sono linearmente indipendenti quando sono tutti indispensabili per generare il sottospazio da essi generato.

■

Osservazione 2.5 - Dal momento che il vettore nullo $\mathbf{0}$ appartiene a qualunque sottospazio di $\mathbf{R}^n$, é evidente che:

(1) se uno dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ é il vettore nullo, allora tali vettori sono linearmente dipendenti;
(2) se i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono linearmente indipendenti, allora tali vettori sono tutti diversi dal vettore nullo.

■

La Proposizione che segue contiene le principali proprietá connesse con i concetti di vettori linearmente dipendenti o indipendenti.

Proposizione 2.6 -. *Se* $X = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r\} \subset \mathbf{R}^n$ *é un insieme di vettori linearmente indipendenti, allora:*

a) *per ogni* $\mathbf{v} \in X$ *i vettori dell'insieme* $X - \{\mathbf{v}\}$ *sono linearmente indipendenti;*
b) *se* $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ *e* $\mathbf{v}$ *sono linearmente dipendenti, allora* $\mathbf{v}$ *appartiene al sottospazio generato da* $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$*;*
c) *se il sottospazio* $\mathcal{L}(X)$ *generato da* X *é diverso da* $\mathbf{R}^n$*, allora per ogni* $\mathbf{v} \in \mathbf{R}^n - \mathcal{L}(X)$ *si ha che i vettori di* $X \cup \{\mathbf{v}\}$ *sono linearmente indipendenti.*

Dim. di a). Se per assurdo esistesse $\mathbf{v} \in X$, (ad esempio $\mathbf{v} = \mathbf{v}_1$), tale che i vettori di $X - \{\mathbf{v}\} = \{\mathbf{v}_2, \ldots \mathbf{v}_r\}$ sono linearmente dipendenti, allora esisterebbe una combinazione lineare nulla $\alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}$ con coefficienti $\alpha_2, \alpha_3, \ldots, \alpha_r$ non tutti nulli. Posto $\alpha_1 = 0$, si otterrebbe una combinazione lineare nulla $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}$ dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ con coefficienti non tutti nulli, e questo é assurdo perché i vettori sono linearmente indipendenti.

Dim. di b). Per ipotesi sappiamo che esistono $r+1$ numeri reali non tutti nulli $\alpha, \alpha_1, \alpha_2, \ldots \alpha_r$ tali che $\alpha \mathbf{v} + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r = \mathbf{0}$. Se fosse $\alpha = 0$ sarebbe $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r = \mathbf{0}$ con $\alpha_1, \alpha_2, \ldots \alpha_r$ non tutti nulli, e quindi i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_r$ sarebbero linearmente dipendenti.
Se ne deduce che é $\alpha \neq 0$ e quindi che $\mathbf{v} = (-\alpha_1/\alpha)\mathbf{v}_1 + (-\alpha_2/\alpha)\mathbf{v}_2 + \ldots + (-\alpha_r/\alpha)\mathbf{v}_r$, cioé che $\mathbf{v} \in \mathcal{L}(X)$.

Dim di c). La tesi é ovvia conseguenza della b).

■

Osservazione 2.7 - Nel prossimo paragrafo impareremo un metodo per stabilire se le colonne o le righe di una matrice sono linearmente dipendenti o indipendenti. Questo ci forniá un metodo per stabilire se r vettori sono linearmente dipendenti o indipendenti: basterá infatti disporli come righe o colonne di una matrice.

■

3 . - Il caso delle righe o colonne di una matrice

Cominciamo con l'osservare che se A é una matrice di tipo $k \times n$, allora le colonne di A sono linearmente dipendenti se e solo se esistono n numeri $\alpha_1, \alpha_2, \ldots, \alpha_n$ non tutti nulli tali che $\alpha_1 A^1 + \alpha_2 A^2 + \ldots + \alpha_n A^n = \mathbf{0}_k$, cioé se e solo se il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ ha soluzioni non banali.
Invece le colonne di A sono linearmente indipendenti se e solo se il sistema $A\mathbf{x} = \mathbf{0}$ ha solo la soluzione nulla.
Analogamente, le righe di A sono vettori linearmente indipendenti in $\mathbf{R}^n$ se e solo se il sistema $A^T\mathbf{x} = \mathbf{0}$ ha solo la soluzione nulla.

Ebbene, sappiamo che (a meno di uno scambio di posizione delle incognite) il sistema $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema $A'\mathbf{x} = \mathbf{0}$, dove A' é stata ottenuta eseguendo su A l'algoritmo di Gauss Jordan completo, e quindi puó essere decomposta a blocchi in una delle forme

$$I_k = I_n, \qquad (\, I_k \quad C \,)\,, \qquad \begin{pmatrix} I_n \\ O \end{pmatrix}, \qquad \begin{pmatrix} I_r & C \\ O & O \end{pmatrix}.$$

Primo caso: A é diventata $A' = I_n$.

Il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema $I_n\mathbf{x} = \mathbf{0}$ che ha l'unica soluzione $\mathbf{x} = \mathbf{0}$ e quindi le colonne di A sono linearmente indipendenti.

Secondo caso: A é diventata $A' = (I_k, C)$

In questo caso, a meno di uno scambio di posizione delle incognite, il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema

$$\begin{cases} x_1 & = -c_{1,1}x_{k+1} - c_{1,2}x_{k+2} - \ldots - c_{1,n-k}x_n, \\ x_2 & = -c_{2,1}x_{k+1} - c_{2,2}x_{k+2} - \ldots - c_{2,n-k}x_n, \\ \ldots & \\ x_k & = -c_{k,1}x_{k+1} - c_{k,2}x_{k+2} - \ldots - c_{k,n-k}x_n, \end{cases}$$

Il sistema ha quindi (in aggiunta alla soluzione nulla) infinite soluzioni non nulle, che si ottengono attribuendo valori arbitrari alle variabili $x_{k+1}, x_{k+2}, \ldots, x_n$, e trovando i corrispondenti valori di $x_1, x_2, \ldots, x_k$; ne segue che le colonne di A sono linearmente dipendenti.
Ad esempio se A é diventata la matrice

$$A' = \begin{pmatrix} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 6 & -5 \\ 0 & 0 & 1 & 4 & -3 \end{pmatrix},$$

allora il sistema $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema

$$\begin{cases} x_1 &= -x_4 + x_5 \\ x_2 &= -6x_4 + 5x_5 \\ x_3 &= -4x_4 + 3x_5. \end{cases}$$

Il sistema $A\mathbf{x} = \mathbf{0}$ ha quindi infinite soluzioni; tali soluzioni si trovano dando valori arbitrari alle variabili x_4 ed x_5 e trovando i corrispondenti valori di x_1, x_2, x_3 .

Ad esempio una soluzione non nulla é data dal vettore $\mathbf{x}$ tale che $x_4 = 1$, $x_5 = 0$, $x_1 = -1$, $x_2 = -6$, $x_3 = -4$. Le colonne di A sono quindi linearmente dipendenti.

Terzo caso: A é diventata $A' = \begin{pmatrix} I_n \\ O \end{pmatrix}$.

Questa volta, per ogni $i > n$, la $i-$sima equazione del sistema omogeneo $A'\mathbf{x} = \mathbf{0}$ é

$$0x_1 + 0x_2 + \ldots + 0x_n = 0;$$

evidentemente tale equazione é soddisfatta da ogni vettore $\mathbf{x}$ di $\mathbf{R}^n$; essa é quindi superflua e puó essere eliminata dal sistema. Eliminate tali equazioni, si ottiene il sistema $I_n\mathbf{x} = \mathbf{0}$, che ha solo la soluzione nulla. Ne segue che le colonne di A sono linearmente indipendenti.

Quarto caso: A é diventata $A' = \begin{pmatrix} I_r & C \\ O & O \end{pmatrix}$. .

Questa volta le ultime $k - r$ equazioni del sistema lineare omogeneo $A'\mathbf{x} = \mathbf{0}$ saranno superflue e il sistema si trasforma in un sistema lineare omogeneo di r equazioni in n incognite del tipo $(I_r, C)\mathbf{x} = \mathbf{0}_r$.

Con lo stesso procedimento illustrato nello studio del "secondo caso" si vede che tale sistema ha infinite soluzioni non banali e quindi le colonne di A sono linearmente dipendenti. .

In definitiva si é provato che:

Proposizione 3.1 -. *Se A é una matrice non nulla di tipo $k \times n$, allora:*

a) *le colonne di A sono linearmente indipendenti se e solo se l'algoritmo di Gauss Jordan completo trasforma A in una matrice del tipo I_n o $\begin{pmatrix} I_n \\ O \end{pmatrix}$;*

b) *le colonne di A sono linearmente dipendenti se e solo se l'algoritmo di Gauss Jordan completo trasforma A in una matrice del tipo (I_k, C) o $\begin{pmatrix} I_r & C \\ O & O \end{pmatrix}$, con $r < n$.*

■

Corollario 3.2 -. *Se A é una matrice quadrata di ordine n, allora le seguenti proprietá sono equivalenti:*

a) *A é invertibile,*

b) $\det(A) \neq 0$,

c) *le colonne di A sono linearmente indipendenti,*

d) *le righe di A sono linearmente indipendenti.*

In particolare, sono linearmente indipendenti le righe e le colonne della matrice identica I_n , cioé i vettori

$$E_1 = (1, 0, 0, \ldots, 0), \qquad E_2 = (0, 1, 0, \ldots, 0), \qquad \ldots\ldots, \qquad E_n = (0, 0, \ldots, 0, 1).$$

Dim. L'equivalenza di a) e b) é contenuta nella Proposizione 5.2 del Capitolo II. L'equivalenza di b) e c) deriva dal fatto che per la Osservazione 2.2 del Capitolo II si ha che $\det(A) \neq 0$ se e solo se l'algoritmo di Gauss Jordan completo trasforma A in I_n e questo (per la Proposizione 3.1) equivale a chiedere che le colonne di A sono linearmente indipendenti.

Infine l'equivalenza di b) e d) discende dal fatto che le righe di A sono le colonne di A^T e quindi sono linearmente indipendenti se e solo se $\det(A^T) \neq 0$ e dunque se e solo se $\det(A) = \det(A^T) \neq 0$.

■

Corollario 3.3 -. *Se A é una matrice di tipo $k \times n$, allora:*

- *le colonne di A sono linearmente dipendenti se risulta $k < n$,*
- *le righe di A sono linearmente dipendenti se risulta $k > n$.*

Dim. Infatti, se risulta $k < n$, (rispett. $k > n$), allora l'algoritmo di Gauss Jordan trasforma certamente A, (rispettivamente A^T), in una matrice del tipo $(B \quad C)$ o $\begin{pmatrix} B & C \\ O & O \end{pmatrix}$.

■

Corollario 3.4 -. *Se $k > n$, allora k vettori di $\mathbf{R}^n$ sono linearmente dipendenti.*

Dim. Basta disporre i vettori come righe di una matrice ed applicare il precedente Corollario 3.3.

■

Osservazione 3.5 - In realtá possiamo essere molto piú precisi ed esaurienti riguardo alla lineare dipendenza o indipendenza delle righe e delle colonne di una matrice A.

Infatti, eseguiamo su A l'algoritmo di Gauss Jordan e trasformiamo A nella matrice A' di uno dei quattro tipi

$$B\,, \qquad (B \quad C)\,, \qquad \begin{pmatrix} B \\ O \end{pmatrix}, \qquad \begin{pmatrix} B & C \\ O & O \end{pmatrix},$$

con B matrice triangolare superiore non singolare e consideriamo le diverse alternative.

Primo caso: $A' = B$.

In questo caso A é una matrice quadrata e si ha che $|\det(A)| = |\det(B)| \neq 0$; ne segue (per il Corollario 3.2) che le colonne e le righe di A e di A' sono linearmente indipendenti.

Secondo caso: $A' = (B \quad C)$.

Questa volta risulta $k < n$ e quindi le colonne di A e di A' sono linearmente dipendenti. Peró, dette B^* e C^* le sottomatrici di A formate dalle colonne corrispondenti rispettivamente alle colonne di B e di C, si ha che l'algoritmo di Gauss Jordan trasformerebbe B^* in B e quindi (per quanto visto al caso precedente) le colonne e le righe di B^* e di B sono linearmente indipendenti.

D'altra parte, se a B^*, (rispett. B), aggiungiamo una colonna di C^*, (rispett. C), otteniamo una matrice di tipo $k \times (k+1)$, che ha le colonne linearmente dipendenti in virtú del Corollario 3.3; ne segue, per la b) della Proposizione 2.6, che qualunque colonna di C^* dipende linearmente dalle colonne di B^* e qualunque colonna di C dipende dalle colonne di B.

Infine le righe di A sono linearmente indipendenti, perché in caso contrario esisterebbero $\alpha_1, \alpha_2, \ldots \alpha_k$ non tutti nulli tali che $\alpha_1 A_1 + \alpha_2 A_2 + \ldots + \alpha_k A_k = \mathbf{0}_n$ e quindi tali che $\alpha_1 B_1^* + \alpha_2 B_2^* + \ldots + \alpha_k B_k^* = \mathbf{0}_k$, contraddicendo il fatto che le righe di B^* sono linearmente indipendenti. In maniera analoga si vede che le righe di A' sono linearmente indipendenti.

In definitiva, le k righe di A e di A' sono linearmente indipendenti; le colonne di B e le corrispondenti k colonne di A sono linearmente indipendenti, ogni altra colonna di A' e di A dipende linearmente da esse.

Terzo caso: $A' = \begin{pmatrix} B \\ O \end{pmatrix}$.

Questa volta risulta $k > n$ e quindi le righe di A ed A' sono linearmente dipendenti per il Corollario 3.3. Invece le colonne di A ed A' sono linearmente indipendenti per la Proposizione 3.1, dal momento che l'algoritmo di Gauss Jordan completo trasformerebbe A ed A' nella matrice $\begin{pmatrix} I_n \\ O \end{pmatrix}$.

D'altra parte, dette B^* e C^* le sottomatrici di A formate dalle righe corrispondenti rispettivamente alle righe di B e di O, si ha che l'algoritmo di Gauss Jordan trasformerebbe B^* in B e quindi (per quanto visto nel primo caso) le righe di B^* e di B sono linearmente indipendenti.

Se a B^* aggiungessimo una riga di C^*, otterremmo una matrice di tipo $(n+1)\times n$ che ha le righe linearmente dipendenti in virtú del Corollario 3.3; ne segue, per la b) della Proposizione 2.6, che le righe di C^* dipendono linearmente dalle righe di B^*. D'altronde le ultime $k-n$ righe di A' sono righe nulle e quindi dipendono linearmente dalle righe di B.

In definitiva le n colonne di A ed A' sono linearmente indipendenti; le righe di B e le corrispondenti n righe di A sono linearmente indipendenti, ogni altra riga di A' e di A dipende linearmente da esse.

Quarto caso: $A' = \begin{pmatrix} B & C \\ O & O \end{pmatrix}$, con B matrice triangolare non singolare di ordine $r < \min(k, n)$.

Questa volta si vede che

a) le r colonne di A' contenenti le colonne di B e le corrispondenti colonne di A sono linearmente indipendenti, le restanti $n-r$ colonne dipendono da esse;
b) le r righe di A' contenenti le righe di B e le corrispondenti righe di A sono linearmente indipendenti, le restanti $k-r$ righe dipendono da esse.
c) le k righe e le n colonne di A ed A' sono linearmente dipendenti;

Infatti, la c) discende chiaramente da a) e b).

Per quanto riguarda la a), osserviamo che l'algoritmo di Gauss Jordan completo trasformerebbe la sottomatrice $\begin{pmatrix} B \\ O \end{pmatrix}$ di A' e la sottomatrice di A formata dalle corrispondenti colonne di A nella matrice $\begin{pmatrix} I_r \\ O \end{pmatrix}$, e quindi tali colonne di A' e di A sono linearmente indipendenti.

D'altra parte, se a tali matrici aggiungessimo una ulteriore colonna di A' e la corrispondente colonna di A, otterremmo delle matrici che l'algoritmo di Gauss Jordan completo trasformerebbe in una matrice del tipo $\begin{pmatrix} I_r & C \\ O & O \end{pmatrix}$, e quindi (per la Proposizione 3.1) le colonne di tali matrici sono linearmente dipendenti. Ne segue, per la b) della Proposizione 2.6, che l'ultima colonna di tali matrici dipende linearmente dalle prime r. Pertanto le ultime $n-r$ colonne di A' dipendono linearmente dalle prime r e le colonne di A corrispondenti alle ultime $n-r$ colonne di A' dipendono linearmente dalle colonne di A corrispondenti alle prime r colonne di A'.

Rimane da provare la b). A tal fine, denotiamo $\hat{A}$ la matrice ottenuta da A disponendo le righe nello stesso ordine in cui sono disposte le corrispondenti righe di A', e proviamo che le prime r righe di $\hat{A}$ ed A' sono linearmente indipendenti e le ultime $k-r$ righe dipendono linearmente dalle prime r.

Infatti, l'algoritmo di Gauss Jordan trasformerebbe la sottomatrice di A formata dalle prime r righe di $\hat{A}$ nella matrice $\begin{pmatrix} B & C \end{pmatrix}$; ne segue, per quanto visto nel secondo caso, che le righe di tali matrici sono linearmente indipendenti.

Inoltre le ultime $k-r$ righe di A' sono righe nulle e quindi ovviamente dipendono dalle precedenti r righe; rimane dunque da provare che le ultime $k-r$ righe di $\hat{A}$ dipendono dalle prime r.

A tal fine, denotiamo $\hat{A}'$ la matrice ottenuta da A' disponendo le colonne nello stesso ordine in cui sono disposte le colonne di A, ed osserviamo che $\hat{A}'$ é stata ottenuta da $\hat{A}$ senza scambi di riga o di colonna, ma solo con r passi di pivot parziali.

Ebbene, nel primo passo di pivot $\hat{A}$ é stata trasformata in una matrice in cui la prima riga é rimasta inalterata e le successive sono state sostituite dalla somma delle stesse righe con un multiplo opportuno della prima.

Nel secondo passo di pivot le prime due righe attuali sono rimaste inalterate e le restanti sono state sostituite dalla somma delle stesse righe con un multiplo opportuno della attuale seconda riga e quindi dalla somma delle righe iniziali di $\hat{A}$ con una opportuna combinazione lineare delle prime due righe iniziali.

In generale, prima dell'i-simo passo di pivot, la prima riga iniziale é rimasta inalterata e tutte le altre righe di $\hat{A}$ sono state sostituite dalla somma delle stesse righe con una opportuna combinazione lineare delle prime $i-1$-righe iniziali di $\hat{A}$. L'i-simo passo di pivot lascia inalterate le attuali prime i righe e sostituisce le righe successive con la somma di tali righe con un multiplo opportuno della i-sima riga, cioé con la somma delle righe iniziali di $\hat{A}$ con una opportuna combinazione lineare delle prime i righe iniziali di $\hat{A}$.

Dopo r passi di pivot si ha che, per ogni $i > r$, la i-sima riga di $\hat{A}$ é stata sostituita dalla somma della stessa riga con una opportuna combinazione lineare delle prime r righe di $\hat{A}$ ed é diventata la i-sima riga di $\hat{A}'$, cioé una riga nulla. Pertanto esistono $\alpha_1, \alpha_2, \dots, \alpha_r$ tali che $\hat{A}_i + \alpha_1 \hat{A}_1 + \alpha_2 \hat{A}_2 + \ldots + \alpha_r \hat{A}_r = \mathbf{0}$ e quindi tali che $\hat{A}_i = -\alpha_1 \hat{A}_1 + (-\alpha_2)\hat{A}_2 + \ldots + (-\alpha_r)\hat{A}_r$.

Questo dimostra che le ultime $k-r$ righe di $\hat{A}$ dipendono linearmente dalle prime r righe, cioé che le righe di A corrispondenti alle righe nulle di A' dipendono linearmente dalla righe di A corrispondenti alle righe non nulle di A'.

■

Ad esempio, se eseguendo l'algoritmo di Gauss Jordan la matrice A é diventata

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

allora le quattro righe e le quattro colonne di A sono linearmente indipendenti.

Se invece le matrici A e B sono state trasformate rispettivamente nelle matrici

$$A' = \left(\begin{array}{rrrr|rr} 2 & -1 & 3 & 0 & -1 & 4 \\ 0 & 4 & 0 & 1 & 2 & -5 \\ 0 & 0 & -1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 3 \end{array}\right), \qquad B' = \left(\begin{array}{rrr} 2 & -1 & 3 \\ 0 & 4 & 2 \\ 0 & 0 & -1 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right),$$

e, nel corso dell'esecuzione dell'algoritmo, la terza colonna di A é stata scambiata con la quinta e la seconda riga di B é stata scambiata con la quarta, allora le quattro righe di A e le tre colonne di B sono linearmente indipendenti. Inoltre la prima, la seconda, la quinta e la quarta colonna di A sono linearmente indipendenti, la terza e la sesta colonna di A dipendono da esse; analogamente la prima, la quarta e la terza riga di B sono linearmente indipendenti, la seconda e la quinta riga dipendono da esse.

Infine se la matrice C é stata trasformata nella matrice

$$C' = \left(\begin{array}{rrr|rr} 2 & -1 & 3 & -2 & -1 \\ 0 & 4 & 2 & -1 & -3 \\ 0 & 0 & -1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

e, nel corso dell'esecuzione dell'algoritmo, sono state scambiate la seconda riga con la quarta e la terza colonna con la quinta, allora la prima, la quarta e la terza riga di C sono linearmente indipendenti, la seconda e la quinta riga dipendono da esse; analogamente la prima, seconda e quinta colonna di C sono linearmente indipendenti, la terza e la quarta colonna dipendono da esse.

■

4. - Sistemi di Generatori e Basi di un Sottospazio Vettoriale di $\mathbf{R}^n$.

Definizione 4.1 -. *Se H é un sottospazio vettoriale di $\mathbf{R}^n$ e $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ sono r vettori di $\mathbf{R}^n$ tali che H coincide con il sottospazio generato da tali vettori, allora si dice che H é generato dai vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ o che i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ generano H o formano un* **sistema di generatori** *di H.*

In altri termini H é generato dai vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ se e solo se per ogni $\mathbf{v} \in H$ esistono r numeri reali $\alpha_1, \alpha_2, \dots, \alpha_r$ tali che $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r$.

■

Esempio 1. - Lo spazio vettoriale $\mathbf{R}^n$ é generato dalle colonne della matrice identica I_n, cioé dai vettori

$$E^1 = (1, 0, 0, \dots, 0)^T, \qquad E^2 = (0, 1, 0, \dots, 0)^T, \qquad \ldots\ldots, \qquad E^n = (0, 0, \dots, 0, 1)^T.$$

Infatti, per ogni vettore $\mathbf{x} = (x_1, x_2, \dots ., x_n)^T$ risulta chiaramente $\mathbf{x} = x_1E^1 + x_2E^2 + \ldots + x_nE^n$.

■

Definizione 4.2 -. *Se $H \neq \{\mathbf{0}\}$ é un sottospazio di $\mathbf{R}^n$, dicesi* **base di** *H un insieme finito $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ di vettori di H che siano linearmente indipendenti e generino lo spazio H.*

■

Esempio 2. - Per quanto visto nel Coroll. 3.2 e nel precedente Esempio 1, i vettori

$$E^1 = (1, 0, 0, \dots, 0)^T, \qquad E^2 = (0, 1, 0, \dots, 0)^T, \qquad \ldots\ldots, \qquad E^n = (0, 0, \dots, 0, 1)^T$$

sono linearmente indipendenti e generano l'intero spazio $\mathbf{R}^n$; pertanto essi formano una base di $\mathbf{R}^n$. Tale base viene detta la **base canonica** di $\mathbf{R}^n$.

■

Si vede facilmente che:

Proposizione 4.3 -. *Se $\mathbf{v}_1, \mathbf{v}_2, \dots ., \mathbf{v}_r$ sono r vettori di H, allora le seguenti proprietá sono equivalenti:*
a) $\{\mathbf{v}_1, \mathbf{v}_2, \dots ., \mathbf{v}_r\}$ *é una base di H;*
b) *per ogni $\mathbf{v} \in H$ esiste uno ed un solo vettore $(\alpha_1, \alpha_2, \dots, \alpha_r) \in \mathbf{R}^r$ tale che*

$$(*) \qquad \mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r.$$

I numeri reali $\alpha_1, \alpha_2, \dots \alpha_r$ diconsi le **coordinate di v rispetto alla base** $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$.

Dim. a) $\Longrightarrow$ b). Fissiamo $\mathbf{v} \in H$; allora, l'esistenza dei numeri reali $\alpha_1, \alpha_2, \dots \alpha_r$ soddisfacenti la (*) discende dal fatto che $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ generano H. L'unicitá discende invece dal fatto che se $\alpha_1, \alpha_2, \dots \alpha_r$ e $\beta_1, \beta_2, \dots, \beta_r$ sono numeri reali tali che:

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r \qquad \text{e} \qquad \mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_r\mathbf{v}_r,$$

allora si ha:

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \ldots + (\alpha_r - \beta_r)\mathbf{v}_r = (\alpha_1\mathbf{v}_1 + \ldots + \alpha_r\mathbf{v}_r) - (\beta_1\mathbf{v}_1 + \ldots + \beta_r\mathbf{v}_r) = \mathbf{v} - \mathbf{v} = \mathbf{0},$$

donde, (per la lineare indipendenza dei vettori $\mathbf{v}_1, \mathbf{v}_2, \dots ., \mathbf{v}_r$), segue che

$$\alpha_1 - \beta_1 = \alpha_2 - \beta_2 = \dots .. = \alpha_r - \beta_r = 0 \qquad \text{e quindi} \qquad \alpha_j = \beta_j \quad \text{per ogni } j = 1, 2, \dots, n.$$

Dim. b) $\Longrightarrow$ a). Da b) si deduce che i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ generano H; inoltre, da b) segue che il vettore $\mathbf{0}$ puó essere rappresentato come combinazione lineare dei vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ solo mediante i coefficienti nulli. Questo dimostra che i suddetti vettori sono linearmente indipendenti e quindi formano una base di H.

■

Ad esempio per ogni $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$ i numeri reali $x_1, x_2, \dots, x_n$ rappresentano le coordinate di $\mathbf{x}$ rispetto alla base canonica $\{E^1, E^2, \dots, E^n\}$ di $\mathbf{R}^n$.

■

Esempio 3 - Sappiamo che se A é una matrice quadrata invertibile di ordine n, allora per ogni $\mathbf{y} \in \mathbf{R}^n$ il sistema lineare non omogeneo $A\mathbf{x} = \mathbf{y}$ ha una ed una sola soluzione, cioé esiste uno ed un solo vettore $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ tale che $x_1 A^1 + x_2 A^2 + \ldots + x_n A^n = \mathbf{y}$.

Questo prova, in base alla Prop. 4.3, che le colonne di A formano una base di $\mathbf{R}^n$ e che per ogni $\mathbf{y} \in \mathbf{R}^n$ l'unica soluzione $\mathbf{x}$ del sistema $A\mathbf{x} = \mathbf{y}$ é il vettore delle coordinate di $\mathbf{y}$ rispetto a tale base.

Analogamente, le righe di A, cioé le colonne di A^T, formano una base di $\mathbf{R}^n$; per ogni $\mathbf{y} \in \mathbf{R}^n$ l'unica soluzione del sistema $A^T\mathbf{x} = \mathbf{y}$ é il vettore delle coordinate di $\mathbf{y}$ rispetto a tale base.

■

La proposizione seguente contiene le principali proprietá delle basi.

Proposizione 4.4 -. *Se $H \neq \{\mathbf{0}\}$ é un sottospazio di $\mathbf{R}^n$, allora:*

(1) *ogni sistema di generatori di H contiene una base di H,*
(2) *ogni insieme finito di vettori linearmente indipendenti di H é contenuto in una base,*
(3) *esiste almeno una base (e quindi almeno un sistema di generatori) di H,*
(4) *se $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ é una base di H, allora k vettori di H, con $k > r$, sono linearmente dipendenti,*
(5) *due basi di H hanno lo stesso numero di elementi.*

Dim. di (1). Sia X un insieme di generatori di H; se gli elementi di X sono linearmente indipendenti, allora X é una base di H, altrimenti uno dei generatori $\mathbf{v}$ puó essere eliminato da X senza modificare lo spazio generato; a questo punto si sostituisce X con $X - \{\mathbf{v}\}$ e si ripete il procedimento.

Se X ha k elementi, dopo al massimo k iterazioni troveremo che i vettori residui generano H e sono diventati linearmente indipendenti; tali vettori formano quindi una base di H contenuta in X.

Dim. di (2). Sia X un insieme di r vettori linearmente indipendenti: se X é un insieme di generatori di H, allora X é una base di H, altrimenti si sceglie un vettore $\mathbf{v}$ in $H - \mathcal{L}(X)$. Per quanto visto nella Proposizione 2.6, i vettori di $X \cup \{\mathbf{v}\}$ sono linearmente indipendenti; a questo punto si sostituisce X con $X \cap \{\mathbf{v}\}$ e si ripete il procedimento.

Dal momento che per il Coroll. 3.4 in H esistono al massimo n vettori linearmente indipendenti, dopo al massimo $n - r$ iterazioni si sará trovato un insieme di vettori linearmente indipendenti che genera H e contiene X, e quindi una base di H che contiene X.

Dim. di (3). Basta fissare un punto $\mathbf{v} \neq \mathbf{0}$ in H ed applicare la (2).

Dim. di (4). Fissiamo k vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ di H, con $k > r$, e sia A la matrice di tipo $r \times k$ avente per colonne i vettori delle coordinate dei vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ rispetto alla base $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$.

Dall'essere $k > r$ si deduce per il Corollario 3.3 che le colonne di A sono linearmente dipendenti; esistono dunque k numeri non tutti nulli $\alpha_1, \alpha_2, \ldots \alpha_k$ tali che $\alpha_1 A^1 + \alpha_2 A^2 + \ldots + \alpha_k A^k = \mathbf{0}$.

Si vede allora che risulta

$$\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \ldots + \alpha_k \mathbf{u}_k = \mathbf{0}$$

e questo prova che i vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ sono linearmente dipendenti.

Infatti la matrice A soddisfa le relazioni

$$\begin{cases} \mathbf{u}_1 = a_{1,1}\mathbf{v}_1 + a_{2,1}\mathbf{v}_2 + \ldots\ldots + a_{r,1}\mathbf{v}_r \\ \mathbf{u}_2 = a_{1,2}\mathbf{v}_1 + a_{2,2}\mathbf{v}_2 + \ldots\ldots + a_{r,2}\mathbf{v}_r \\ \ldots\ldots \\ \mathbf{u}_k = a_{1,k}\mathbf{v}_1 + a_{2,k}\mathbf{v}_2 + \ldots\ldots + a_{r,k}\mathbf{v}_r \end{cases} \qquad \begin{cases} \alpha_1 a_{1,1} + \alpha_2 a_{1,2} + \ldots\ldots + \alpha_k a_{1,k} = 0 \\ \alpha_1 a_{2,1} + \alpha_2 a_{2,2} + \ldots\ldots + \alpha_k a_{2,k} = 0 \\ \ldots\ldots \\ \alpha_1 a_{r,1} + \alpha_2 a_{r,2} + \ldots\ldots + \alpha_k a_{r,k} = 0, \end{cases}$$

donde segue che

$$\begin{aligned}\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\ldots+\alpha_k\mathbf{u}_k =&(\alpha_1 a_{1,1}+\alpha_2 a_{1,2}+\ldots+\alpha_k a_{1,k})\mathbf{v}_1+\\ &+(\alpha_1 a_{2,1}+\alpha_2 a_{2,2}+\ldots+\alpha_k a_{2,k})\mathbf{v}_2+\\ &+\ldots\ldots+(\alpha_1 a_{r,1}+\alpha_2 a_{r,2}+\ldots+\alpha_k a_{r,k})\mathbf{v}_r=\\ =&\,0\cdot\mathbf{v}_1+0\cdot\mathbf{v}_2+\ldots+0\cdot\mathbf{v}_r=\mathbf{0}.\end{aligned}$$

Dim. di (5). Se $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_r\}$ e $\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_s\}$ sono due basi di H si ha $r=s$. Infatti dalla (4) si deduce che non puó essere $r<s$, (altrimenti i vettori $\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_s$ sarebbero linearmente dipendenti), e nemmeno puó essere $r>s$, (altrimenti i vettori $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_r$ sarebbero linearmente dipendenti).

■

5 . - Dimensione di un sottospazio vettroriale

La (5) della Prop. 4.4 giustifica la seguente

Definizione 5.1 -. *Se $H\neq\{\mathbf{0}\}$ é un sottospazio vettoriale di $\mathbf{R}^n$, dicesi* **dimensione** *di H, (in simboli* $\dim(H)$*), il numero di elementi di una qualunque base di H.*
Se $H=\{\mathbf{0}\}$, si dice che H é un sottospazio vettoriale di dimensione 0.

■

Esempio 1 - L'insieme $\mathbf{R}^n$ é uno spazio vettoriale di dimensione n , dal momento che una sua base é formata dai vettori

$$E^1=(1,0,0,\ldots,0)^T,\qquad E^2=(0,1,0,\ldots,0)^T,\qquad \ldots\ldots,\qquad E^n=(0,0,\ldots,0,1)^T.$$

■

Il concetto di dimensione di un sottospazio di $\mathbf{R}^n$ é di fondamentale importanza; il teorema seguente illustra le principali proprietá connesse a tale nozione.

Teorema 5.3 -. *Se $H\neq\{\mathbf{0}\}$ é un sottospazio vettoriale di dimensione r, allora:*

(1) *r é il minimo numero di generatori di H,*
(2) *r é il massimo numero di vettori linearmente indipendenti di H,*
(3) *r vettori linearmente indipendenti di H formano una base di H,*
(4) *r generatori di H formano una base di H.*

Dim di (1). Dalla (1) della Prop. 4.4 si deduce che ogni sistema di generatori contiene una base e quindi ha almeno r elementi, e d'altra parte ogni base di H é un insieme di generatori formato da r elementi.

Dim di (2). Infatti ogni base di H é un insieme di r vettori linearmente indipendenti e (per la (4) della Prop. 4.4) qualunque insieme di vettori linearmente indipendenti ha al massimo r elementi.

Dim di (3). Sia $X=\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_r\}$ un insieme di r vettori linearmente indipendenti di H; se X non fosse una base di H, per la (2) della Prop. 4.4, esisterebbe una base di H che contiene strettamente X e quindi ha piú di r elementi, contraddicendo l'ipotesi che $\dim(H)=r$.

Dim di (4). Sia $X=\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_r\}$ un sistema di generatori di H; se X non fosse una base, allora (per la (1) della Prop. 4.4) esisterebbe una base di H che é strettamente contenuta in X e quindi ha meno di r elementi, contraddicendo l'ipotesi che $\dim(H)=r$.

■

Osservazione 5.4 - Se $H = \{\mathbf{0}\}$, allora H non possiede vettori linearmente indipendenti e quindi il massimo numero di vettori linearmente indipendenti di H é 0; questo giustifica il fatto che abbiamo posto per definizione $\dim(\{\mathbf{0}\}) = 0$.

■

6. - Caratteristica o Rango di una Matrice

Sia A una matrice non nulla di tipo $k \times n$ e siano U il sottospazio di $\mathbf{R}^n$ generato dalle righe di A e V il sottospazio di $\mathbf{R}^k$ generato dalle colonne di A. Ci proponiamo di trovare la dimensione ed una base di U e di V.

A tal fine eseguiamo su A l'algoritmo di Gauss Jordan che trasforma la matrice A in una matrice A' del tipo

$$B, \qquad (\,B \quad C\,)\,, \qquad \begin{pmatrix} B \\ O \end{pmatrix}, \qquad \begin{pmatrix} B & C \\ O & O \end{pmatrix},$$

(con B matrice triangolare superiore non singolare), e siano U' e V' i sottospazi generati rispettivamente dalle righe e dalle colonne di A'.

Primo caso: $A' = B$**.**

In questo caso si ha che le $k = n$ righe di A e di A' sono linearmente indipendenti e formano una base degli spazi U ed U'; le $k = n$ colonne di A ed A' sono linearmente indipendenti e formano una base di V e V'; si ha quindi

$$\dim(U) = \dim(U') = \dim(V) = \dim(V') = k = n.$$

Secondo caso: $A' = (B \quad C)$**.**

Questa volta le k righe di A ed A' sono linearmente indipendenti e formano una base degli spazi U ed U'; si ha quindi $\dim(U) = \dim(U') = k$.

D'altra parte le k colonne di B e le k colonne di A corrispondenti alle colonne di B sono linearmente indipendenti e ogni altra colonna di A' e di A dipende linearmente da esse. Le suddette k colonne di B e le corrispondenti k colonne di A formano quindi una base rispettivamente di V' e di V e di conseguenza si ha: $\dim(V') = dim(V) = k$.

Terzo caso: $A' = \begin{pmatrix} B \\ O \end{pmatrix}$.

Questa volta le n colonne di A ed A' sono linearmente indipendenti e quindi formano una base rispettivamente degli spazi V e V'; si ha quindi $\dim(V) = \dim(V') = n$.

D'altra parte le righe di B e le corrispondenti n righe di A sono linearmente indipendenti e ogni altra riga di A' o di A dipende linearmente da esse. Le suddette n righe di A' e di A formano quindi una base di U' e di U e di conseguenza si ha: $\dim(U) = \dim(U') = n$.

Quarto caso: $A' = \begin{pmatrix} B & C \\ O & O \end{pmatrix}$, con B matrice triangolare invertibile di ordine $r < \min(k, n)$.

Questa volta le prime r righe e colonne di A' e le corrispondenti r righe e colonne di A sono linearmente indipendenti, le restanti $k - r$ righe e le restanti $n - r$ colonne dipendono da esse.

L'insieme delle prime r righe e l'insieme delle prime r colonne di A' formano una base di U' e V' rispettivamente; l'insieme delle r righe di A corrispondenti alle prime r righe di A' formano una base di U, l'insieme delle r colonne di A corrispondenti alle prime r colonne di A' formano una base di V.

Si ha dunque $\dim(U) = \dim(V) = \dim(U') = \dim(V') = r$.

In definitiva si é dimostrato che

Proposizione 6.1 -. *Se A é una matrice non nulla di tipo $k \times n$ e l'algoritmo di Gauss Jordan trasforma A in A', allora detti U ed U' i sottospazi di $\mathbf{R}^n$ generati dalle righe di A e di A' e detti V e V' i sottospazi di $\mathbf{R}^k$ generati dalle colonne di A ed A', si ha:*

$$\dim(U) = \dim(U') = \dim(V) = \dim(V') = \textit{numero delle righe non nulle di } A' \ .$$

D'altra parte, se A é la matrice nulla, si ha $U = \{\mathbf{0}_n\}$, $V = \{\mathbf{0}_k\}$ e quindi $\dim(U) = \dim(V) = 0$.

■

La Proposizione 6.1 giustifica la seguente definizione che introduce una nozione fondamentale dell'Algebra lineare.

Definizione 6.2 -. *Se A é una matrice di tipo $k \times n$, dicesi* **caratteristica** *o* **rango** *di A, (in simboli $car(A)$), la dimensione dello spazio generato dalle righe o dalle colonne di A.*
Pertanto, se A é una matrice non nulla, la caratteristica di A rappresenta il massimo numero di righe o di colonne linearmente indipendenti di A.

■

Osservazione 6.3 - Dalla Definizione 6.2 e dalla Proposizione 6.1 si deduce che:

a) la caratteristica di A coincide con la caratteristica della matrice A^T trasposta di A;
b) se A é una matrice non nulla e l'algoritmo di Gauss Jordan trasforma A in A', allora si ha che $car(A) = car(A')$.

Pertanto l'algoritmo di Gauss Jordan non fa altro che trasformare A in una matrice A' avente la stessa caratteristica di A, di cui sia immediato trovare la caratteristica, perché la caratteristica di A' é semplicemente il numero di righe non nulle di A'.

■

Esempio 1 - Trovare (al variare del parametro α) la caratteristica della matrice

$$A = \begin{pmatrix} 1 & 1 & -\alpha & 1 \\ \alpha & \alpha+1 & -1 & \alpha \\ 2 & 1 & 2 & \alpha^2+\alpha+2 \\ \alpha+1 & \alpha+2 & -\alpha-1 & 1-\alpha^2 \end{pmatrix}$$

Effettuando un primo passo di pivot parziale su $a_{1,1} = 1$, cioé sostituendo la riga A_2 con $A_2 - \alpha A_1$, la riga A_3 con $A_3 - 2A_1$ e la riga A_4 con $A_4 - (\alpha+1)A_1$, si ottiene la matrice:

$$\begin{pmatrix} 1 & 1 & -\alpha & 1 \\ 0 & 1 & \alpha^2-1 & 0 \\ 0 & -1 & 2+2\alpha & \alpha^2+\alpha \\ 0 & 1 & \alpha^2-1 & -\alpha^2-\alpha \end{pmatrix}.$$

Effettuando un secondo passo di pivot parziale su $a_{2,2} = 1$, cioé sostituendo la riga A_3 con $A_3 + A_2$ e la riga A_4 con $A_4 - A_2$, si ottiene la matrice:

$$A' = \begin{pmatrix} 1 & 1 & -\alpha & 1 \\ 0 & 1 & \alpha^2-1 & 0 \\ 0 & 0 & \alpha^2+2\alpha+1 & \alpha^2+\alpha \\ 0 & 0 & 0 & -\alpha^2-\alpha \end{pmatrix}.$$

Se α é diverso da 0 e da -1, si ha A' é una matrice triangolare superiore non singolare: le quattro righe e le quattro colonne di A sono linearmente indipendenti; esse formano una base rispettivamente dello spazio

U generato dalle righe e dello spazio V generato dalle colonne di A; la dimensione di tali sottospazi é 4 e quindi si ha che $car(A) = 4$.

Se $\alpha = 0$, la matrice A' diventa

$$A' = \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

Le quattro righe e le quattro colonne di A sono linearmente dipendenti; le prime tre righe (colonne) sono indipendenti, la quarta riga (colonna) dipende dalle prime tre. Una base di U é data dalle prime tre righe, una base di V é data dalle prime tre colonne e quindi si ha $car(A) = \dim(U) = \dim(V) = 3$.
Infine se $\alpha = -1$ la matrice A' diventa

$$A' = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Questa volta sia la terza che la quarta riga (colonna) di A dipendono linearmente dalle prime due, che sono invece linearmente indipendenti. Una base di U é data dalle prime due righe, una base di V é data dalle prime due colonne e quindi si ha $car(A) = \dim(U) = \dim(V) = 2$.

In conclusione la caratteristica di A é

$$r = \begin{cases} 4 & \text{se} \quad \alpha \text{ é diverso da } 0 \text{ e } -1, \\ 3 & \text{se} \quad \alpha = 0, \\ 2 & \text{se} \quad \alpha = -1. \end{cases}$$

In ogni caso la caratteristica di A é uguale al numero delle righe non nulle di A'.

■

La proposizione seguente consente di vedere sotto un'altra luce il concetto di caratteristica di una matrice.

Proposizione 6.4 -. *Se A é una matrice non nulla di tipo $k \times n$, allora la caratteristica di A rappresenta l'ordine massimo delle sottomatrici quadrate invertibili di A, cioé l'ordine massimo dele matrici quadrate non singolari che é possibile estrarre da A.*

Dim. Sia $r = car(A)$, cioé il massimo numero di righe o colonne linearmente indipendenti di A, e proviamo che:

a) se B é una sottomatrice quadrata invertibile di ordine h di A, allora le h righe e le h colonne di A che contengono le righe e le colonne di B sono linearmente indipendenti, e quindi si ha si ha che $h \leq r$;
b) A possiede una sottomatrice quadrata invertibile di ordine $r = car(A)$.

Ebbene la a) deriva dal fatto che se ci fosse una loro combinazione lineare nulla con coefficienti non tutti nulli, allora ci sarebbe una combinazione lineare nulla delle righe o colonne di B con gli stessi coefficienti e quindi le righe o colonne di B sarebbero linearmente dipendenti, contraddicendo il Corollario 3.2.
D'altra parte la b) discende dal fatto che l'algoritmo di Gauss Jordan trasforma A in una matrice A' del tipo

$$B, \qquad (B \quad C), \qquad \begin{pmatrix} B \\ O \end{pmatrix}, \qquad \begin{pmatrix} B & C \\ O & O \end{pmatrix}$$

(con B matrice triangolare superiore non singolare) e la caratteristica r di A é uguale al numero delle righe non nulle di A', cioé all'ordine della matrice B. Ebbene la sottomatrice quadrata di ordine r di A ottenuta intersecando le r righe e le r colonne corrispondenti alle righe e colonne di B é invertibile, poiché verrebbe trasformata in B dall'algoritmo di Gauss Jordan.

■

Osservazione 6.5 - Se chiamiamo **minore di ordine p estratto** da A il determinante di una sottomatrice quadrata di ordine p ottenuta intersecando p righe e p colonne di A, allora la tesi della precedente Proposizione puó essere riformulata dicendo che la caratteristica di A é l'ordine massimo dei minori non nulli che é possibile estrarre da A.

■

Osservazione 6.6 - La Prop. 6.4 suggerisce un metodo di calcolo della caratteristica di una matrice non nulla che é alternativo all'uso dell'algoritmo di Gauss Jordan. L'idea é quella di costruire progressivamente la sottomatrice quadrata invertibile di ordine massimo, aggiungendo alla sottomatrice invertibile giá costruita una riga ed una colonna, in modo che la nuova matrice sia ancora invertibile.

Ci si ferma quando sono state esaurite le righe o le colonne di A oppure quando sono nulli tutti i minori ottenuti aggiungendo alla sottomatrice B giá costruita una ulteriore riga ed una ulteriore colonna.

■

Possiamo formalizzare meglio il procedimento.

Se B é una sottomatrice quadrata di ordine h di A, dicesi **matrice orlata** di B la sottomatrice di ordine $h+1$ ottenuta aggiungendo, alle righe e alle colonne di A corrispondenti alle righe e colonne di B, una riga ed una colonna residua di A ed intersecando le nuove righe e colonne. Il determinante di una matrice orlata di B dicesi **minore orlato**.

Si ha allora il seguente

Metodo degli "orlati" di Kronecker.

(0) *si sceglie in A un elemento $a_{i,j} \neq 0$ e si pone $B = (\, a_{i,j}\,)$ e $r = 1$.*
(1) *se $r = k$, STOP: risulta $car(A) = r = k$;*
(2) *se $r = n$, STOP: risulta $car(A) = r = n$;*
(3) *se é possibile orlare B in modo che il determinante della matrice orlata sia diverso da 0, allora si sostituisce B con la matrice orlata, si sostituisce r con $r+1$ e si torna al passo (1);*
(4) *se tutti i minori ottenuti orlando B sono nulli, STOP: risulta $car(A) = r$;*

■

Esempio 2 - Consideriamo la matrice $A = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 4 & 2 & 1 & 3 \\ -1 & 2 & 1 & -3 \end{pmatrix}$.

Si pone $B = (a_{1,1}) = (2)$, $r = 1$; B fornisce un minore di ordine 1 diverso da 0 estratto da A.

Il minore ottenuto orlando B con la seconda riga e la seconda colonna, cioé il determinante della matrice $\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$, é nullo e quindi va scartato. Invece il minore ottenuto orlando B con la seconda riga e la terza colonna, cioé il determinante della matrice $\begin{pmatrix} 2 & -1 \\ 4 & 1 \end{pmatrix}$, é diverso da 0 e quindi va accettato.

Si pone dunque $B = \begin{pmatrix} 2 & -1 \\ 4 & 1 \end{pmatrix}$, $r = 2$.

La matrice ottenuta orlando B con la terza riga e la quarta colonna é tale che

$$\det \begin{pmatrix} 2 & -1 & 0 \\ 4 & 1 & 3 \\ -1 & 1 & -3 \end{pmatrix} = \det \begin{pmatrix} 2 & -1 & 0 \\ 3 & 2 & 0 \\ -1 & 1 & -3 \end{pmatrix} = (-3)(4+3) = -21 \neq 0.$$

Si pone dunque $B = \begin{pmatrix} 2 & -1 & 0 \\ 4 & 1 & 3 \\ -1 & 1 & -3 \end{pmatrix}$, $r = 3$. Avendo esaurito tutte le righe di A, si ha che $car(A) = 3$.

■

Esempio 3 - Consideriamo la matrice $A = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 4 & 2 & 1 & 3 \\ -2 & -1 & 4 & 3 \end{pmatrix}$.

Come prima si trova che $B = \begin{pmatrix} 2 & -1 \\ 4 & 1 \end{pmatrix}$ é una sottomatrice con determinante diverso da 0 e quindi provvisoriamente é $r = 2$.
Le uniche sottomatrici che é possibile estrarre da A orlando B sono quelle che si ottengono orlando B con la terza riga e la seconda colonna oppure con la terza riga e la quarta colonna.
Ebbene si ha

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

$$\det \begin{pmatrix} 2 & -1 & 0 \\ 4 & 1 & 3 \\ -2 & 4 & 3 \end{pmatrix} = \det \begin{pmatrix} 0 & -1 & 0 \\ 6 & 1 & 3 \\ 6 & 4 & 3 \end{pmatrix} = -(-1) \cdot \det \begin{pmatrix} 6 & 3 \\ 6 & 3 \end{pmatrix} = 18 - 18 = 0.$$

Pertanto tutte le sottomatrici che é possibile estrarre da A orlando B hanno determinante nullo e quindi si ha $car(A) = 2$.

■

7. - Rette, Piani e Iperpiani (Affini)

Definizione 7.1 -. *Un sottospazio di* $\mathbf{R}^n$ *di dimensione* 1, *(cioé generato da un unico vettore non nullo* $\mathbf{u}$), *dicesi anche* **retta**; *il vettore* $\mathbf{u}$ *dicesi la* **direzione** *della retta.*
Invece, un sottospazio di dimensione 2, *(cioé generato da due vettori non allineati e quindi linearmente indipendenti,* $\mathbf{u}$ *e* $\mathbf{v}$*) dicesi* **piano**. *La coppia di vettori* $(\mathbf{u}, \mathbf{v})$ *dicesi* **giacitura** *del piano.*

Ovviamente la direzione di una retta non é unica, poiché se $\mathbf{u}$ genera U, anche $\alpha\mathbf{u}$ genera U per ogni $\alpha \in \mathbf{R}$. Analogamente la giacitura di un piano non é univocamente determinata.

■

Osserviamo che nello spazio vettoriale $\mathbf{R}^3$ un piano é un sottospazio di dimensione $2 = dim(\mathbf{R}^3) - 1$; per analogia si dá la seguente

Definizione 7.2 -. *Un sottospazio di dimensione* $n-1$ *di* $\mathbf{R}^n$ *dicesi* **iperpiano** *di* $\mathbf{R}^n$.

■

Definizione 7.3 -. *Se* U *é un sottospazio vettoriale di dimensione* k *di* $\mathbf{R}^n$ *ed* $\mathbf{a} \in \mathbf{R}^n$, *l'insieme*

$$a + U = \{\mathbf{x} \in \mathbf{R}^n \mid \exists \mathbf{u} \in U \text{ tale che } \mathbf{x} = \mathbf{a} + \mathbf{u}\} = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} - \mathbf{a} \in U\},$$

ottenuto traslando il sottospazio U *a partire dal vettore* $\mathbf{a}$, *dicesi* **sottospazio affine** *di dimensione* k *di* $\mathbf{R}^n$, *passante per* $\mathbf{a}$ *e parallelo ad* U .

Definizione 7.4 -. *Un sottospazio affine di dimensione* 1 *dicesi* **retta affine**; *un sottospazio affine di dimensione* 2 *dicesi* **piano affine.** *Un sottospazio affine di dimensione* $n-1$ *di* $\mathbf{R}^n$ *dicesi* **iperpiano affine** *di* $\mathbf{R}^n$.
Due sottospazi affini si dicono **paralleli** *se sono paralleli ad uno stesso sottospazio vettoriale di* $\mathbf{R}^n$.

∎

Osservazione 7.5 - *(Equazioni parametriche di una retta affine di $\mathbf{R}^n$).*
Se $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$ é un punto di $\mathbf{R}^n$ ed $\mathbf{u} = (u_1, u_2, \dots, u_n)^T$ é un vettore non nullo di $\mathbf{R}^n$, la retta per $\mathbf{a}$ parallela al vettore $\mathbf{u}$ é dato da tutti e soli i punti $\mathbf{x}$ di $\mathbf{R}^n$ del tipo

$$\mathbf{x} = \mathbf{a} + \alpha \mathbf{u}, \qquad \text{con } \alpha \in \mathbf{R}, \tag{1}$$

cioé tutti e soli i punti $\mathbf{x} = (x_1, x_2, \dots, x_n)$ di $\mathbf{R}^n$ soddisfacenti le relazioni

$$\begin{cases} x_1 = & a_1 + \alpha u_1, \\ x_2 = & a_2 + \alpha u_2, \\ \dots & \dots \\ x_n = & a_n + \alpha u_n, \end{cases} \qquad \text{con } \alpha \in \mathbf{R}, \tag{2}$$

che diconsi le **equazioni parametriche** della retta passante per $\mathbf{a}$ e parallela ad $\mathbf{u}$, mentre la (1) dicesi l'**equazione parametrica vettoriale** della stessa retta.
Ad esempio, in $\mathbf{R}^3$ le equazioni parametriche della retta passante per $\mathbf{a} = (1, 0, -2)^T$ e parallela al vettore $\mathbf{u} = (1, -2, 4)^T$ sono

$$\begin{cases} x_1 = 1 + \alpha, \\ x_2 = -2\alpha, \\ x_3 = -2 + 4\alpha, \end{cases} \qquad \text{con } \alpha \in \mathbf{R}.$$

Due rette affini di equazioni parametriche vettoriali $\mathbf{x} = \mathbf{a} + \alpha \mathbf{u}$, $\mathbf{x} = \mathbf{b} + \alpha \mathbf{v}$, sono parallele, se il sottospazio generato da $\mathbf{u}$ coincide con il sottospazio generato da $\mathbf{v}$, e quindi se e solo se esiste λ tale che $\mathbf{v} = \lambda \mathbf{u}$.

∎

Osservazione 7.6 - (*Equazioni parametriche di un piano affine).*
Il piano per $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$ parallelo al piano generato dai vettori $\mathbf{u} = (u_1, u_2, \dots, u_n)^T$ e $\mathbf{v} = (v_1, v_2, \dots, v_n)^T$ é dato da tutti e soli i punti di $\mathbf{R}^n$ che soddisfano l' equazione vettoriale

$$\mathbf{x} = \mathbf{a} + \alpha \mathbf{u} + \beta \mathbf{v} \tag{3}$$

cioé le equazioni

$$\begin{cases} x_1 = & a_1 + \alpha u_1 + \beta v_1, \\ x_2 = & a_2 + \alpha u_2 + \beta v_2, \\ \dots & \dots \\ x_n = & a_n + \alpha u_n + \beta v_n, \end{cases} \qquad \text{con } \alpha, \beta \in \mathbf{R}, \tag{4}$$

che diconsi le **equazioni parametriche** del suddetto piano.
Ad esempio le equazioni parametriche del piano passante per $\mathbf{a} = (1, 0, -2)^T$ e parallela al piano generato dai vettori $\mathbf{u} = (1, -2, 4)^T$ e $\mathbf{v} = (0, 2, -3)^T$, sono

$$\begin{cases} x_1 = 1 + \alpha, \\ x_2 = -2\alpha + 2\beta, \\ x_3 = -2 + 4\alpha - 3\beta, \end{cases} \qquad \text{con } \alpha,\ \beta \in \mathbf{R}.$$

Due piani affini di equazioni parametriche vettoriali $\mathbf{x} = \mathbf{a} + \alpha \mathbf{u} + \beta \mathbf{v}$, $\mathbf{x} = \mathbf{b} + \alpha \mathbf{u}' + \beta \mathbf{v}'$, sono paralleli, se il sottospazio generato da $\{\mathbf{u}, \mathbf{v}\}$ coincide con il sottospazio generato da $\{\mathbf{u}', \mathbf{v}'\}$.

■

Osservazione 7.7 - *(Retta passante per due punti distinti).*
Se ora $\mathbf{a}, \mathbf{b}$ sono due punti distinti di $\mathbf{R}^n$, allora la retta passante per tali punti passa per $\mathbf{a} = \mathbf{a} + 0(\mathbf{b} - \mathbf{a})$ e per $\mathbf{b} = \mathbf{a} + 1(\mathbf{b} - \mathbf{a})$; ne segue che essa non é altro che la retta passante per $\mathbf{a}$ e parallelo al vettore $\mathbf{u} = \mathbf{b} - \mathbf{a}$; l'equazione parametrica vettoriale di tale retta é quindi:

$$\mathbf{x} = \mathbf{a} + \alpha(\mathbf{b} - \mathbf{a}) = \alpha\mathbf{b} + (1 - \alpha)\mathbf{a}, \qquad \text{con } \alpha \in \mathbf{R}.$$

I punti $\mathbf{a}$ e $\mathbf{b}$ si trovano in corrispondenza del valore $\alpha = 0$ e $\alpha = 1$ del parametro α; per $0 < \alpha < 1$ si trovano i punti del segmento congiungente $\mathbf{a}$ con $\mathbf{b}$; per $\alpha < 0$, (rispett. $\alpha > 1$), si ottengono i punti della semiretta che precedono $\mathbf{a}$, (rispett. che seguono $\mathbf{b}$).
Ad esempio le equazioni parametriche della retta congiungente i punti $(0, 3, -2)^T$ e $(1, -3, 5)^T$ sono:

$$\mathbf{x} = \alpha \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix} \qquad \text{cioé} \qquad \begin{cases} x_1 = \alpha \\ x_2 = -3\alpha + 3(1 - \alpha) = 3 - 6\alpha \\ x_3 = 5\alpha - 2(1 - \alpha) = -2 + 7\alpha \end{cases} \qquad \text{con } \alpha \in \mathbf{R}.$$

■

Osservazione 7.8 - *(Piano passante per tre punti non allineati).*
Analogamente il piano passante per tre punti non allineati, $\mathbf{a}, \mathbf{b}$ e $\mathbf{c}$ di $\mathbf{R}^n$, passa per i punti

$$\mathbf{a} = \mathbf{c} + 1(\mathbf{a} - \mathbf{c}) + 0(\mathbf{b} - \mathbf{c}), \quad \mathbf{b} = \mathbf{c} + 0(\mathbf{a} - \mathbf{c}) + 1(\mathbf{b} - \mathbf{c}), \quad \mathbf{c} = \mathbf{c} + 0(\mathbf{a} - \mathbf{c}) + 0(\mathbf{b} - \mathbf{c});$$

esso non é altro che il piano passante per $\mathbf{c}$ e parallelo al piano generato dai vettori $\mathbf{u} = \mathbf{a} - \mathbf{c}$ e $\mathbf{v} = \mathbf{b} - \mathbf{c}$; l'equazione parametrica vettoriale di tale piano é quindi:

$$\mathbf{x} = \mathbf{c} + \alpha(\mathbf{a} - \mathbf{c}) + \beta(\mathbf{b} - \mathbf{c}) = \alpha\mathbf{a} + \beta\mathbf{b} + (1 - \alpha - \beta)\mathbf{c} \qquad \text{con } \alpha,\, \beta \in \mathbf{R}.$$

Ad esempio, l'equazione vettoriale del piano in $\mathbf{R}^3$ passante per i punti $\mathbf{a} = (0, 2, -4)$, $\mathbf{b} = (1, 0, -2)$ e $\mathbf{c} = (1, 2, -3)$ é

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + \alpha \begin{pmatrix} 0 - 1 \\ 2 - 2 \\ -4 + 3 \end{pmatrix} + \beta \begin{pmatrix} 1 - 1 \\ 0 - 2 \\ -2 + 3 \end{pmatrix} = \begin{pmatrix} 1 - \alpha \\ 2 - 2\beta \\ -3 - \alpha + \beta \end{pmatrix};$$

le equazioni parametriche dello stesso piano sono

$$\begin{cases} x_1 = 1 - \alpha, \\ x_2 = 2 - 2\beta, \\ x_3 = -3 - \alpha + \beta, \end{cases} \qquad \text{con } \alpha, \beta \text{ arbitrari numeri reali.}$$

■

Osservazione 7.9 - *(Equazioni cartesiane di una retta o di un piano)* Dalle equazioni parametriche di una retta in $\mathbf{R}^n$, ricavando il parametro α da una delle equazioni e sostituendolo nelle rimanenti, si ha che la retta é data da tutti e soli i punti di $\mathbf{R}^n$ che soddisfano un sistema di $n - 1$ equazioni lineari nelle n incognite $x_1, x_2, \ldots, x_n$, che diconsi le **equazioni cartesiane della retta** .
Ad esempio, dalle equazioni parametriche della retta passante per $(1, 3, -1)$ parallela al vettore $(1, -2, 3)$, cioé

$$\begin{cases} x_1 = 1 + \alpha, \\ x_2 = 3 - 2\alpha, \\ x_3 = -1 + 3\alpha \end{cases}$$

si ricava
$$\begin{cases} \alpha = x_1 - 1, \\ x_2 = 3 - 2x_1 + 2 = 5 - 2x_1, \\ x_3 = -1 + 3(x_1 - 1) = 3x_1 - 4, \end{cases}$$
e quindi le equazioni cartesiane della retta sono
$$\begin{cases} 2x_1 + x_2 = 5, \\ 3x_1 - x_3 = 4. \end{cases}$$

Analogamente, nel caso di un piano affine in $\mathbf{R}^n$, ricavando i parametri α e β da due delle equazioni parametriche e sostituendoli nelle rimanenti $n-2$ equazioni, si ricava che il piano é dato da tutti e soli i punti di $\mathbf{R}^n$ che soddisfano un sistema di $n-2$ equazioni lineari nelle incognite $x_1, x_2, \dots, x_n$, che diconsi le **equazioni cartesiane del piano**.

Ad esempio dalle equazioni parametriche del piano passante per $\mathbf{a} = (1, 0, -2, 0)^T$ e parallela al piano generato dai vettori $\mathbf{u} = (1, -2, 4, 1)^T$ e $\mathbf{v} = (0, 2, -3, -1)^T$,
$$\begin{cases} x_1 = 1 + \alpha, \\ x_2 = -2\alpha + 2\beta, \\ x_3 = -2 + 4\alpha - 3\beta, \\ x_4 = \alpha - \beta \end{cases} \qquad \text{con } \alpha,\ \beta \in \mathbf{R},$$
ricavando i parametri $\alpha = x_1 - 1$ e $\beta = \alpha - x_4 = x_1 - x_4 - 1$ dalla prima e quarta equazione e sostituendoli nelle restanti due, si ottengono le equazioni cartesiane del piano in $\mathbf{R}^4$:
$$\begin{cases} x_2 = -2\alpha + 2\beta = -2x_1 + 2 + 2x_1 - 2x_4 - 2 = -2x_4, \\ x_3 = -2 + 4\alpha - 3\beta = -2 + 4x_1 - 4 - 3x_1 + 3x_4 + 3 = x_1 + 3x_4 - 3 \end{cases}$$
cioé
$$\begin{cases} x_2 + 2x_4 = 0 \\ x_1 - x_3 + 3x_4 = 3. \end{cases}$$
■

E S E R C I Z I

n. 1 - Dopo avere ricordato la definizione di vettori linearmente indipendenti, verificare (mediante la definizione) che i vettori $\mathbf{u} = (1, 3, -3, 0)^T$, $\mathbf{v} = (1, 2, 0, -3,)^T$ e $\mathbf{w} = (0, -1, 4, -5)^T$ sono linearmente indipendenti. **n. 2 -** Date le matrici

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 2 & -1 & 4 \\ -3 & -1 & -6 \\ 4 & -7 & 8 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 1 & 3 & -1 & 4 \\ 2 & 5 & -3 & 3 \end{pmatrix},$$

trovare (con l'algoritmo di Gauss Jordan e con il metodo di Kronecker) la caratteristica di A, di B, di $A + B^T$, di $A^T + B$, di AB e BA, indicando in ogni caso se le righe e le colonne sono linearmente dipendenti o indipendenti, nonché una base del sottospazio generato dalle righe ed una base del sottospazio generato dalle colonne.

n. 3 - Trovare la dimensione ed una base degli spazi generati dalle righe e dalle colonne delle matrici:

$$A = \begin{pmatrix} 0 & 2 & 3 & -1 & 4 & 0 \\ 2 & 0 & 2 & -4 & -6 & 0 \\ 1 & 3 & 2 & -4 & 0 & -2 \\ 0 & -2 & 0 & 1 & 0 & 3 \\ 3 & 0 & 0 & -1 & 2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & -2 & 0 & -3 & 0 \\ -1 & 2 & 0 & 0 & -6 \\ 1 & -3 & 2 & -4 & 0 \\ 0 & 2 & 0 & 0 & 3 \\ 3 & 0 & 0 & 2 & 1 \end{pmatrix}.$$

Di tali matrici trovare la caratteristica con il metodo degli orlati di Kronecker.

n. 4 - Al variare del parametro α trovare la dimensione ed una base degli spazi generati dalle righe e dalle colonne delle matrici:

$$A = \begin{pmatrix} 1 & 0 & \alpha & \alpha+2 \\ 0 & 1 & 2 & \alpha+1 \\ -1 & 2 & 3 & 1 \\ \alpha & 0 & \alpha^2 & 2\alpha+1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ -2 & 1-2\alpha & -2 & \alpha+2 \\ -1 & 1-\alpha & -1 & 2\alpha \\ 0 & -1 & 1 & 0 \end{pmatrix}.$$

n. 5 - Dati i vettori $\mathbf{a} = (1, , 2, -3, 0)^T$, $\mathbf{u} = (1, , -1, 0, -3,)^T$ e $\mathbf{v} = (2, -1, 4, -5)^T$ trovare le equazioni parametriche e cartesiane:

a) della retta affine passante per **a** e parallela ad **u** e della retta affine passante per **u** e **v**,
b) del piano affine passante per **a** e parallelo al piano generato da **u** e **v** e del piano affine passante per **a**, **u** e **v**.
c) dell'iperpiano H generato dai vettori **a**, **u** e **v**.

n. 6 - Dati i vettori $\mathbf{a} = (1, 3, 2, -4)^T$, $\mathbf{u} = (2, -2, 3, 5)^T$ e $\mathbf{v} = (-1, 3, -4, 0)$, trovare le equazioni parametriche e cartesiane della retta affine passante per **a** e parallela al vettore **u** e del piano affine passante per **a** e parallelo al piano generato da **u** e **v**.

Trovare poi le equazioni parametriche e cartesiane della retta passante per **a** e **u** e del piano passante per **a**, **u** e **v**.

Capitolo IV - SISTEMI LINEARI

1. - Il teorema fondamentale sui sistemi lineari.

Siamo ora in grado di studiare compiutamente i sistemi lineari introdotti nel n. 7 del Capitolo I.

Vedremo come affrontare i problemi enunciati lí, e precisamente:

(1) sotto quali condizioni un sistema omogeneo ha solo la soluzione banale?
(2) sotto quali condizioni un sistema nonomogeneo é compatibile o incompatibile ?
(3) se il sistema é compatibile, sotto quali condizioni é determinato o indeterminato?
(4) se il sistema é indeterminato, qual'é la struttura dell'insieme delle sue soluzioni?

Il seguente Teorema fornisce una risposta esauriente a tali problemi e la sua dimostrazione contiene un algoritmo pratico di risoluzione del sistema.

Teorema 1.1 -. *Sia A una matrice di tipo $k \times n$ e sia $\mathbf{b} \in \mathbf{R}^k$; si ha allora quanto segue.*

(1) *Il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ ha solo la soluzione banale se e solo se la caratteristica di A é uguale al numero n delle incognite.*

(2) **(Teorema di Rouché Capelli) -** *Il sistema nonomogeneo $A\mathbf{x} = \mathbf{b}$ é compatibile se e solo se la matrice A dei coefficienti e la matrice $\hat{A} = (A, \mathbf{b})$ dei coefficienti e termini noti hanno la stessa caratteristica.*

(3) *L'insieme U delle soluzioni del sistema omogeneo $A\mathbf{x} = \mathbf{0}$ é un sottospazio vettoriale di $\mathbf{R}^n$.*

(4) *Se il sistema non omogeneo $A\mathbf{x} = \mathbf{b}$ é compatibile, allora l'insieme delle soluzioni di tale sistema é un sottospazio affine parallelo ad U.*

(5) *Il sistema $A\mathbf{x} = \mathbf{b}$ é* $\begin{cases} \textit{determinato se risulta } car(A) = car(A, \mathbf{b}) = n, \\ \textit{indeterminato se risulta } car(A) = car(A, \mathbf{b}) < n. \end{cases}$

(6) *L'insieme U delle soluzioni del sistema $A\mathbf{x} = \mathbf{0}$ é un sottospazio vettoriale di dimensione $n - r = n - car(A)$.*

Pertanto si ha che $U = \{\mathbf{0}\}$, (cioé il sistema ha solo la soluzione nulla), se e solo se $car(A) = n$. Se invece risulta $car(A) = r < n$, allora esistono $n - r$ soluzioni linearmente indipendenti, (e quindi non banali), $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}$ di tale sistema omogeneo che generano U, cioé tali che le soluzioni del sistema sono tutti e soli i vettori di $\mathbf{R}^n$ del tipo :

$$\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_{n-r}\mathbf{u}_{n-r}, \qquad \textit{con } \alpha_1, \alpha_2, \ldots, \alpha_{n-r} \textit{ numeri reali arbitrari.}$$

(7) *Se il sistema non omogeneo $A\mathbf{x} = \mathbf{b}$ é indeterminato, allora le soluzioni del sistema sono tutti e soli i vettori di $\mathbf{R}^n$ del tipo:*

$$\mathbf{x} = \mathbf{x}^* + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_{n-r}\mathbf{u}_{n-r}, \qquad \textit{con } \alpha_1, \alpha_2, \ldots, \alpha_{n-r} \textit{ numeri reali arbitrari ,}$$

dove $\mathbf{x}^*$ é una soluzione del sistema lineare nonomogeneo ed $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-r}$ sono $n - r$ soluzioni linearmente indipendenti del sistema lineare omogeneo associato $A\mathbf{x} = \mathbf{0}$.

Nota bene - Il fatto che il sistema omogeneo o non omogeneo ha infinite soluzioni dipendenti da $n - r$ parametri $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ viene espresso dicendo che il sistema ha ∞^{n-r} soluzioni.

Dim. di (1). Sappiamo che il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ ha solo la soluzione nulla se e solo se le colonne di A sono linearmente indipendenti e quindi se e solo se la dimensione dello spazio generato dalle colonne di A, (cioé la caratteristica di A), é uguale ad n.

Dim. di (2). Infatti il sistema $A\mathbf{x} = \mathbf{b}$ é compatibile se e solo se esistono n numeri reali $x_1, x_2, \ldots, x_n$ tali che $\mathbf{b} = x_1 A^1 + x_2 A^2 + \ldots + x_n A^n$ e quindi se e solo se $\mathbf{b}$ appartiene allo spazio generato dalle colonne di A. Pertanto il sistema $A\mathbf{x} = \mathbf{b}$ é compatibile se e solo se risulta $\mathcal{L}(A^1, A^2, \ldots, A^n, \mathbf{b}) = \mathcal{L}(A^1, A^2, \ldots, A^n)$ e quindi se e solo se $car(\hat{A}) = \dim(\mathcal{L}(A^1, A^2, \ldots, A^n, \mathbf{b})) = \dim(\mathcal{L}(A^1, A^2, \ldots, A^n) = car(A)$.

Dim. di (3). Infatti se $\mathbf{u}$ e $\mathbf{v}$ sono due elementi di U, (cioé due soluzioni del sistema omogeneo), ed $\alpha \in \mathbf{R}$, allora risulta $A\mathbf{u} = \mathbf{0}$ e $A\mathbf{v} = \mathbf{0}$ e quindi

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0} \qquad \text{e} \qquad A(\alpha u) = \alpha \cdot A\mathbf{u} = \alpha \cdot \mathbf{0} = \mathbf{0}.$$

Pertanto $\mathbf{u} + \mathbf{v}$ ed $\alpha\mathbf{u}$ sono elementi di U e questo prova che U é un sottospazio vettoriale di $\mathbf{R}^n$.

Dim. di (4). Sia $\mathbf{x}^*$ una soluzione del sistema non omogeneo $A\mathbf{x} = \mathbf{b}$ e proviamo che l'insieme X delle soluzioni del suddetto sistema non omogeneo é dato da

$$X = \mathbf{x}^* + U = \{\mathbf{x} \in \mathbf{R}^n | \exists \mathbf{u} \in U \ t.c. \ \mathbf{x} = \mathbf{x}^* + \mathbf{u}\}$$

Infatti si ha $A\mathbf{x}^* = \mathbf{b}$ e quindi

$$\mathbf{u} \in U \quad \Longrightarrow \quad A\mathbf{u} = \mathbf{0} \quad \Longrightarrow A(\mathbf{x}^* + \mathbf{u}) = A\mathbf{x}^* + A\mathbf{u} = \mathbf{b} + \mathbf{0} = \mathbf{b} \quad \Longrightarrow \mathbf{x}^* + \mathbf{u} \in X$$

e viceversa

$$\mathbf{x} \in X \implies A\mathbf{x} = \mathbf{b} \implies A(\mathbf{x} - \mathbf{x}^*) = A\mathbf{x} - A\mathbf{x}^* = \mathbf{b} - \mathbf{b} = \mathbf{0} \implies \mathbf{u} = \mathbf{x} - \mathbf{x}^* \in U \ \text{ ed } \ \mathbf{x} = \mathbf{x}^* + \mathbf{u}.$$

Dim. di (5) . Da 4) si deduce che se il sistema non omogeneo $A\mathbf{x} = \mathbf{b}$ é compatibile, allora esso é determinato se e solo se risulta $U = \{\mathbf{0}\}$, (cioé se il sistema omogeneo $A\mathbf{x} = \mathbf{0}$ ha solo la soluzione nulla). La tesi discende allora dalla (1) e (2).

Dim. di (6) e (7). Per quanto visto nella Osservazione 7.4 del Capitolo I, (a meno di uno scambio di colonne, cioé a meno di uno scambio di posizione delle incognite), i sistemi $A\mathbf{x} = \mathbf{0}$ e $A\mathbf{x} = \mathbf{b}$ sono equivalenti rispettivamente ai sistemi $A'\mathbf{x} = \mathbf{0}$ ed $A'\mathbf{x} = \mathbf{b}'$, dove A' ed $(A', \mathbf{b}')$ sono state ottenute da A ed $(A, \mathbf{b})$ eseguendo l'algoritmo di Gauss Jordan completo.

Pertanto la matrice A' puó essere decomposta a blocchi in una delle forme

$$I_k = I_n, \qquad (\, I_k \quad C \,), \qquad \begin{pmatrix} I_n \\ O \end{pmatrix}, \qquad \begin{pmatrix} I_r & C \\ O & O \end{pmatrix}.$$

Primo caso: A é diventata $A' = I_k = I_n$ e $(A, \mathbf{b})$ é diventata $(I_k, \mathbf{b}')$.

In questo caso, si ha $car(A, \mathbf{b}) = car(A) = k = n$. Inoltre il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema $I_n\mathbf{x} = \mathbf{0}$ e quindi ha solo la soluzione nulla; invece il sistema lineare non omogeneo $A\mathbf{x} = \mathbf{b}$ é equivalente al sistema $I_n\mathbf{x} = \mathbf{b}'$, che ha l'unica soluzione $\mathbf{x} = \mathbf{b}'$.

Secondo caso: A é diventata $A' = (I_k \quad C)$ e $(A \quad \mathbf{b})$ é diventata $(I_k \quad C \quad \mathbf{b}')$.

In questo caso, si ha $car(A, \mathbf{b}) = car(A) = k < n$. Inoltre, a meno di uno scambio di posizione delle incognite, il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ é equivalente al sistema

$$(*) \qquad \begin{cases} x_1 = -c_{1,1}x_{k+1} - c_{1,2}x_{k+2} - \ldots - c_{1,n-k}x_n, \\ x_2 = -c_{2,1}x_{k+1} - c_{2,2}x_{k+2} - \ldots - c_{2,n-k}x_n, \\ \cdots\cdots\cdots \\ x_k = -c_{k,1}x_{k+1} - c_{k,2}x_{k+2} - \ldots - c_{k,n-k}x_n. \end{cases}$$

Il sistema ha quindi infinite soluzioni che si ottengono attribuendo valori arbitrari alle variabili $x_{k+1}, x_{k+2}, \ldots, x_n$, e trovando i corrispondenti valori di $x_1, x_2, \ldots, x_k$.
In particolare, una prima soluzione non banale $\mathbf{u}_1$ si ottiene attribuendo valore 1 ad x_{k+1}, valore 0 alle variabili $x_{k+2}, \ldots, x_n$ e trovando i corrispondenti valori di $x_1, x_2, \ldots, x_k$, cioé

$$\mathbf{u}_1 = (-c_{1,1},\ -c_{2,1},\ \ldots,\ -c_{k,1}, 1, 0, \ldots, 0)^T.$$

Attribuendo invece valore 1 ad x_{k+2} e valore 0 alle variabili $x_{k+1}, x_{k+3}, \ldots, x_n$ e trovando i corrispondenti valori di $x_1, x_2, \ldots, x_k$, si ottiene una seconda soluzione non banale

$$\mathbf{u}_2 = (-c_{1,2},\ -c_{2,2},\ \ldots,\ -c_{k,2}, 0, 1, 0, \ldots, 0)^T.$$

In modo analogo si trovano le soluzioni

$$\begin{aligned}
&\mathbf{u}_3 = (-c_{1,3},\ -c_{2,3},\ \ldots,\ -c_{k,3},\ 0, 0, 1, 0, \ldots, 0)^T,\\
&\mathbf{u}_4 = (\ -c_{1,4},\ -c_{2,4},\ \ldots,\ -c_{k,4},\ 0, 0, 0, 1, 0, \ldots, 0)^T,\\
&\ldots\ldots\ldots\\
&\mathbf{u}_{n-k} = (-c_{1,n-k},\ -c_{2,n-k},\ \ldots,\ -c_{k,n-k}, 0, 0, \ldots, 0, 1)^T.
\end{aligned}$$

Tali soluzioni sono linearmente indipendenti, poiché la matrice avente per colonne tali vettori é la matrice $\begin{pmatrix} -C \\ I_{n-k} \end{pmatrix}$, che ha caratteristica $n-k$, poiché le ultime $n-k$ righe sono linearmente indipendenti.
Infine le soluzioni $\mathbf{x}$ del sistema $A\mathbf{x} = \mathbf{0}$ sono tutti e soli i vettori $\mathbf{x}$ ottenuti da $(*)$, attribuendo valori arbitrari $\alpha_1, \alpha_2, \ldots, \alpha_{n-k}$ alle variabili $x_{k+1}, x_{k+2}, \ldots, x_n$ e trovando i corrispondenti valori di $x_1, x_2, \ldots x_k$, cioé i vettori

$$\mathbf{x} = \begin{pmatrix} -c_{1,1}\alpha_1 - c_{1,2}\alpha_2 - \ldots - c_{1,n-k}\alpha_{n-k} \\ -c_{2,1}\alpha_1 - c_{2,2}\alpha_2 - \ldots - c_{2,n-k}\alpha_{n-k} \\ \ldots\ldots\ldots \\ -c_{k,1}\alpha_1 - c_{k,2}\alpha_2 - \ldots - c_{k,n-k}\alpha_{n-k} \\ \alpha_1 \quad + \quad 0\alpha_2 + \ldots + 0\alpha_{n-k} \\ 0\alpha_1 \quad + \quad \alpha_2 + \ldots + 0\alpha_{n-k} \\ 0\alpha_1 \quad + \quad 0\alpha_2 + \ldots + \alpha_{n-k} \end{pmatrix} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_{n-k}\mathbf{u}_{n-k}.$$

D'altra parte il sistema lineare non omogeneo $A\mathbf{x} = \mathbf{b}$ é equivalente al sistema

$$(**) \qquad \begin{cases} x_1 = b'_1 - c_{1,1}x_{k+1} - c_{1,2}x_{k+2} - \ldots - c_{1,n-k}x_n \\ x_2 = b'_2 - c_{2,1}x_{k+1} - c_{2,2}x_{k+2} - \ldots - c_{2,n-k}x_n \\ \ldots\ldots\ldots \\ x_k = b'_k - c_{k,1}x_{k+1} - c_{k,2}x_{k+2} - \ldots - c_{k,n-k}x_n \end{cases}.$$

Pertanto, le soluzioni $\mathbf{x}$ del sistema $A\mathbf{x} = \mathbf{b}$ sono tutti e soli i vettori $\mathbf{x}$ ottenuti da $(**)$, attribuendo valori arbitrari $\alpha_1, \alpha_2, \ldots, \alpha_{n-k}$ alle variabili $x_{k+1}, x_{k+2}, \ldots, x_n$, cioé tutti e soli i vettori del tipo

$$\mathbf{x} = \begin{pmatrix} b'_1 - c_{1,1}\alpha_1 - c_{1,2}\alpha_2 - \ldots - c_{1,n-k}\alpha_{n-k} \\ b'_2 - c_{2,1}\alpha_1 - c_{2,2}\alpha_2 - \ldots - c_{2,n-k}\alpha_{n-k} \\ \ldots\ldots\ldots \\ b'_k - c_{k,1}\alpha_1 - c_{k,2}\alpha_2 - \ldots - c_{k,n-k}\alpha_{n-k} \\ 0 \ + \ \alpha_1 + \ 0\alpha_2 \ + \ldots + \ 0\alpha_{n-k} \\ 0 \ + \ 0\alpha_1 + \ \alpha_2 \ + \ldots + \ 0\alpha_{n-k} \\ 0 \ + \ 0\alpha_1 + \ 0\alpha_2 \ + \ldots + \ \alpha_{n-k} \end{pmatrix} = \mathbf{x}^* + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 \quad + \ldots + \alpha_{n-k}\mathbf{u}_{n-k},$$

dove $\mathbf{x}^* = (b'_1, b'_2, \dots, b'_k, 0, 0, \dots, 0)^T$ é la soluzione ottenuta attribuendo valore nullo a tutte le variabili $x_{k+1}, x_{k+2}, \dots, x_n$.

Terzo caso: A é diventata $A' = \begin{pmatrix} I_n \\ O \end{pmatrix}$, e quindi $car(A) = n < k$.

Questa volta, per ogni $i > n$ la $i-$sima equazione del sistema omogeneo $A'\mathbf{x} = \mathbf{0}$ é $0x_1 + 0x_2 + \ldots + 0x_n = 0$; evidentemente tale equazione é soddisfatta da ogni vettore $\mathbf{x}$ di $\mathbf{R}^n$ e quindi é superflua e puó essere eliminata dal sistema. Eliminate tali equazioni, si ottiene il sistema $I_n\mathbf{x} = \mathbf{0}$, e quindi il sistema omogeneo ha solo la soluzione nulla.

Per quanto riguarda invece il sistema non omogeneo, per ogni $i > n$ la $i-$esima equazione del sistema é diventata

$$0x_1 + 0x_2 + \ldots + 0x_n = b'_i.$$

Ebbene, se risulta $b'_i \neq 0$, allora la $i-$esima equazione é incompatibile e quindi il sistema é incompatibile.

Se invece risulta $b'_i = 0$ per ogni $i > n$, allora le ultime $k - n$ equazioni del sistema sono superflue; eliminate tali equazioni, si ottiene il sistema $I_n\mathbf{x} = (b'_1, b'_2, \dots, b'_n)^T$, che ha l'unica soluzione $\mathbf{x} = (b'_1, b'_2, \dots, b'_n)^T$.

Pertanto il sistema é determinato o incompatibile a seconda che sia

$b'_i = 0$ per ogni $i > n$ oppure $b'_i \neq 0$ per qualche $i > n$,

e queste condizioni corrispondono esattamente a chiedere che

$car(A, \mathbf{b}) = n = car(A)$ oppure $car(A, \mathbf{b}) = n + 1 > n = car(A)$.

Quarto caso: A é diventata $A' = \begin{pmatrix} I_r & C \\ O & O \end{pmatrix}$, e quindi $car(A) = r < \min(k, n)$

Questa volta le ultime $k - r$ equazioni del sistema lineare omogeneo $A'\mathbf{x} = \mathbf{0}$ saranno superflue e il sistema si trasforma in un sistema lineare omogeneo di r equazioni in n incognite del tipo $(I_r \quad C)\mathbf{x} = \mathbf{0}_r$. Con lo stesso procedimento illustrato nello studio del "secondo caso", di tale sistema si trovano $n - r$ soluzioni linearmente indipendenti $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{n-r}$, tali che tutte le altre soluzioni sono combinazioni lineari di esse.

Per quanto riguarda invece il sistema non omogeneo $A'\mathbf{x} = \mathbf{b}'$, si possono ripetere le considerazioni fatte nello studio del "Terzo caso" e concludere che il sistema é compatibile se e solo se risulta $b'_i = 0$ per ogni $i > r$. In tal caso le ultime $k - r$ equazioni del sistema sono superflue; eliminate tali equazioni, si ottiene un sistema di r equazioni in n incognite, in cui la matrice dei coefficienti é del tipo $(I_r \quad C)$, che si risolve come visto nel "secondo caso".

Il sistema é dunque indeterminato o incompatibile a seconda che sia

$b'_i = 0$ per ogni $i > r$ oppure $b'_i \neq 0$ per qualche $i > r$,

e queste condizioni corrispondono esattamente a chiedere che

$car(A, \mathbf{b}) = r = car(A)$ oppure $car(A, \mathbf{b}) = r + 1 > r = car(A)$.

Inoltre, le soluzioni del sistema nonomogeneo sono del tipo

$$\mathbf{x} = \mathbf{x}^* + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_{n-r}\mathbf{u}_{n-r}, \qquad \text{con} \qquad \mathbf{x}^* = (b'_1, b'_2, \dots, b'_r, 0, 0, \dots, 0)^T.$$

In definitiva, a conferma di quanto visto al punto (1), il sistema lineare omogeneo ha solo la soluzione nulla solo nel primo e terzo caso, cioé se e solo se $car(A) = n$. Invece nel secondo e quarto caso l'insieme delle soluzioni del sistema lineare omogeneo é un sottospazio vettoriale di dimensione $n - car(A)$.

Per quanto riguarda il sistema non omogeneo, a conferma di quanto visto al punto (2), abbiamo visto che il sistema é compatibile se e solo se $car(A, \mathbf{b}) = car(A)$, ed é determinato nel primo e terzo caso, (cioé quando $car(A) = n$), é indeterminato nel secondo e quarto caso, (cioé quando $car(A) < n$); in tale ultimo caso l'insieme delle soluzioni del sistema é esattamente come previsto dalla (5).

■

Osservazione 1.2 - In particolare l'insieme delle soluzioni di una singola equazione lineare omogenea

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$$

con coefficienti $a_1, a_2, \ldots, a_n$ non tutti nulli é un sottospazio vettoriale di $\mathbf{R}^n$ di dimensione $n-1$, cioé un iperpiano in $\mathbf{R}^n$.
Analogamente l'insieme delle soluzioni di una singola equazione lineare non omogenea

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

con $a_1, a_2, \ldots, a_n$ non tutti nulli, é un sottospazio affine di dimensione $n-1$, cioé un iperpiano affine.
Di conseguenza l'insieme delle soluzioni di un sistema lineare non omogeneo di k equazioni in n incognite puó essere visto come l'intersezione di k iperpiani affini di $\mathbf{R}^n$.

■

Esempio 1 - Risolvere, al variare del parametro k, il sistema

$$\begin{cases} x_1 + 2x_2 + 3x_3 & = 1, \\ 2x_1 - x_2 + (1-k)x_3 + 2x_4 & = -3, \\ 3x_1 - 4x_2 + (1-k)x_3 + k^2x_4 & = 0, \\ kx_1 + 2kx_2 + 3kx_3 + (k^2-4)x_4 & = 2. \end{cases}$$

Consideriamo la matrice $\hat{A} = (A, \mathbf{b})$ dei coefficienti e termini noti

$$\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 2 & 3 & 0 & 1 \\ 2 & -1 & 1-k & 2 & -3 \\ 3 & -4 & 1-k & k^2 & 0 \\ k & 2k & 3k & k^2-4 & 2 \end{array}$$

ed effettuiamo un primo passo di pivot sull' elemento $a_{1,1} = 1$, cioé dividiamo per 1 la prima riga e sostituiamo la seconda riga $\hat{A}_2$ con $\hat{A}_2 - 2\hat{A}_1$, la terza riga $\hat{A}_3$ con $\hat{A}_3 - 3\hat{A}_1$ ed $\hat{A}_4$ con $\hat{A}_4 - k\hat{A}_1$. Si ottiene allora la matrice

$$\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 2 & 3 & 0 & 1 \\ 0 & -5 & -5-k & 2 & -5 \\ 0 & -10 & -8-k & k^2 & -3 \\ 0 & 0 & 0 & k^2-4 & 2-k. \end{array}$$

Effettuando un secondo passo di pivot sull' elemento $a_{2,2} = -5$, cioé dividendo per -5 la seconda riga e sostituendo $\hat{A}_1$ con $\hat{A}_1 + (2/5)\hat{A}_2$, ed $\hat{A}_3$ con $\hat{A}_3 - 2\hat{A}_2$, si ottiene la matrice

$$(*) \qquad \begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 1 & 0 & 1-(2/5)k & 4/5 & -1 \\ 0 & 1 & 1+(k/5) & -2/5 & 1 \\ 0 & 0 & 2+k & k^2-4 & 7 \\ 0 & 0 & 0 & k^2-4 & 2-k. \end{array}$$

A questo punto distinguiamo i casi : $k \neq \pm 2$, $k = 2$ e $k = -2$.

Primo caso: k diverso da 2 e da -2.

In questo caso, potremo effettuare un passo di pivot dapprima su $a_{3,3} = k+2$ e poi su $a_{4,4} = k^2 - 4$ ed otterremo le matrici

$$\begin{array}{cccc|c}
x_1 & x_2 & x_3 & x_4 & \\
\hline
1 & 0 & 0 & (2k^2-9k+14)/5 & (9k-45)/(5k+10) \\
0 & 1 & 0 & k^2+3k-8 & (2k+25)/(k+2) \\
0 & 0 & 1 & k-2 & 7/(k+2) \\
0 & 0 & 0 & k^2-4 & 2-k
\end{array},$$

$$\begin{array}{cccc|c}
x_1 & x_2 & x_3 & x_4 & \\
\hline
1 & 0 & 0 & 0 & (2k^2-31)/(5k+10) \\
0 & 1 & 0 & 0 & -(k^2+5k+17)/(5k+10) \\
0 & 0 & 1 & 0 & (k+5)/(k+2) \\
0 & 0 & 0 & 1 & -1/(k+2).
\end{array}$$

E' evidente quindi che il sistema omogeneo ha solo la soluzione nulla, mentre il sistema non omogeneo ha l'unica soluzione

$$\mathbf{x} = \left(\frac{2k^2-31}{5k+10}, \; -\frac{k^2+5k+17}{5k+10}, \; \frac{k+5}{k+2}, \; -\frac{1}{k+2} \right)^T.$$

Questa volta la matrice dei coefficienti e la matrice completa hanno caratteristica 4, uguale al numero delle incognite e il Teorema 1.1 prevede appunto che il sistema omogeneo abbia solo la soluzione banale e il sistema non omogeneo sia determinato.

Secondo caso: $k = 2$. Questa volta la matrice dei coefficienti e termini noti del sistema é diventata

$$\begin{array}{cccc|c}
x_1 & x_2 & x_3 & x_4 & \\
\hline
1 & 0 & 1/5 & 4/5 & -1 \\
0 & 1 & 7/5 & -2/5 & 1 \\
0 & 0 & 4 & 0 & 7 \\
0 & 0 & 0 & 0 & 0.
\end{array}$$

Effettuando un passo di pivot su $a_{3,3} = 4$, cioé dividendo per 4 la terza riga e sostituendo $\hat{A}_1$ con $\hat{A}_1 - (1/20)\hat{A}_3$, ed $\hat{A}_2$ con $\hat{A}_2 - (7/20)\hat{A}_3$, la matrice diventa

$$\begin{array}{cccc|c}
x_1 & x_2 & x_3 & x_4 & \\
\hline
1 & 0 & 0 & 4/5 & -27/20 \\
0 & 1 & 0 & -2/5 & -29/20 \\
0 & 0 & 1 & 0 & 7/4 \\
0 & 0 & 0 & 0 & 0
\end{array}.$$

Questa volta si ha che $car(A) = car(A, \mathbf{b}) = 3$ minore del numero delle incognite e quindi ci aspettiamo che il sistema non omogeneo sia indeterminato, che l'insieme delle soluzioni del sistema omogeneo sia un sottospazio di dimensione $4 - car(A) = 4 - 3 = 1$ e che l'insieme delle soluzioni del sistema non omogeneo sia un sottospazio affine di dimensione 1.

In effetti l'ultima equazione é diventata $0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$, e quindi é superflua e puó essere eliminata. Dalle prime tre si ricava invece che il sistema omogeneo e quello non omogeneo sono rispettivamente equivalenti ai sistemi

$$\begin{cases} x_1 = -(4/5)x_4 \\ x_2 = (2/5)x_4 \\ x_3 = 0 \end{cases} \qquad \begin{cases} x_1 = -27/20 - (4/5)x_4 \\ x_2 = -29/20 + (2/5)x_4 \\ x_3 = 7/4. \end{cases}$$

Attribuendo un valore arbitrario α ad x_4 si trovano i corrispondenti valori di x_1, x_2, x_3 e dunque il sistema omogeneo ha le infinite soluzioni

$$\mathbf{x} = \begin{pmatrix} -(4/5)\alpha \\ (2/5)\alpha \\ 0 \\ \alpha \end{pmatrix} = \alpha\mathbf{u}, \quad \text{con} \quad \mathbf{u} = \begin{pmatrix} -4/5 \\ 2/5 \\ 0 \\ 1 \end{pmatrix} \quad \text{ed } \alpha \text{ numero reale arbitrario .}$$

L'insieme delle soluzioni del sistema omogeneo é quindi un sottospazio di dimensione 1.
Invece le soluzioni del sistema non omogeneo sono tutti e soli i vettori della forma

$$\mathbf{x} = \begin{pmatrix} -27/20 - (4/5)\alpha \\ -29/20 + (2/5)\alpha \\ 7/4 \\ \alpha \end{pmatrix} = \begin{pmatrix} -27/20 \\ -29/20 \\ 7/4 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -4/5 \\ 2/5 \\ 0 \\ 1 \end{pmatrix} = \mathbf{x}^* + \alpha\mathbf{u}$$

con $\mathbf{x}^* = (-27/20, -29/20, 7/4, 0)^T$ soluzione del sistema non omogeneo, $\mathbf{u} = (-4/5, 2/5, 0, 1)^T$ soluzione non nulla del sistema omogeneo ed α numero reale arbitrario.

Terzo caso: $k = -2$. In questo caso la matrice $(A, \mathbf{b})$ é diventata

x_1	x_2	x_3	x_4	
1	0	9/5	4/5	−1
0	1	3/5	−2/5	1
0	0	0	0	7
0	0	0	0	4,

cioé le ultime due equazioni sono diventate

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 7, \qquad 0x_1 + 0x_2 + 0x_3 + 0x_4 = 4,$$

che sono chiaramente incompatibili e dunque il sistema é incompatibile.
In effetti, in questo caso la matrice A dei coefficienti ha caratteristica 2, e la matrice completa ha invece caratteristica 3, e il Teorema di Rouché Capelli prevede appunto che il sistema sia incompatibile.

Per quanto riguarda il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$, tale sistema sarebbe stato trasformato in un sistema in cui le ultime due equazioni sono

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0, \qquad 0x_1 + 0x_2 + 0x_3 + 0x_4 = 0,$$

e quindi sono superflue. Eliminate tali equazioni, il sistema si riduce a

$$x_1 = -(9/5)x_3 - (4/5)x_4, \qquad x_2 = -(3/5)x_3 + (2/5)x_4,$$

le cui infinite soluzioni si ottengono dando valori arbitrari α_1, α_2 ad x_3, x_4 e ricavando i corrispondenti valori di x_1 ed x_2; le soluzioni sono dunque i vettori $\mathbf{x}$ del tipo:

$$\mathbf{x} = \begin{pmatrix} -(9/5)\alpha_1 - (4/5)\alpha_2 \\ -(3/5)\alpha_1 + (2/5)\alpha_2 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} -9/5 \\ -3/5 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -4/5 \\ 2/5 \\ 0 \\ 1 \end{pmatrix} \qquad \text{con } \alpha_1, \alpha_2 \text{ arbitrari.}$$

L'insieme delle soluzioni del sistema omogeneo é quindi il sottospazio vettoriale di dimensione $2 = n - car(A)$ generato dai vettori linearmente indipendenti $\mathbf{u}_1 = (-9/5,\ -3/5,\ 1,\ 0)^T$ ed $\mathbf{u}_2 = (-4/5,\ 2/5,\ 0,\ 1)^T$.

■

Esempio 2. - Risolvere il sistema

$$\begin{cases} x + ky + z = k - 2 \\ 2x - y + 3z = k^2 + 1 \\ kx + 4y + kz = 0 \\ (1 + 2k)y - z = -5. \end{cases}$$

Consideriamo la matrice completa

$$\begin{array}{ccc|c} x & y & z & \\ \hline 1 & k & 1 & k-2 \\ 2 & -1 & 3 & k^2+1 \\ k & 4 & k & 0 \\ 0 & 1+2k & -1 & -5. \end{array}$$

Effettuando un passo di pivot su $a_{1,1} = 1$ tale matrice diventa

$$\begin{array}{ccc|c} x & y & z & \\ \hline 1 & k & 1 & k-2 \\ 0 & -1-2k & 1 & k^2-2k+5 \\ 0 & 4-k^2 & 0 & -k^2+2k \\ 0 & 1+2k & -1 & -5. \end{array}$$

Effettuando un passo di pivot su $a_{2,3} = 1$ e scambiando la seconda con la terza colonna si ottiene la matrice

$$\begin{array}{ccc|c} x & z & y & \\ \hline 1 & 0 & 1+3k & -k^2+3k-7 \\ 0 & 1 & -1-2k & k^2-2k+5 \\ 0 & 0 & 4-k^2 & -k^2+2k \\ 0 & 0 & 0 & k^2-2k. \end{array}$$

Se $4 - k^2 \neq 0$ e $k^2 - 2k \neq 0$, cioé se k é diverso da $0, 2$ e -2, il sistema é incompatibile poiché A ha rango 3, mentre $\hat{A}$ ha rango 4. (In effetti l' ultima equazione é diventata $0x + 0y + 0z = k^2 - 2k$ che é chiaramente incompatibile).

Invece il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ ha solo la soluzione banale, poiché l'ultima equazione é diventata $0x + 0y + 0z = 0$ e quindi é superflua, e dalle prime tre equazioni si ricava

$$\begin{cases} x = 0 - (1 + 3k)y \\ z = 0 + (1 + 2k)y \\ (4 - k^2)y = 0 \end{cases} \quad \text{e quindi} \quad \begin{cases} y = 0 \\ x = 0 \\ z = 0. \end{cases}$$

Del resto questo risultato é coerente con la (1) del Teorema 1.3, poiché risulta $car(A) = 3 = n$.

Nel caso $k = 0$, la matrice dei coefficienti e termini noti é diventata

$$\begin{array}{ccc|c} x & z & y & \\ \hline 1 & 0 & 1 & -7 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array} \quad \text{e quindi} \quad \begin{array}{ccc|c} x & z & y & \\ \hline 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

dopo un ulteriore passo di pivot su $a_{3,3}$. Quindi A ed $\hat{A} = (A, \mathbf{b})$ hanno entrambe caratteristica 3; pertanto il sistema é determinato. In effetti l' ultima equazione é superflua e dalle altre tre si ricava l' unica soluzione $(x, y, z)^T = (-7, 0, 5)^T$.

Si noti come in questo caso il sistema omogeneo avrebbe solo la soluzione nulla, coerentemente con quanto previsto dalla (1) del Teorema 1.1.

Se $k=-2$ la matrice dei coefficienti e termini noti é diventata

$$\begin{array}{ccc|c} x & z & y & \\ \hline 1 & 0 & -5 & -17 \\ 0 & 1 & 3 & 13 \\ 0 & 0 & 0 & -8 \\ 0 & 0 & 0 & 8. \end{array}$$

Quindi A ha rango 2, mentre $\hat{A}$ ha rango 3; il sistema é perció incompatibile. In effetti le ultime due equazioni sono diventate

$$0x+0y+0z=-8 \qquad \text{e} \qquad 0x+0y+0z=8$$

che sono chiaramente incompatibili.

Invece il sistema omogeneo sarebbe stato trasformato in un sistema in cui le ultime due equazioni sono superflue e le soluzioni del sistema si ricavano quindi dalle prime due equazioni: $x-5y=0$ e $z+3y=0$. L'insieme di tali soluzioni é quindi l'insieme dei vettori del tipo

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5\alpha \\ \alpha \\ -3\alpha \end{pmatrix} = \alpha \begin{pmatrix} 5 \\ 1 \\ -3 \end{pmatrix} \qquad \text{con } \alpha \in \mathbf{R} \text{ arbitrario}$$

ed é quindi un sottospazio vettoriale di dimensione $1 = n - car(A)$.

Infine se $k=2$ la matrice é diventata

$$\begin{array}{ccc|c} x & z & y & \\ \hline 1 & 0 & 7 & -5 \\ 0 & 1 & -5 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

Questa volta A ed $\hat{A}$ hanno entrambe rango $2 < n = 3$ e quindi l'insieme U delle soluzioni del sistema omogeneo é un sottospazio vettoriale di dimensione $n - car(A) = 1$ e l'insieme X delle soluzioni di quello non omogeneo é un sottospazio affine di dimensione 1 parallelo ad U.

In effetti le ultime due equazioni sono superflue e dalle prime due si ricava che il sistema omogeneo ha le infinite soluzioni:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -7\alpha \\ \alpha \\ 5\alpha \end{pmatrix} = \alpha \begin{pmatrix} -7 \\ 1 \\ 5 \end{pmatrix} \qquad \text{con } \alpha \text{ numero reale arbitrario}$$

e quello non omogeneo ha le infinite soluzioni

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -5-7\alpha \\ \alpha \\ 5+5\alpha \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 5 \end{pmatrix} + \alpha \begin{pmatrix} -7 \\ 1 \\ 5 \end{pmatrix} \qquad \text{con } \alpha \text{ numero reale arbitrario.}$$

■

E S E R C I Z I

n. 1 - Data la matrice $A = \begin{pmatrix} 1 & 0 & \alpha & \alpha+2 \\ 0 & 1 & 2 & \alpha+1 \\ -1 & 2 & 3 & 1 \\ \alpha & 0 & \alpha^2 & 2\alpha+1 \end{pmatrix}$ e il vettore $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -\alpha \\ -1 \end{pmatrix}$,

(1) calcolare il determinante di A e (al variare del parametro α), la caratteristica delle matrici A e $(A, \mathbf{b})$,
(2) risolvere il sistema lineare omogeneo $Ax = 0$,
(3) dire per quale valore del parametro α il sistema lineare $A\mathbf{x} = \mathbf{b}$ é determinato, indeterminato o incompatibile; se il sistema é compatibile, trovare la o le soluzioni.

n. 2 - Dire se sono determinati, indeterminati o incompatibili i sistemi

$$\begin{cases} x+2y-z = 0 \\ y+2z = 1 \\ 3x+5y-5z = -1 \end{cases}, \qquad \begin{cases} x-2y+z = 0 \\ x-2y+2z = 1 \\ -2x+4y-z = -1 \end{cases},$$

(indicando esplicitamente la caratteristica della matrice A dei coefficienti e della matrice completa $\hat{A}$), e trovarne le eventuali soluzioni.

Dire se il sistema lineare omogeneo associato a tali sistemi ha soluzioni non banali, e in tal caso trovare tali soluzioni.

n. 3 - Dire per quali valori del parametro α i sistemi lineari

$$\begin{cases} x-y-2z = 4 \\ 2x-y+\alpha^2 z = \alpha-6 \\ -3x+2y-\alpha z = 0 \end{cases} \qquad \begin{cases} x + \alpha z = \alpha+2 \\ y+2z = \alpha+1 \\ -x+2y+3z = 1 \\ \alpha x + \alpha^2 z = 2\alpha+1 \end{cases} \qquad \begin{cases} x+3y+\alpha z-t = 0 \\ 3x+8y-3t = 4 \\ \alpha x+(1+3\alpha)y+4\alpha z = \alpha-2, \end{cases}$$

$$\begin{cases} 4x-\alpha y-z = 1 \\ \alpha x-y+3z = 2 \\ -x+y-3z = -1 \end{cases} \qquad \begin{cases} x-2y+\alpha z+2t = 0 \\ -2x+5y+t = 4 \\ -\alpha x+(1+2\alpha)y+4\alpha z = \alpha-2, \end{cases}$$

sono incompatibili, determinati o indeterminati, indicando esplicitamente la caratteristica della matrice A dei coefficienti e della matrice completa $\hat{A}$.

Se il sistema è compatibile, trovarne la o le soluzioni; trovare infine le soluzioni del sistema omogeneo associato a tali sistemi.

n. 4 - Ripetere il precedente esercizio per i sistemi

$$\begin{cases} x+3y+\alpha z-t = 0 \\ 3x+8y-3t = 4 \\ \alpha x+(1+3\alpha)y+4\alpha z = \alpha-2, \end{cases} \qquad \begin{cases} 4x-\alpha y-z = 1 \\ \alpha x-y+3z = 2 \\ -x+y-3z = -1 \end{cases}$$

n. 5 - Date le matrici

$$M = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ -2 & 1-2\alpha & -2 & \alpha+2 \\ -1 & 1-\alpha & -1 & 2\alpha \\ 0 & -1 & 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & \alpha & 0 \\ -2 & 1-2\alpha & -2 \\ -1 & 1-\alpha & -1 \\ 0 & -1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ \alpha+2 \\ 2\alpha \\ 0 \end{pmatrix},$$

a) calcolare il determinante di M,
b) trovare (al variare del parametro α) la caratteristica delle matrici M ed A,
c) dire per quali valori di α il sistema $Ax = b$ è incompatibile, determinato o indeterminato.
Se il sistema è compatibile, trovarne la o le soluzioni, indicando anche esplicitamente le soluzioni del sistema lineare omogeneo $Ax = 0$.

n. 6 - Ripetere il precedente esercizio nel caso:

$$M = \begin{pmatrix} 1 & 0 & \alpha & \alpha+2 \\ 0 & 1 & 2 & \alpha+1 \\ 0 & 2 & \alpha+3 & \alpha+3 \\ \alpha & 0 & \alpha^2 & 2\alpha+1 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 2 \\ 0 & 2 & \alpha+3 \\ \alpha & 0 & \alpha^2 \end{pmatrix}, \qquad b = \begin{pmatrix} \alpha+2 \\ \alpha+1 \\ \alpha+3 \\ 2\alpha+1 \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 0 & \alpha & \alpha+2 \\ 0 & 1 & 2 & \alpha+1 \\ -1 & 2 & 3 & 1 \\ \alpha & 0 & \alpha^2 & 2\alpha+1 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 2 \\ -1 & 2 & 3 \\ \alpha & 0 & \alpha^2 \end{pmatrix}, \qquad b = \begin{pmatrix} \alpha+2 \\ \alpha+1 \\ 1 \\ 2\alpha+1 \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 0 & \alpha-4 & 2-\alpha \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 2\alpha & \alpha^2 \\ \alpha & -2 & (\alpha-2)^2 & 2 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 & \alpha-4 \\ 0 & 1 & -2 \\ 0 & 0 & 2\alpha \\ \alpha & -2 & (\alpha-2)^2 \end{pmatrix}, \qquad b = \begin{pmatrix} 2-\alpha \\ -1 \\ \alpha^2 \\ 2 \end{pmatrix},$$

Capitolo V - TRASFORMAZIONI LINEARI

1. - Definizioni e prime proprietá

Definizione 1.1 -. *Dicesi* **applicazione lineare** *o* **trasformazione lineare** *di* $\mathbf{R}^n$ *in* $\mathbf{R}^k$ *una qualunque funzione* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^k$ *che soddisfa le condizioni:*

a) $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$ *per ogni* $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$*,*
b) $L(\alpha\mathbf{v}) = \alpha L(\mathbf{v})$ *per ogni* $\mathbf{v} \in \mathbf{R}^n$ *e per ogni* $\alpha \in \mathbf{R}$ *.*

■

Osservazione 1.2 - L'espressione "trasformazione lineare" é dovuta al fatto che una trasformazione lineare L trasforma:

(1) l'origine $\mathbf{0}_n$ di $\mathbf{R}^n$ nell'origine $\mathbf{0}_k$ di $\mathbf{R}^k$, poiché risulta $L(\mathbf{0}_n) = L(0 \cdot \mathbf{0}_n) = 0 \cdot L(\mathbf{0}_n) = \mathbf{0}_k$;
(2) due punti $\mathbf{u}$ e $\mathbf{v} = \alpha\mathbf{u}$ allineati con l'origine in due punti $L(\mathbf{u})$ e $L(\mathbf{v}) = \alpha L(\mathbf{u})$ allineati con l'origine, e quindi trasforma la retta di $\mathbf{R}^n$ di direzione $\mathbf{u}$ nella retta di $\mathbf{R}^k$ di direzione $L(\mathbf{u})$;
(3) il quarto vertice $\mathbf{u}+\mathbf{v}$ del parallelogramma di vertici $\mathbf{0}_n, \mathbf{u}, \mathbf{v}$ nel quarto vertice $L(\mathbf{u}) + L(\mathbf{v})$ del parallelogramma di vertici $L(\mathbf{0}_n) = \mathbf{0}_k, L(\mathbf{u}), L(\mathbf{v})$;
(4) la retta affine passante per $\mathbf{a}$ parallela ad $\mathbf{u}$ nella retta affine passante per $L(\mathbf{a})$ parallela ad $L(\mathbf{u})$;
(5) due rette affini parallele in due rette affini parallele;
(6) il segmento di estremi $\mathbf{a}$ e $\mathbf{b}$ nel segmento di estremi $L(\mathbf{a})$ ed $L(\mathbf{b})$;
(7) il parallelogramma di vertici $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ nel parallelogramma di vertici $L(\mathbf{a}), L(\mathbf{b}), L(\mathbf{c}), L(\mathbf{d})$.

Naturalmente, in generale, L non conserva l'angolo tra due vettori, cosí come non conserva la norma o lunghezza dei vettori, né la distanza tra due punti.

■

Ad esempio, non é una trasformazione lineare la funzione L che associa ad ogni vettore $\mathbf{x} = (x_1, x_2, x_3)^T$ di $\mathbf{R}^3$ il vettore $\mathbf{y} = (y_1, y_2)^T \in \mathbf{R}^2$ tale che $y_1 = x_1 + 2x_2 - 1$ e $y_2 = x_2 - x_3$, poiché risulta $L(0,0,0) = (-1,0) \neq (0,0)$.

Analogamente non é una trasformazione lineare la funzione che associa ad ogni vettore $\mathbf{x} = (x_1, x_2)^T \in \mathbf{R}^2$ il vettore $\mathbf{y} = (y_1, y_2)^T \in \mathbf{R}^2$ tale che $y_1 = x_1 + x_2$ e $y_2 = x_1 x_2$.

Infatti, ad esempio, essa trasforma la retta di equazione $x_2 = x_1$ nella curva di equazioni parametriche $y_1 = 2x_1$ ed $y_2 = x_1^2$, cioé nella parabola di equazione $y_2 = x_1^2 = y_1^2/4$.

■

Esempio 1 - La funzione L che ad ogni $\mathbf{x} \in \mathbf{R}^n$ associa il vettore $\mathbf{0}_k \in \mathbf{R}^k$ é una trasformazione lineare tra $\mathbf{R}^n$ ed $\mathbf{R}^k$.

Infattti, per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ e per ogni $\alpha \in \mathbf{R}$ si ha: $L(\mathbf{x}) = \mathbf{0}_k,\ L(\mathbf{y}) = \mathbf{0}_k,\ L(\mathbf{x}+\mathbf{y}) = \mathbf{0}_k,\ L(\alpha\mathbf{x}) = \mathbf{0}_k$, e quindi $L(\mathbf{x}) + L(\mathbf{y}) = \mathbf{0}_k + \mathbf{0}_k = \mathbf{0}_k = L(\mathbf{x}+\mathbf{y})$, $\alpha \cdot L(\mathbf{x}) = \alpha \cdot \mathbf{0}_k = \mathbf{0}_k = L(\alpha\mathbf{x})$.

■

Esempio 2 - La funzione L che ad ogni $\mathbf{x} \in \mathbf{R}^n$ associa il vettore $L(\mathbf{x}) = \mathbf{x} \in \mathbf{R}^n$ é una trasformazione lineare di $\mathbf{R}^n$ in sé, detta **trasformazione identica**.

Infattti si ha: $L(\mathbf{x}+\mathbf{y}) = \mathbf{x}+\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y})$ ed $L(\alpha \cdot \mathbf{x}) = \alpha \cdot \mathbf{x} = \alpha \cdot L(\mathbf{x})$. , per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ e per ogni $\alpha \in \mathbf{R}$.

■

Osservazione 1.3 - Se A é una matrice di tipo $k \times n$, allora la funzione L_A che associa ad ogni vettore $\mathbf{x}$ di $\mathbf{R}^n$ il vettore $\mathbf{L}_A(\mathbf{x}) = A\mathbf{x}$ di $\mathbf{R}^k$ é una trasformazione lineare tra $\mathbf{R}^n$ ed $\mathbf{R}^k$, che viene detta la **trasformazione lineare associata alla matrice** A.

Infatti risulta $L_A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = L_A(\mathbf{x}) + L_A(\mathbf{y})$ ed $L_A(\alpha\mathbf{x}) = A(\alpha\mathbf{x}) = \alpha(A\mathbf{x}) = \alpha L_A(\mathbf{x})$, per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ e per ogni $\alpha \in \mathbf{R}$.

■

Osservazione 1.4 - Un esempio particolarmente significativo di trasformazioni lineari é dato dalle **rotazioni del piano**.

Assegnato un riferimento cartesiano nel piano euclideo, fissiamo un numero reale $\theta \in]0, 2\pi[$ e consideriamo la trasformazione L di $\mathbf{R}^2$ in sé che lascia invariata l'origine $(0,0)$ ed associa ad ogni $\mathbf{x} = (x_1, x_2) \neq (0,0)$, (e quindi al punto P di coordinate (x_1, x_2)), il vettore $\mathbf{x}' = (x_1', x_2')$ delle coordinate del punto P' ottenuto facendo ruotare in senso antiorario intorno all'origine il segmento OP di un angolo di ampiezza θ.

Si vede facilmente che L é una trasformazione lineare.

Infatti per ogni $\mathbf{u}, \mathbf{v} \in \mathbf{R}^2$ l'intero parallelogramma di vertici $\mathbf{0}, \mathbf{u}, \mathbf{v}$ ed $\mathbf{u}+\mathbf{v}$ verrá ruotato di un angolo di ampiezza θ intorno all'origine e trasformato nel parallelogramma di vertici $L(\mathbf{0}) = \mathbf{0}, L(\mathbf{u}), L(\mathbf{v})$, ed $L(\mathbf{u}+\mathbf{v})$, sicché risulta $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$.

D'altra parte, se $\mathbf{u}$ e $\mathbf{v}$ sono i vettori delle coordinate di due punti allineati con l'origine, allora i segmenti $[\mathbf{0}, \mathbf{u}]$ e $[\mathbf{0}, \mathbf{v}]$ verranno ruotati entrambi di un angolo di ampiezza θ, e quindi:

- i punti di coordinate $L(\mathbf{u})$ ed $L(\mathbf{v})$ rimarranno allineati con l'origine,
- se i segmenti orientati $[\mathbf{0}, \mathbf{u}]$ e $[\mathbf{0}, \mathbf{v}]$ hanno lo stesso verso (rispett. verso opposto), allora anche $[\mathbf{0}, L(\mathbf{u})]$ e $[\mathbf{0}, L(\mathbf{v})]$ hanno lo stesso verso (rispett. verso opposto);
- il rapporto tra le lunghezze dei segmenti $[\mathbf{0}, \mathbf{v}]$ e $[\mathbf{0}, \mathbf{u}]$ coincide con il rapporto tra le lunghezze dei segmenti $[\mathbf{0}, L(\mathbf{v})]$ ed $[\mathbf{0}, L(\mathbf{u})]$.

Questo mostra che risulta $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u})$ per ogni $\alpha \in \mathbf{R}$ e per ogni $\mathbf{u} \in \mathbf{R}^2$ e quindi che la trasformazione L di $\mathbf{R}^2$ in sé che produce una rotazione intorno all'origine di un angolo di ampiezza θ é una trasformazione lineare.

■

2. - Proprietá delle trasformazioni lineari

Il teorema seguente contiene le principali proprietá delle trasformazioni lineari.

Teorema 2.1 -. *Se $L : \mathbf{R}^n \longrightarrow \mathbf{R}^k$ é una trasformazione lineare, allora si ha quanto segue:*

(1) *per ogni $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r \in \mathbf{R}^n$ e per ogni $\alpha_1, \alpha_2, \dots \alpha_r \in \mathbf{R}$ si ha:*

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r);$$

(2) *se $H \subseteq \mathbf{R}^n$ é un sottospazio vettoriale di $\mathbf{R}^n$, allora l'insieme*

$$L(H) = \{\mathbf{y} \in \mathbf{R}^k | \exists \mathbf{x} \in H \ \textit{tale che}\ \mathbf{y} = L(\mathbf{x})\} \subseteq \mathbf{R}^k$$

é un sottospazio vettoriale di $\mathbf{R}^k$;

(3) *se H' é un sottospazio di $\mathbf{R}^k$, allora l'insieme*

$$L^{-1}(H') = \{\mathbf{x} \in \mathbf{R}^n | L(\mathbf{x}) \in H'\}$$

é un sottospazio di $\mathbf{R}^n$;

(4) *in particolare sono sottospazi vettoriali (rispettivamente di $\mathbf{R}^k$ ed $\mathbf{R}^n$) gli insiemi*

$$Im(L) = L(\mathbf{R}^n) = \{\mathbf{y} \in \mathbf{R}^k | \exists x \in \mathbf{R}^n \ \textit{tale che}\ \mathbf{y} = L(\mathbf{x})\} \subseteq \mathbf{R}^k, \quad \textit{(detto } \textbf{immagine di } L\textit{)},$$

$$Ker(L) = L^{-1}(\{\mathbf{0}_k\}) = \{\mathbf{x} \in \mathbf{R}^n | L(\mathbf{x}) = \mathbf{0}_k\}, \quad \textit{(detto } \textbf{nucleo di } L\textit{)}.$$

(5) *se H é il sottospazio di $\mathbf{R}^n$ generato dai vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, allora $L(H)$ é il sottospazio di $\mathbf{R}^k$ generato dai vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)$;*
(6) *se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono vettori linearmente dipendenti di $\mathbf{R}^n$, allora $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)$ sono vettori linearmente dipendenti di $\mathbf{R}^k$;*
(7) *se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono vettori linearmente indipendenti di $\mathbf{R}^n$ ed L é iniettiva, allora i vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)$ sono linearmente indipendenti in $\mathbf{R}^k$;*
(8) *se L é iniettiva ed H é un sottospazio vettoriale di $\mathbf{R}^n$, allora H ed $L(H)$ hanno la stessa dimensione.*
(9) *L é iniettiva se e solo se $Ker(L) = \{\mathbf{0}_n\}$.*

Dim. di (1). Risulta infatti

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r) = L(\alpha_1\mathbf{v}_1) + L(\alpha_2\mathbf{v}_2) + \ldots + L(\alpha_r\mathbf{v}_r) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r).$$

Dim. di (2). Si deve dimostrare che se $\mathbf{y}, \mathbf{z} \in L(H)$ ed $\alpha \in \mathbf{R}$, allora $\mathbf{y} + \mathbf{z}$ ed $\alpha\mathbf{y}$ appartengono ad $L(H)$.
Infatti, per ipotesi esistono $\mathbf{u}, \mathbf{v} \in H$ tali che $L(\mathbf{u}) = \mathbf{y}$ ed $L(\mathbf{v}) = \mathbf{z}$; ne segue (essendo H un sottospazio) che $\mathbf{u} + \mathbf{v} \in H$ ed $\alpha\mathbf{u} \in H$, e quindi che $\mathbf{y} + \mathbf{z} = L(\mathbf{u}) + L(\mathbf{v}) = L(\mathbf{u} + \mathbf{v}) \in L(H)$ ed $\alpha\mathbf{y} = \alpha L(\mathbf{u}) = L(\alpha\mathbf{u}) \in L(H)$.

Dim. di (3). Si deve dimostrare che se $\mathbf{u}, \mathbf{v} \in L^{-1}(H')$, ed $\alpha \in \mathbf{R}$, allora si ha che $\mathbf{u}+\mathbf{v}$ ed $\alpha\mathbf{u}$ appartengono ad $L^{-1}(H')$, cioé che $L(\mathbf{u} + \mathbf{v})$ ed $L(\alpha\mathbf{u})$ appartengono ad H'.
Infatti per ipotesi si ha che $L(\mathbf{u}) \in H'$, $L(\mathbf{v}) \in H'$ e quindi (essendo H' un sottospazio) che $L(\mathbf{u}) + L(\mathbf{v})$ ed $\alpha L(\mathbf{u})$ appartengono ad H'. La tesi discende allora dal fatto che $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u})+L(\mathbf{v})$ ed $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u})$.

Dim. di (4). La tesi discende da (2) e (3), dal momento che $\mathbf{R}^n$ é un sottospazio di sé stesso e $\{\mathbf{0}_k\}$ é un sottospazio di $\mathbf{R}^k$.

Dim. di (5). Occorre dimostrare che per ogni $\mathbf{y} \in L(H)$ esistono $\alpha_1, \alpha_2, \ldots, \alpha_r$ tali che

$$\mathbf{y} = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r).$$

Infatti, fissato $\mathbf{y} \in L(H)$, esiste $\mathbf{x} \in H$ tale che $\mathbf{y} = L(\mathbf{x})$; d'altra parte, essendo $H = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$, esistono $\alpha_1, \alpha_2, \ldots, \alpha_r$ tali che $\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r$. Ne segue che

$$\mathbf{y} = L(\mathbf{x}) = L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r).$$

Dim. di (6). Infatti se i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono linearmente dipendenti, allora esistono r numeri reali non tutti nulli $\alpha_1, \alpha_2, \ldots, \alpha_r$ tali che $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}_n$; ne segue che risulta:

$$\alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r) = L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r) = L(\mathbf{0}_n) = \mathbf{0}_k,$$

e questo prova che i vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)$ sono linearmente dipendenti.

Dim. di (7). Infatti se i vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)$ fossero linearmente dipendenti, allora esisterebbero r numeri non tutti nulli $\alpha_1, \alpha_2, \ldots, \alpha_r$ tali che $\alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r) = \mathbf{0}_k$.
Ne seguirebbe che $L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_r L(\mathbf{v}_r) = \mathbf{0}_k = L(\mathbf{0}_n)$, e quindi, (per la iniettivitá di L), che $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}_n$ con $\alpha_1, \alpha_2, \ldots, \alpha_r$ non tutti nulli. I vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sarebbero quindi linearmente dipendenti, contraddicendo l'ipotesi.

Dim. di (8). Da (6) e (7) segue che se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ é una base di H, allora $\{L(\mathbf{v}_1), L(\mathbf{v}_2), \ldots, L(\mathbf{v}_r)\}$ é una base di $L(H)$, e questo prova la tesi.

Dim. di (9). Infatti se L é iniettiva, allora risulta $Ker(L) = \{\mathbf{0}_n\}$, dal momento che $\mathbf{u} \in Ker(L)$ se e solo se $L(\mathbf{u}) = \mathbf{0}_k = L(\mathbf{0}_n)$ e quindi se e solo se $\mathbf{u} = \mathbf{0}_n$.
Viceversa se $Ker(L) = \{\mathbf{0}_n\}$, allora L é iniettiva, poiché risulta:

$$\mathbf{u} \neq \mathbf{v} \Longrightarrow \mathbf{u} - \mathbf{v} \neq \mathbf{0}_n \Longrightarrow \mathbf{u} - \mathbf{v} \notin Ker(L) \Longrightarrow L(\mathbf{u} - \mathbf{v}) \neq \mathbf{0}_k \Longrightarrow L(\mathbf{u}) - L(\mathbf{v}) \neq \mathbf{0}_k \Longrightarrow L(\mathbf{u}) \neq L(\mathbf{v}).$$

■

3. - Matrice di una trasformazione lineare

Nella Osservazione 1.3 si é visto che ad ogni matrice A di tipo $k \times n$ puó essere associata una trasformazione lineare, precisamente la trasformazione lineare L_A che associa ad ogni vettore $\mathbf{x}$ di $\mathbf{R}^n$ il vettore $L_A(\mathbf{x}) = A\mathbf{x}$ di $\mathbf{R}^k$. La seguente Proposizione mostra viceversa che ad ogni trasformazione lineare puó essere associata una matrice.

Proposizione 3.1. -. *Per ogni trasformazione lineare $L : \mathbf{R}^n \mapsto \mathbf{R}^k$ esiste una ed una sola matrice A di tipo $k \times n$ tale che $L = L_A$, cioé tale che $L(\mathbf{x}) = A\mathbf{x}$ per ogni $\mathbf{x} \in \mathbf{R}^n$.*

Dim. L'unicitá deriva dal fatto che se A é una matrice tale che $L(\mathbf{x}) = A\mathbf{x}$ per ogni $\mathbf{x}$, allora detta $\{E^1, E^2, \dots, E^n\}$ la base canonica di $\mathbf{R}^n$, (cioé le colonne della matrice identica di ordine n), si ha:

$$L(E^1) = A \cdot E^1 = A^1, \qquad L(E^2) = A \cdot E^2 = A^2, \qquad \dots \quad L(E^n) = A \cdot E^n = A^n,$$

e quindi A é la matrice avente per colonne i vettori $L(\mathbf{E}^1), L(\mathbf{E}^2), \dots, L(\mathbf{E}^n)$.
D'altra parte, la matrice A avente per colonne i vettori $L(E^1), L(E^2), \dots, L(E^n)$, soddisfa la tesi, dal momento che per ogni $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$ si ha $\mathbf{x} = x_1E^1 + x_2E^2 + \dots + x_nE^n$ e quindi :

$$\begin{aligned} L(\mathbf{x}) &= L(x_1E^1 + x_2E^2 + \dots + x_nE^n) = x_1L(E^1) + x_2L(E^2) + \dots + x_nL(E^n) = \\ &= x_1A^1 + x_2A^2 + \dots + x_nA^n = A\mathbf{x} = L_A(\mathbf{x}). \end{aligned}$$

■

Osservazione 3.2 - La matrice A tale che $L(\mathbf{x}) = A\mathbf{x}$ per ogni $\mathbf{x} \in \mathbf{R}^n$ prende il nome di **matrice della trasformazione lineare L.** Il legame esistente tra il generico vettore $\mathbf{x}$ ed il corrispondente vettore $\mathbf{y} = L(\mathbf{x})$ é descritto dalla relazione matriciale $\mathbf{y} = A\mathbf{x}$, o equivalentemente dalle relazioni

$$(1) \qquad \begin{cases} y_1 &= a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ y_2 &= a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ & \dots\dots\dots\dots \\ y_k &= a_{k,1}x_1 + a_{k,2}x_2 + \dots + a_{k,n}x_n \end{cases}$$

che diconsi le **equazioni della trasformazione lineare** L .

■

Ad esempio le equazioni

$$\begin{cases} y_1 = 2x_1 - x_2 + 3x_3 \\ y_2 = 2x_2 - x_3, \end{cases} \qquad \begin{cases} y_1 = 3x_1 - 2x_2 \\ y_2 = x_1 + 2x_2 \\ y_3 = 4x_1 - 5x_2 \end{cases} \qquad \begin{cases} y_1 = 2x_1 - x_2 + 3x_3 \\ y_2 = 2x_2 - x_3 + x_4 \\ y_3 = x_1 - 2x_2 + 3x_4 \end{cases}$$

rappresentano trasformazioni lineari rispettivamente di $\mathbf{R}^3$ in $\mathbf{R}^2$, di $\mathbf{R}^2$ in $\mathbf{R}^3$ e di $\mathbf{R}^4$ in $\mathbf{R}^3$.
Le matrici di tali trasformazioni lineari sono rispettivamente le matrici

$$\begin{pmatrix} 2 & -1 & 3 \\ 0 & 2 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 3 & -2 \\ 1 & 2 \\ 4 & -5 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -1 & 3 & 0 \\ 0 & 2 & -1 & 1 \\ 1 & -2 & 0 & 3 \end{pmatrix}.$$

Non sono invece trasformazioni lineari le funzioni descritte dalle equazioni:

$$\begin{cases} y_1 = 2x_1 - x_2 + 3x_3 \\ y_2 = 2x_2 - x_3 + 1, \end{cases} \qquad \begin{cases} y_1 = x_1^2 - 2x_2 \\ y_2 = x_1 \cdot x_2 \\ y_3 = 3x_1 - 5x_2 \end{cases} \qquad \begin{cases} y_1 = 2x_1 - x_2 + 3x_3 \\ y_2 = 2x_2 - x_3 + x_4 \\ y_3 = \cos(x_1 - 2x_2 + 3x_4) \\ y_4 = -2x_1 + \sqrt{x_3^2 + 2x_4^2}. \end{cases}$$

■

Osservazione 3.3 - Dalla dimostrazione della Proposizione 3.1 si deduce che se L é una trasformazione lineare tra $\mathbf{R}^n$ ed R^k, allora i valori che L assume dipendono esclusivamente dai valori che L assume sui vettori $\{E^1, E^2, \ldots, E^n\}$ della base canonica di $\mathbf{R}^n$.

Infatti, la matrice di L é la matrice A che ha per colonne i vettori $L(E^1), L(E^2), \ldots, L(E^n)$ e per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha: $L(\mathbf{x}) = A\mathbf{x}$.

Ad esempio, la matrice della trasformazione lineare L di $\mathbf{R}^2$ in $\mathbf{R}^3$ tale che $L((1,0)^T) = (1,2,-3)^T$ ed $L((0,1)^T) = (-2,3,-1)^T$ é

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

Essa trasforma il generico vettore $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbf{R}^2$ nel vettore $\mathbf{y} = \begin{pmatrix} x_1 - 2x_2 \\ 2x_1 + 3x_2 \\ -3x_1 - x_2 \end{pmatrix}$.

■

Osservazione 3.4 - Possiamo facilmente trovare la matrice e le equazioni della rotazione di un angolo θ.

Infatti, evidentemente L trasforma i vettori della base canonica $E^1 = (1,0)^T$ ed $E^2 = (0,1)^T$ rispettivamente nei vettori

$$L(E^1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad \text{ed} \qquad L(E^2) = \begin{pmatrix} \cos(\pi/2 + \theta) \\ \sin(\pi/2 + \theta) \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

La matrice della rotazione intorno all'origine di un angolo θ é dunque la matrice:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Le equazioni della rotazione di un angolo θ sono

$$\begin{cases} y_1 = x_1\cos\theta - x_2\sin\theta \\ y_2 = x_1\sin\theta + x_2\cos\theta. \end{cases}$$

In effetti, per ogni $\mathbf{x} = (x_1, x_2)^T \neq \mathbf{0}$, posto $\rho = ||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2}$ e detta α l'ampiezza dell'angolo che la semiretta congiungente l'origine con il punto $\mathbf{x}$ forma con il semiasse positivo delle ascisse, si ha:

$$\mathbf{x} = \begin{pmatrix} \rho\cos\alpha \\ \rho\sin\alpha \end{pmatrix}, \qquad A\mathbf{x} = \begin{pmatrix} \rho\cos\alpha\cos\theta - \rho\sin\alpha\sin\theta \\ \rho\cos\alpha\sin\theta + \rho\sin\alpha\cos\theta \end{pmatrix} = \begin{pmatrix} \rho\cos(\theta + \alpha) \\ \rho\sin(\theta + \alpha) \end{pmatrix}.$$

Dunque, per ogni $\mathbf{x} \neq 0$ si ha che $A\mathbf{x}$ é esattamente il punto del piano ottenuto facendo ruotare il punto $\mathbf{x}$ di un angolo di ampiezza θ intorno all'origine, cioé $L(\mathbf{x})$.

■

4. - Nucleo ed immagine di una trasformazione lineare

Nello studio di una trasformazione lineare L un ruolo centrale viene giocato dal nucleo e dall'immagine di L. Sussiste infatti il seguente fondamentale

Teorema 4.1 - (Teorema della dimensione). *Se L é una trasformazione lineare tra $\mathbf{R}^n$ ed $\mathbf{R}^k$, allora risulta:*

$$\dim\big(Ker(L)\big) + \dim\big(Im(L)\big) = n = \dim\big(dom(L)\big).$$

Dim. Sia A la matrice della trasformazione lineare L, cioé la matrice di tipo $k \times n$ tale che $L(\mathbf{x}) = A\mathbf{x}$ per ogni $\mathbf{x} \in \mathbf{R}^n$. Si ha allora:

$$\mathbf{y} \in Im(L) \Longleftrightarrow \exists \mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n \text{ tale che } \mathbf{y} = L(\mathbf{x}) = A\mathbf{x} = x_1A^1 + x_2A^2 + \ldots + x_nA^n,$$

Pertanto $Im(L)$ coincide con lo spazio generato dalle colonne di A ed é dunque un sottospazio di $\mathbf{R}^k$ di dimensione $r = car(A)$. D'altra parte si ha

$$\mathbf{x} \in Ker(L) \Longleftrightarrow L(\mathbf{x}) = \mathbf{0}_k \Longleftrightarrow A\mathbf{x} = \mathbf{0}_k.$$

Ne segue che $Ker(L)$ coincide con l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$ ed é dunque un sottospazio di $\mathbf{R}^n$ di dimensione $n - r = n - car(A)$.

Si ha dunque $\quad \dim\big(Ker(L)\big) + dim\big(Im(L)\big) = (n - car(A)) + car(A) = n, \quad$ e la tesi é provata.

■

Osservazione 4.2 - Il precedente teorema in sostanza afferma che quanto piú é "grande"il nucleo di L tanto piú é "piccola"l'immagine $Im(L)$, cioé quanto piú numerosi sono i vettori linearmente indipendenti che vengono trasformati nell'origine $\mathbf{0}_k$ tanto meno numerosi sono i vettori linearmente indipendenti di $Im(L)$.

■

Osservazione 4.3 - Le argomentazioni sviluppate nella dimostrazione del precedente Teorema 4.1 ci consentono di dire:

- l'immagine $Im(L)$ di L coincide con lo spazio V generato dalle colonne della matrice A che la rappresenta; scopriremo invece nel prossimo capitolo il legame profondo che esiste tra il nucleo di L e il sottospazio U generato dalle righe di A e impareremo a descrivere in maniera efficace i sottospazi U e V;.
- la caratteristica di una matrice puó essere interpretata come la dimensione dello spazio immagine della trasformazione lineare di cui A é la matrice.
- l'insieme delle soluzioni di un sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$ puó essere interpretato come il nucleo della suddetta trasformazione lineare, cioé come l'insieme dei punti di $\mathbf{R}^n$ che L trasforma nell'origine $\mathbf{0}_k$ di $\mathbf{R}^k$.

■

Osservazione 4.4 - Sappiamo che L é iniettiva se e solo se $Ker(L) = \{\mathbf{0}_n\}$ cioé se e solo se il sistema $A\mathbf{x} = \mathbf{0}_k$ ha solo la soluzione nulla e sappiamo che questo accade se e solo se risulta $car(A) = n$.

Invece L é suriettiva se e solo se risulta $Im(L) = \mathbf{R}^k$ e quindi se e solo se risulta

$$car(A) = \dim(Im(L)) = \dim(\mathbf{R}^k) = k.$$

Ebbene, se risulta $car(A) = k = n$, allora L é iniettiva e suriettiva e quindi bigettiva; per ogni $\mathbf{y} \in \mathbf{R}^k$ esiste uno ed un solo $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{y} = L(\mathbf{x}) = A\mathbf{x}$, cioé il sistema $A\mathbf{x} = \mathbf{y}$ é determinato per ogni $\mathbf{y} \in \mathbf{R}^k$.

Se invece risulta $car(A) = n < k$, allora L é iniettiva, ma non suriettiva. Lo spazio immagine $Im(L)$ é un sottospazio proprio di dimensione $n < k$ di $\mathbf{R}^k$ ed L é una bigezione tra $\mathbf{R}^n$ ed $Im(L)$; il sistema $A\mathbf{x} = \mathbf{y}$ é determinato se $\mathbf{y} \in Im(L)$, incompatibile se $\mathbf{y} \notin Im(L)$.

Se risulta $car(A) = k < n$, allora L é suriettiva, ma non iniettiva. Il nucleo di L, cioé l'insieme delle soluzioni del sistema $A\mathbf{x} = \mathbf{0}_k$, é un sottospazio di $\mathbf{R}^n$ di dimensione $n - k$. La trasformazione L trasforma tale sottospazio nell'origine di $\mathbf{R}^k$ ed ogni sottospazio affine $x^* + Ker(L)$ parallelo a $Ker(L)$ nell'unico punto $L(x^*)$ e viceversa per ogni $\mathbf{y} \in \mathbf{R}^k$ l'insieme $\{\mathbf{x} \in \mathbf{R}^n | L(\mathbf{x}) = \mathbf{y}\}$, (cioé l'insieme delle soluzioni del sistema lineare non omogeneo $A\mathbf{x} = \mathbf{y}$), é un sottospazio affine parallelo a $Ker(L)$.

Infine se risulta $car(A) = r < \min(k, n)$, allora L non é iniettiva e nemmeno suriettiva. Il nucleo di L, cioé l'insieme delle soluzioni del sistema $A\mathbf{x} = \mathbf{0}_k$, é un sottospazio di $\mathbf{R}^n$ di dimensione $s = n - car(A)$. La trasformazione L trasforma tale sottospazio nell'origine di $\mathbf{R}^k$ ed ogni sottospazio affine $x^* + Ker(L)$ parallelo a $Ker(L)$ nell'unico punto $L(x^*)$ di $Im(L)$. Invece $Im(L)$ é un sottospazio di $\mathbf{R}^k$ di dimensione $r = car(A) = n - s$; per ogni $\mathbf{y} \in \mathbf{R}^k$ l'insieme $\{\mathbf{x} \in \mathbf{R}^n | L(\mathbf{x}) = \mathbf{y}\}$, (cioé l'insieme delle soluzioni del sistema lineare non omogeneo $A\mathbf{x} = \mathbf{y}$) é l'insieme vuoto se $\mathbf{y} \notin Im(L)$ ed é un sottospazio affine parallelo a $Ker(L)$ se $\mathbf{y} \in Im(L)$.

■

4 - Cambio di riferimento (o di coordinate)

Osservazione 4.1 - Sia A una matrice quadrata invertibile di ordine n e sia L_A la trasformazione lineare associata ad A; tale trasformazione associa ad ogni punto $\mathbf{x}$ di $\mathbf{R}^n$ uno ed un solo punto $\mathbf{y} = A\mathbf{x}$ di $\mathbf{R}^n$.

Peró tale trasformazione puó essere interpretata diversamente: infatti, sappiamo che le colonne di A sono linearmente indipendenti e quindi formano una base di $\mathbf{R}^n$; la soluzione $\mathbf{x}$ del sistema $A\mathbf{x} = \mathbf{y}$ puó essere allora interpretato come il vettore delle coordinate del vettore $\mathbf{y}$ rispetto alla base $\{A^1, A^2, \dots, A^n\}$ di $\mathbf{R}^n$.

In altri termini L_A trasforma il vettore $\mathbf{x}$ delle coordinate del generico punto di $\mathbf{R}^n$ rispetto alla base $\{A^1, A^2, \dots, A^n\}$ nel vettore $\mathbf{y}$ delle coordinate dello stesso punto rispetto alla base canonica $\{E^1, E^2, \dots, E^n\}$.

In generale, se $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ é una base di $\mathbf{R}^n$, allora detta A la matrice avente per colonne i vettori $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$, la trasformazione lineare associata ad A permette di passare dalle coordinate del generico vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ alle coordinate dello stesso vettore rispetto alla base canonica.

Viceversa l'inversa di A trasforma il vettore $\mathbf{y}$ delle coordinate del generico punto rispetto alla base canonica $\{E^1, E^2, \dots, E^n\}$ nel vettore $\mathbf{x} = A^{-1}\mathbf{y}$ delle coordinate dello stesso punto rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$.

■

Osservazione 4.2 - In generale, se $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$ sono due basi di $\mathbf{R}^n$, allora ad un vettore $\mathbf{v} \in \mathbf{R}^n$ possono essere associati due vettori $\mathbf{x}$ e $\mathbf{y}$ di $\mathbf{R}^n$ che rappresentano le coordinate di $\mathbf{v}$ rispetto alle due basi.

Ci chiediamo quale legame esista tra le coordinate di $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e le coordinate dello stesso vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$, cioé tra i vettori $\mathbf{x}$ ed $\mathbf{y}$.

A tal fine siano A e B le matrici quadrate invertibili aventi per colonne rispettivamente i vettori $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ e $\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}$. Per quanto detto sopra, possiamo affermare che se $\mathbf{x}$ é il vettore delle coordinate di $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$, allora $A\mathbf{x}$ rappresenta il vettore delle coordinate dello stesso punto $\mathbf{v}$ rispetto alla base canonica; ma allora $B^{-1}(A\mathbf{x})$ rappresenta il vettore delle coordinate dello stesso vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$, cioé $\mathbf{y}$.

Il legame tra le coordinate di un vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e le coordinate dello stesso vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$ é quindi fornito dalle relazioni

$$\mathbf{y} = B^{-1}(A\mathbf{x}) = (B^{-1}A)\mathbf{x}.$$

La matrice quadrata invertibile $C = B^{-1}A$ che descrive il passaggio dalle coordinate $\mathbf{x}$ di un vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ alle coordinate $\mathbf{y}$ dello stesso vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$ dicesi la matrice del **cambio di coordinate** o del **cambio di riferimento** .

Si noti che le colonne di $C = B^{-1}A$, cioé:

$$C^1 = B^{-1}A^1 = B^{-1}\mathbf{v}_1, \quad C^2 = B^{-1}A^2 = B^{-1}\mathbf{v}_2, \quad \dots, C^n = B^{-1}A^1 = B^{-1}\mathbf{v}_n$$

rappresentano i vettori delle coordinate dei vettori $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$.
In altri termini, la matrice del cambio di coordinate é la matrice che ha per colonne i vettori delle coordinate (rispetto alla nuova base) degli elementi della vecchia base.

■

5 - Matrice di una trasformazione lineare rispetto ad assegnate basi

Nella Proposizione 3.1 abbiamo visto come ad ogni trasformazione lineare $L : \mathbf{R}^n \longrightarrow \mathbf{R}^k$ possa essere associata una matrice A tale che $L(\mathbf{x}) = A\mathbf{x}$ per ogni $\mathbf{x}$, e precisamente la matrice che ha per colonne i vettori $L(E^1), L(E^2), \dots, L(E^n)$.

Ebbene in questo discorso i generici punti $\mathbf{x}$ ed $L(\mathbf{x})$ sono implicitamente riferiti alle basi canoniche di $\mathbf{R}^n$ ed $\mathbf{R}^k$, nel senso che le componenti di un vettore di $\mathbf{R}^n$ o di $\mathbf{R}^k$ sono esattamente le coordinate rispetto alle relative basi canoniche. Vediamo cosa accade se cambiamo riferimento sia sullo spazio di partenza $\mathbf{R}^n$ che su quello di arrivo $\mathbf{R}^k$.

Siano dunque $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$ basi rispetivamente di $\mathbf{R}^n$ ed $\mathbf{R}^k$, sia $\mathbf{v}$ il generico punto di $\mathbf{R}^n$ e siano $\mathbf{x}$ e $\mathbf{y}$ i vettori delle coordinate di $\mathbf{v}$ e di $L(\mathbf{v})$ rispetto alle basi $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$ rispettivamente; vogliamo stabilire quale legame esiste tra $\mathbf{x}$ e $\mathbf{y}$.
A tal fine siano

- A la matrice di tipo $k \times n$ avente per colonne i vettori $L(E^1), L(E^2), \dots, L(\mathbf{E}^n)$,
- B la matrice quadrata di ordine n avente per colonne i vettori $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$,
- C la matrice quadrata di ordine k avente per colonne i vettori $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$;

per quanto visto nella Osservazione 4.1, ne segue che $B\mathbf{x}$ e $C\mathbf{y}$ rappresentano rispettivamente i vettori delle coordinate di $\mathbf{v}$ ed $L(\mathbf{v})$ rispetto alle basi canoniche di $\mathbf{R}^n$ e di $\mathbf{R}^k$ e quindi che risulta $C\mathbf{y} = AB\mathbf{x}$, ovvero $\mathbf{y} = C^{-1}AB\mathbf{x}$.

Definizione 5.1. *La matrice* $A' = C^{-1}AB$ *che descrive il passaggio dal vettore* $\mathbf{x}$ *delle coordinate del generico vettore* $\mathbf{v} \in \mathbf{R}^n$ *rispetto alla base* $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ *al vettore* $\mathbf{y} = A'\mathbf{x}$ *delle coordinate del vettore* $L(\mathbf{v})$ *rispetto alla base* $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$ *dicesi la* **matrice della trasformazione lineare L rispetto alle basi $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$.**

■

Osservazione 5.2 - Scegliendo la base canonica sia in $\mathbf{R}^n$ che in $\mathbf{R}^k$, si ha che le matrici B e C descritte sopra coincidono con le matrici identiche di ordine n e k rispettivamente, e quindi la matrice di L rispetto alle basi canoniche di $\mathbf{R}^n$ ed $\mathbf{R}^k$ coincide esattamente con la matrice A avente per colonne i vettori $L(E^1),\ L(E^2), \dots, L(E^n)$.

■

Osservazione 5.3 - Si noti che le colonne della matrice A' sono i vettori delle coordinate dei vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \dots, L(\mathbf{v}_n)$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_k}\}$.
Ad esempio la prima colonna di A' é il vettore $(C^{-1}A)B^1 = C^{-1}(AB^1) = C^{-1}(A\mathbf{v}_1)$; ebbene $A\mathbf{v}_1$ rappresenta il vettore delle coordinate di $L(\mathbf{v}_1)$ rispetto alla base canonica di $\mathbf{R}^k$ e $C^{-1}(A\mathbf{v}_1)$ rappresenta il vettore delle coordinate dello stesso vettore rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}\}$.

■

Osservazione 5.4 - In definitiva una trasformazione lineare tra $\mathbf{R}^n$ ed $\mathbf{R}^k$ puó essere rappresentata da infinite matrici di tipo $k \times n$ a seconda delle basi scelte sugli spazi di partenza e di arrivo $\mathbf{R}^n$ ed $\mathbf{R}^k$.

Si puó dunque cercare di scegliere la base su $\mathbf{R}^n$ e/o su $\mathbf{R}^k$ in modo che L sia rappresentata da una matrice quanto piú semplice possibile. Questo fatto fornisce una spiegazione profonda ad esempio per l'algoritmo di Gauss Jordan o di Gauss Jordan completo.

In effetti, quando su una matrice non nulla A eseguiamo l'algoritmo di Gauss Jordan o l'algoritmo di Gauss Jordan completo, implicitamente stiamo utilizzando una base in $\mathbf{R}^n$ ed una base in $\mathbf{R}^k$, in modo che la trasformazione lineare L che é rappresentata da A rispetto alle basi canoniche sia rappresentata dalla matrice piú semplice A' rispetto a tali basi.

Questo fatto fornisce una spiegazione profonda del fatto che, ad esempio, le matrici A ed A' hanno la stessa caratteristica: infatti, $car(A)$ e $car(A')$ rappresentano la dimensione dello stesso spazio immagine di L.

■

6 - Matrice di una trasformazione lineare di $\mathbf{R}^n$ in sé rispetto ad una base di $\mathbf{R}^n$

Consideriamo ora il caso in cui l'insieme di arrivo della trasformazione lineare L coincide con l'insieme di partenza $\mathbf{R}^n$; in questo caso é possibile scegliere la stessa base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ sia nell'insieme di partenza che in quello di arrivo. Si puó allora dare la seguente:

Definizione 6.1. *Data la trasformazione lineare L di $\mathbf{R}^n$ in sé ed una base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ di $\mathbf{R}^n$, la matrice A' della trasformazione L rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ di $\mathbf{R}^n$ (considerato spazio di partenza) e la stessa base di $\mathbf{R}^n$ (considerato spazio di arrivo) dicesi* **matrice della trasformazione lineare** L **rispetto alla base** $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$.

Si ha allora evidentemente che:

(1) la matrice quadrata A di ordine n avente per colonne i vettori $L(E^1),\ L(E^2), \dots, L(E^n)$ non é altro che la matrice di L rispetto alla base canonica di $\mathbf{R}^n$; in generale la matrice A' di L rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ é la matrice quadrata di ordine n avente per colonne i vettori delle coordinate dei vettori $L(\mathbf{v}_1), L(\mathbf{v}_2), \dots, L(\mathbf{v}_n)$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$;
(2) la relazione $\mathbf{x} \mapsto \mathbf{y} = A'\mathbf{x}$ descrive il passaggio dalle coordinate del generico vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ al vettore $\mathbf{y}$ delle coordinate del vettore $L(\mathbf{v})$ rispetto alla stessa base;
(3) la matrice A' é legata alla matrice A dalla relazione $A' = B^{-1}AB$, dove B é la matrice avente per colonne i vettori $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$;
(4) viceversa se B é una matrice quadrata invertibile, allora $A' = B^{-1}AB$ é la matrice della trasformazione lineare L rispetto alla base di $\mathbf{R}^n$ formata dalle colonne di B.

■

In definitiva, la stessa trasformazione lineare L puó essere rappresentata da infinite matrici quadrate, e precisamente tutte e sole le matrici del tipo $B^{-1}AB$. Questo ci porta ad introdurre la seguente

Definizione 6.2 -. *Se A ed A' sono matrici quadrate di ordine n, si dice che A' é* **simile** *ad A se esiste una matrice quadrata invertibile B tale che $A' = B^{-1}AB$.*

Da quanto visto sopra, si deduce che le matrici associate ad una trasformazione lineare L di $\mathbf{R}^n$ in sé (rispetto ad una data base) sono tutte e sole le matrici simili alla matrice A che rappresenta L rispetto alla base canonica. Naturalmente converrá scegliere la base di $\mathbf{R}^n$ in modo che L sia rappresentata da una matrice quanto piú semplice possibile, ed é evidente che una matrice é tanto piú semplice quanto maggiore é il numero di elementi nulli della matrice.

L'ideale sarebbe riuscire addirittura a costruire una base di $\mathbf{R}^n$ rispetto a cui L é rappresentata da una matrice diagonale, perché in tal caso, detti $\lambda_1, \lambda_2, \dots, \lambda_n$ gli elementi diagonali di tale matrice, la trasformazione lineare L si limita a trasformare il generico vettore $\mathbf{v}$ di coordinate $(x_1, x_2, \dots, x_n)^T$ nel vettore $L(\mathbf{v})$ di coordinate $(\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n)^T$.

Ovviamente questo é possibile se e solo se la matrice A di L é simile ad una matrice diagonale, e quindi sorge il problema di stabilire sotto quali condizioni una matrice quadrata A é simile ad una matrice diagonale. A questo problema verrá fornita una soluzione nel Capitolo VII.

E S E R C I Z I

n. 1 - Dopo aver dato la definizione di trasformazione lineare, dire se sono trasformazioni lineari le funzioni L descritte dalle equazioni:

$$\begin{cases} y_1 &= x_1 - 3x_2 - 2x_3 \\ y_2 &= 2x_1 - 4x_2 + 3x_3 \end{cases} \qquad \begin{cases} y_1 &= x_1 - 3x_2 - 2x_3 \\ y_2 &= x_1 - 4x_2 + 3x_3 - 2, \end{cases} \qquad \begin{cases} y_1 &= x_1 - x_2 - 2x_3 \\ y_2 &= x_1 x_2 + x_3 \end{cases} .$$

$$\begin{cases} y_1 &= x_1 - x_2 - 2x_3 \\ y_2 &= 2x_1 - x_2 + 3x_3 \\ y_3 &= -3x_1 + 2x_2 - x_3 \end{cases} \qquad \begin{cases} y_1 &= x_1^2 + 3x_2 + x_3 \\ y_2 &= 3x_1 + 8x_2 - 2 \\ y_3 &= x_1 + x_2 + 4x_3. \end{cases}$$

Se L é una trasformazione lineare, verificare che il nucleo di L é un sottospazio vettoriale di $\mathbf{R}^n$.

n. 2 - Dati i vettori $\mathbf{u}_1 = (1, 1, -1, 0)$, $\mathbf{u}_2 = (-1, 3, 0, 2)$ e $\mathbf{u}_3 = (2, 1, 3, 3)$,

(1) dire se tali vettori sono linearmente dipendenti o indipendenti;
(2) trovare la dimensione ed una base dello spazio U generato da tali vettori;
(3) trovare le equazioni e la matrice A della trasformazione lineare L di $\mathbf{R}^3$ in $\mathbf{R}^4$ tale che
$L(E^1) = \mathbf{u}_1$, $L(E^2) = \mathbf{u}_2$, $L(E^3) = \mathbf{u}_3$;
(4) trovare la dimensione ed una base del nucleo e dell'immagine di tale trasformazione lineare e della trasformazione lineare associata alla trasposta di A; dire quindi se tali trasformazioni lineari sono iniettive, suriettive, bigettive;

Ripetere l'esercizio per i vettori

$$\mathbf{u}_1 = (1, 1, 0, -1)^T \ , \quad \mathbf{u}_2 = (-1, 0, 2, 2)^T \ , \quad \mathbf{u}_3 = (0, 1, 3, -2)^T \quad .$$

n. 3 - Date le matrici

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 & 1 & 0 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & -1 & 2 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & -1 & 0 & 1 \\ \alpha & 1 & 0 & 1 \\ 0 & -1 & 2 & 1 \\ 2 & \alpha & -4 & 0 \end{pmatrix},$$

a) trovare la loro caratteristica,
b) trovare la dimensione ed una base del nucleo e dell'immagine della trasformazione lineare associata a tali matrici;
c) dire se tale trasformazione lineare é iniettiva, suriettiva, bigettiva.

n. 4 - Dati i vettori $\mathbf{u}_1 = (h, 1, 2-h)^T$, $\mathbf{u}_2 = (1, 0, -1)^T$, $\mathbf{u}_3 = (2, 1, h)^T$,

a) scrivere le equazioni e la matrice A della trasformazione lineare L tale che
$L(E^1) = \mathbf{u}_1, \qquad L(E^2) = \mathbf{u}_2, \qquad L(E^3) = \mathbf{u}_3;$
b) dire per quale valore del parametro h le righe di A sono linearmente dipendenti o indipendenti,
b) trovare il rango di A al variare del parametro h,
d) trovare la dimensione ed una base del nucleo e dell'immagine della trasformazione lineare L.

n. 5 - Trovare la matrice de cambio di riferimento

dalla base $\{(-1,2,3)^T,\ (1,0,2)^T,\ (-2,1,4)^T\}$ alla base $\{(1,0,-1)^T,\ (0,2,3)^T,\ (1,2,-1)^T\}$.

n. 6 - Trovare la matrice della trasformazione lineare $L : \mathbf{R}^4 \mapsto \mathbf{R}^3$ di equazioni

$$\begin{cases} y_1 &= 2x_1 - x_2 + x_4 \\ y_2 &= x_1 + x_3 - 2x_4 \\ y_3 &= x_2 + x_3 + x_4 \end{cases}$$

rispetto alle basi

$$\{(-1,0,2,-1)^T,\ (0,1,0,2)^T,\ (1,-2,1,0)^T,\ (0,1,0,-1)^T\} \quad \text{e} \quad \{(1,0,-1)^T,\ (0,2,3)^T,\ (1,2,-1)^T\}.$$

n. 7 - Trovare la matrice della trasformazione lineare $L : \mathbf{R}^3 \mapsto \mathbf{R}^3$ di equazioni

$$\begin{cases} y_1 &= 2x_1 - x_2 + x_3 \\ y_2 &= x_1 + x_2 - 2x_3 \\ y_3 &= x_1 + x_2 + x_3 \end{cases}$$

rispetto alla base $\{(1,0,-1)^T,\ (0,2,3)^T,\ (1,2,-1)^T\}$.

Capitolo VI - ORTOGONALITA'

1. - Sottospazio ortogonale

In questo capitolo approfondiremo il concetto di ortogonalitá tra vettori. Ricordiamo che due vettori $\mathbf{u} = (u_1, u_2, \dots, u_n)$ e $\mathbf{v} = (v_1, v_2, \dots, v_n)$ di $\mathbf{R}^n$ si dicono **ortogonali,** se il loro prodotto scalare $(\mathbf{u}|\ \mathbf{v}) = u_1v_1 + u_2v_2 + \dots + u_nv_n$ é nullo.
Il concetto di ortogonalitá consente di vedere sotto un'altra luce l'insieme delle soluzioni di un'equazione o di un sistema lineare omogeneo.
Infatti, l'insieme delle soluzioni dell'equazione lineare omogenea $a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$ non é altro che l'insieme dei vettori di $\mathbf{R}^n$ che sono ortogonali al vettore $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$. In generale, se A é una matrice di tipo $k \times n$, allora l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ puó essere interpretato come l'insieme dei vettori di $\mathbf{R}^n$ che sono ortogonali alle righe di A.

La seguente definizione consente di precisare meglio i concetti.

Definizione 1.1. *Si dice che un vettore* $\mathbf{u}$ *é ortogonale ad un insieme* $X \subseteq \mathbf{R}^n$*, se é ortogonale a tutti gli elementi di* X*. Indicheremo con il simbolo* $X^\perp$ *l'insieme dei vettori ortogonali ad* X*, cioé l'insieme*

$$X^\perp = \{\mathbf{u} \in \mathbf{R}^n | (\mathbf{x}|\mathbf{u}) = 0 \textit{ per ogni } x \in X\}$$

∎

Si dimostra allora facilmente la seguente

Proposizione 1.2. *Per ogni* $X \subseteq \mathbf{R}^n$ *l'insieme* $X^\perp$ *dei vettori ortogonali ad* X *é un sottospazio vettoriale di* $\mathbf{R}^n$ *, che viene detto il* **sottospazio ortogonale** *ad* X*.*

Dim. Siano $\mathbf{u}, \mathbf{v} \in X^\perp$ ed $\alpha \in \mathbf{R}$; si deve provare che $\mathbf{u} + \mathbf{v}$ e $\alpha\mathbf{u}$ appartengono ad $X^\perp$.
Infatti per ogni $\mathbf{x} \in X$ si ha $(\mathbf{u}|\mathbf{x}) = 0$ e $(\mathbf{v}|\mathbf{x}) = 0$ e quindi $(\mathbf{u} + \mathbf{v}|\mathbf{x}) = (\mathbf{u}|\mathbf{x}) + (\mathbf{v}|\mathbf{x}) = 0 + 0 = 0$ e $(\alpha\mathbf{u}|\mathbf{x}) = \alpha(\mathbf{u}|\mathbf{x}) = 0$. Questo prova la tesi.

∎

Proposizione 1.3. *In particolare si ha* $\{\mathbf{0}_n\}^\perp = \mathbf{R}^n$, $(\mathbf{R}^n)^\perp = \{\mathbf{0}_n\}$.

Dim. E' evidente che $\{\mathbf{0}_n\}^\perp = \mathbf{R}^n$, poiché qualunque vettore di $\mathbf{R}^n$ é ortogonale al vettore nullo. D'altra parte, si ha $(\mathbf{R}^n)^\perp = \{\mathbf{0}\}$, poiché $\mathbf{0}$ é ortogonale ad ogni $\mathbf{x} \in \mathbf{R}^n$ e per ogni $\mathbf{x} \in (\mathbf{R}^n)^\perp$ si ha $(\mathbf{x}|\mathbf{x}) = 0$ e quindi $\mathbf{x} = \mathbf{0}$.

∎

La Proposizione che segue contiene le principali proprietá del sottospazio ortogonale ad un sottospazio.

Proposizione 1.4 -. *Se* U *é un sottospazio di* $\mathbf{R}^n$*, allora si ha quanto segue:*

a) *se* $X = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ *é un insieme di generatori di* U*, allora si ha* $U^\perp = X^\perp$*;*
b) *se* U *ha dimensione* k*, allora* $U^\perp$ *ha dimensione* $n - k$*,*
c) $(U^\perp)^\perp = U$*.*
d) $U \cap U^\perp = \{0_n\}$
e) $U + U^\perp = \mathbf{R}^n$*, nel senso che per ogni* $\mathbf{x} \in \mathbf{R}^n$ *esistono* $\mathbf{u} \in U$ *e* $\mathbf{v} \in U^\perp$ *tali che* $\mathbf{x} = \mathbf{u} + \mathbf{v}$*.*

Dim. di a). Se $\mathbf{x} \in U^\perp$, allora $\mathbf{x}$ é ortogonale a tutti gli elementi di U e quindi a tutti gli elementi di X. Questo prova che $x \in X^\perp$ e quindi che $U^\perp \subseteq X^\perp$.

Viceversa, sia $\mathbf{x} \in X^\perp$, cioé sia $(\mathbf{x}|\mathbf{u}_i) = 0$ per ogni $i = 1, 2, \dots, k$ e proviamo che $\mathbf{x} \in U^\perp$. Infatti, fissato $\mathbf{u} \in U$, esistono k numeri reali $\alpha_1, \alpha_2, \dots \alpha_k$ tali che $\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k$; ne segue che

$$(\mathbf{x}|\mathbf{u}) = (\mathbf{x} \mid \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k) = \alpha_1(\mathbf{x}|\mathbf{u}_1) + \alpha_2(\mathbf{x}|\mathbf{u}_2) + \dots + \alpha_k(\mathbf{x}|\mathbf{x}_k) = 0 + 0 + \dots + 0 = 0.$$

Dim. di b). Dalla Proposizione 1.3 si deduce che la tesi é ovvia se $k = 0$ oppure $k = n$. Sia dunque $0 < k < n$ e sia A una matrice di tipo $k \times n$ e caratteristica k avente per righe gli elementi di una base di U. Allora, per la precedente proprietá a), si ha che lo spazio ortogonale ad U coincide con lo spazio ortogonale all'insieme delle righe di A, cioé con l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$. Sappiamo che tale insieme é un sottospazio di dimensione $n - car(A) = n - k$, e quindi $U^\perp$ ha dimensione $n - k$.

Dim. di c). Fissato arbitrariamente $\mathbf{u} \in U$, per ogni $\mathbf{x} \in U^\perp$ si ha $(\mathbf{x}|\mathbf{u}) = 0$ e quindi $\mathbf{u} \in (U^\perp)^\perp$; questo prova che $U \subseteq (U^\perp)^\perp$. D'altra parte, dalla b) si deduce che

$$\dim((U^\perp)^\perp) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U).$$

Ne segue che una base di U é base anche di $(U^\perp)^\perp$ e quindi che $U = (U^\perp)^\perp$.

Dim. di d). Il vettore nullo appartiene a qualunque sottospazio di $\mathbf{R}^n$ e quindi $\mathbf{0} \in U \cap U^\perp$; viceversa se $\mathbf{x} \in U \cap U^\perp$, allora si ha $(\mathbf{x}|\mathbf{x}) = 0$, e quindi $\mathbf{x} = \mathbf{0}$.

Dim. di e). Per la Proposizione 1.3 la tesi é ovvia se la dimensione di U é 0 o n. Sia dunque $k = \dim(U)$ con $0 < k < n$ e sia $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ una base di U; sia poi $\{\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \dots, \mathbf{u}_n\}$ una base di $U^\perp$ e proviamo che i vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ sono linearmente indipendenti.

Infatti, se $\alpha_1, \alpha_2, \dots \alpha_n$ sono numeri reali tali che $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n = \mathbf{0}_n$. si ha che

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots \alpha_k\mathbf{u}_k = -\big(\alpha_{k+1}\mathbf{u}_{k+1} + \alpha_{k+2}\mathbf{u}_{k+2} + \dots + \alpha_n\mathbf{u}_n\big) \in U \cap U^\perp = \{\mathbf{0}_n\}$$

e quindi

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots \alpha_k\mathbf{u}_k = \mathbf{0}_n \qquad \text{e} \qquad \alpha_{k+1}\mathbf{u}_{k+1} + \alpha_{k+2}\mathbf{u}_{k+2} + \dots + \alpha_n\mathbf{u}_n = \mathbf{0}_n;$$

ne segue che $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ ed $\alpha_{k+1} = \alpha_{k+2}0 \dots = \alpha_n = 0$, dal momento che i vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ ed $\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \dots, \mathbf{u}_n$ sono linearmente indipendenti.

Pertanto i vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ sono n vettori linearmente indipendenti di $\mathbf{R}^n$ e quindi formano una base di tale spazio. Ne segue che per ogni $\mathbf{x} \in \mathbf{R}^n$ esistono n numeri reali $\alpha_1, \alpha_2, \dots \alpha_n$ tali che

$$\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n.$$

Posto $\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k \in U$ e $\mathbf{v} = \alpha_{k+1}\mathbf{u}_{k+1} + \alpha_{k+2}\mathbf{u}_{k+2} + \dots + \alpha_n\mathbf{u}_n \in U^\perp$, si ha che $\mathbf{x} = \mathbf{u} + \mathbf{v}$, come volevasi.

■

Osservazione 1.5 - Se U e V sono due sottospazi di $\mathbf{R}^n$ tali che $U \cap V = \{\mathbf{0}\}$ e $U + V = \mathbf{R}^n$, si dice che U e V sono **sottospazi complementari**.

Se U é un sottospazio vettoriale di dimensione k di $\mathbf{R}^n$, allora $U^\perp$ é uno degli infiniti sottospazi complementari ad U, e viene detto il **complemento ortogonale** ad U.

Per trovare gli altri sottospazi complementari ad U, basta fissare una base $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ di U e poi scegliere altri $n - k$ vettori $\mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n$ in modo che $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ formino una base di $\mathbf{R}^n$; allora lo spazio V generato dai vettori $\mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n$ é un sottospazio complementare ad U.

■

Osservazione 1.6 - Se U ha dimensione 1, cioé se U é una retta, allora il complemento ortogonale di U ha dimensione $n-1$ ed é dunque un iperpiano; tale iperpiano viene detto **l'iperpiano ortogonale** a tale retta. Viceversa, se U ha dimensione $n-1$, cioé é un iperpiano, il suo complemento ortogonale ha dimensione 1 e quindi é una retta, che viene detta la **retta ortogonale** all'iperpiano.

L' iperpiano ortogonale alla retta di direzione $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T \neq \mathbf{0}$ é dato dall'insieme dei punti $\mathbf{x}$ di $\mathbf{R}^n$ che soddisfano l'equazione $c_1x_1+c_2x_2+\ldots+c_nx_n = 0$. Viceversa l'equazione $c_1x_1+c_2x_2+\ldots+c_nx_n = 0$ con $c_1, c_2, \ldots, c_n$ non tutti nulli, rappresenta l'iperpiano ortogonale alla retta di direzione $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$.

■

Se ora $\mathbf{a}$ é un punto di $\mathbf{R}^n$ e $\mathbf{c}$ é un vettore non nullo di $\mathbf{R}^n$, allora dicesi **iperpiano affine passante per a e ortogonale al vettore c, (o alla retta generata da** c**)** l'iperpiano affine passante per $\mathbf{a}$ e parallelo all'iperpiano ortogonale al vettore $\mathbf{c}$. Esso é l'insieme dei punti $\mathbf{x}$ di $\mathbf{R}^n$ tali che $\mathbf{x} - \mathbf{a}$ sia ortogonale a $\mathbf{c}$, cioé l'insieme dei punti $\mathbf{x} \in \mathbf{R}^n$ soddisfacenti l'equazione

$$c_1(x_1 - a_1) + c_2(x_2 - a_2) + \ldots + c_n(x_n - a_n) = 0,$$

o equivalentemente l'equazione

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n = c_0, \quad \text{con} \quad c_0 = c_1a_1 + c_2a_2 + \ldots + c_na_n.$$

Viceversa l'insieme dei punti di $\mathbf{R}^n$ che soddisfano un'equazione del tipo $c_1x_1 + c_2x_2 + \ldots + c_nx_n = c_0$, con $c_1, c_2, \ldots, c_n$ non tutti nulli é un iperpiano affine ortogonale alla retta di direzione $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$.

■

2. - Proiezione ortogonale

Osservazione 2.1 - (Teorema di Pitagora) - Anche in $\mathbf{R}^n$ sussiste il "Teorema di Pitagora ", nel senso che se $\mathbf{x}$ e $\mathbf{y}$ sono vettori ortogonali di $\mathbf{R}^n$, allora risulta

$$||\mathbf{x} + \mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2.$$

Infatti risulta

$$||\mathbf{x} + \mathbf{y}||^2 = (\mathbf{x} + \mathbf{y}|\mathbf{x} + \mathbf{y}) = (\mathbf{x}|\mathbf{x} + \mathbf{y}) + (\mathbf{y}|\mathbf{x} + \mathbf{y}) = (\mathbf{x}|\mathbf{x}) + (\mathbf{x}|\mathbf{y}) + (\mathbf{y}|\mathbf{x}) + (\mathbf{y}|\mathbf{y}) = ||\mathbf{x}||^2 + 0 + 0 + ||\mathbf{y}||^2.$$

■

Osservazione 2.2 - Dalla e) della Prop. 1.4 si deduce che, se U é un sottospazio di $\mathbf{R}^n$, allora per ogni $\mathbf{x} \in U$ esiste un vettore $\mathbf{u} \in U$ tale che $\mathbf{x} - \mathbf{u} = \mathbf{v} \in U^\perp$.

Anzi tale vettore $\mathbf{u}$ é unico, perché se $\mathbf{u}', \mathbf{u}''$ sono elementi di U tali che $\mathbf{x}-\mathbf{u}' = \mathbf{v}' \in U^\perp$ ed $\mathbf{x}-\mathbf{u}'' = \mathbf{v}'' \in U^\perp$, allora si ha $\mathbf{u}' + \mathbf{v}' = \mathbf{x} = \mathbf{u}'' + \mathbf{v}''$; ne segue che $\mathbf{u}' - \mathbf{u}'' = \mathbf{v}'' - \mathbf{v}' \in U \cap U^\perp = \{\mathbf{0}\}$ e quindi che $\mathbf{u}' = \mathbf{u}''$.

Ebbene l'unico $\mathbf{u} \in U$ tale che $\mathbf{x} - \mathbf{u} \in U^\perp$ prende il nome di **proiezione ortogonale di** $\mathbf{x}$ su U e viene denotato con il simbolo $P_U(\mathbf{x})$.

■

Si ha allora che

Proposizione 2.3 -. *Se U é un sottospazio vettoriale di $\mathbf{R}^n$, si ha quanto segue.*

(1) *la funzione P_U che associa ad ogni vettore $\mathbf{x}$ la sua proiezione ortogonale su U é una trasformazione lineare di $\mathbf{R}^n$ in sé, detta la* **proiezione ortogonale su U***;*

(2) $\mathbf{x} - P_U(\mathbf{x}) \in U^\perp$, *per ogni* $\mathbf{x} \in \mathbf{R}^n$,

(3) $||\mathbf{x}||^2 = ||P_U(\mathbf{x})||^2 + ||\mathbf{x} - P_U(\mathbf{x})||^2$ *per ogni* $\mathbf{x} \in \mathbf{R}^n$*;*

(4) $P_U(\mathbf{x}) = \mathbf{x}$ *se e solo se* $\mathbf{x} \in U$*;*

(5) $P_U(\mathbf{x}) = \mathbf{0}$ *se e solo se* $\mathbf{x} \in U^\perp$ *;*

(6) $||\mathbf{x} - P_U(\mathbf{x})|| \leq ||\mathbf{x} - \mathbf{u}||$ *per ogni* $\mathbf{u} \in U$, *(cioé $P_U(\mathbf{x})$ é il punto di U piú vicino ad $\mathbf{x}$).*

Dim. di (1). Si vuole provare che per ogni $\mathbf{x},\ \mathbf{y} \in \mathbf{R}^n$ e per ogni $\alpha \in \mathbf{R}$ si ha:

$$P_U(\mathbf{x}+\mathbf{y}) = P_U(\mathbf{x}) + P_U(\mathbf{y}), \qquad P_U(\alpha\mathbf{x}) = \alpha P_U(\mathbf{x}).$$

Infatti, fissati $\mathbf{x},\ \mathbf{y} \in \mathbf{R}^n$, si ha che $P_U(\mathbf{x})$ e $P_U(\mathbf{y})$ appartengono ad U, mentre $\mathbf{x} - P_U(\mathbf{x})$ ed $\mathbf{y} - P_U(\mathbf{y})$ appartengono ad $U^\perp$; ne segue che

$$P_U(\mathbf{x}) + P_U(\mathbf{y}) \in U, \qquad (\mathbf{x}+\mathbf{y}) - (P_U(\mathbf{x}) + P_U(\mathbf{y})) = (\mathbf{x} - P_U(\mathbf{x})) + (\mathbf{y} - P_U(\mathbf{y})) \in U^\perp,$$

e questo prova che $P_U(\mathbf{x}) + P_U(\mathbf{y})$ é l'unico vettore $\mathbf{w}$ di U tale che $(\mathbf{x}+\mathbf{y}) - \mathbf{w} \in U^\perp$, cioé $P_U(\mathbf{x}+\mathbf{y})$.
D'altra parte per ogni $\alpha \in \mathbf{R}$ si ha che $\alpha P_U(\mathbf{x}) \in U$ ed $\alpha\mathbf{x} - \alpha P_U(\mathbf{x}) = \alpha \cdot (\mathbf{x} - P_U(\mathbf{x})) \in U^\perp$;, donde segue che $\alpha P_U(\mathbf{x}) = P_U(\alpha\mathbf{x})$.

Dim. di (2). La tesi é ovvia, perché (per definizione) $P_U(\mathbf{x})$ é l'unico elemento $\mathbf{u}$ di U tale che $\mathbf{x} - \mathbf{u} \in U^\perp$.

Dim. di (3). La tesi discende dal teorema di Pitagora, dal momento che $P_U(\mathbf{x})$ e $\mathbf{x} - P_U(\mathbf{x})$ sono ortogonali.

Dim. di (4). Se $\mathbf{x} \in U$, dall'essere $\mathbf{x} - \mathbf{x} = \mathbf{0} \in U^\perp$, si deduce che $P_U(\mathbf{x}) = \mathbf{x}$. Viceversa se $\mathbf{x} = P_U(\mathbf{x})$, allora é $\mathbf{x} \in U$, dal momento che $P_U(\mathbf{x}) \in U$.

Dim. di (5). Se $\mathbf{x} = \mathbf{x} - \mathbf{0} \in U^\perp$, dall'essere $\mathbf{0} \in U$ si deduce che $P_U(\mathbf{x}) = \mathbf{0}$. Viceversa, dalla (2) si deduce che se $P_U(\mathbf{x}) = \mathbf{0}$, allora é $\mathbf{x} \in U^\perp$.

Dim. di (6). Per ogni $\mathbf{u} \in U$ si ha che $P_U(\mathbf{x}) - \mathbf{u}$ appartiene ad U ed $\mathbf{x} - P_U(\mathbf{x})$ appartiene ad $U^\perp$; ne segue, per il Teorema di Pitagora, che

$$||\mathbf{x} - \mathbf{u}||^2 = ||(\mathbf{x} - P_U(\mathbf{x}) + (P_U(\mathbf{x}) - \mathbf{u}||^2 = ||\mathbf{x} - P_U(\mathbf{x})||^2 + ||P_U(\mathbf{x}) - \mathbf{u}||^2 \geq ||\mathbf{x} - P_U(\mathbf{x})||^2.$$

■

3 . - Complemento ortogonale e sistemi lineari -

Osservazione 3.1 - Abbiamo giá osservato che se A é una matrice di tipo $k \times n$, allora l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$ puó essere interpretato come il sottospazio ortogonale all'insieme delle righe di A e dunque come il complemento ortogonale allo spazio generato da tali righe.

Viceversa, se U é un sottospazio di $\mathbf{R}^n$, allora per trovare il complemento ortogonale ad U é sufficiente:
(1) trovare un insieme di generatori (possibilmente una base) di U,
(2) disporre tali vettori come righe di una matrice A,
(3) risolvere il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$.
Il complemento ortogonale $U^\perp$ di U é proprio l'insieme delle soluzioni di tale sistema.

■

Esempio 1 - Si voglia trovare il complemento ortogonale al sottospazio U di $\mathbf{R}^5$ generato dai vettori $\mathbf{u}_1 = (1, 1, -1, 2, -3)^T$, $\mathbf{u}_2 = (-2, 1, 2, 2, -3)^T$ ed $\mathbf{u}_3 = (3, 1, -3, 2, -3)^T$.
A tal fine consideriamo la matrice A avente per righe tali vettori e risolviamo il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$. Ebbene, con due successivi passi di pivot completi la matrice A viene trasformata nelle matrici

$$\begin{pmatrix} 1 & 1 & -1 & 2 & -3 \\ -2 & 1 & 2 & 2 & -3 \\ 3 & 1 & -3 & 2 & -3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & -1 & 2 & -3 \\ 0 & 3 & 0 & 6 & -9 \\ 0 & -2 & 0 & -4 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il sistema $A\mathbf{x} = \mathbf{0}$ é quindi equivalente al sistema

$$\begin{cases} x_1 = x_3, \\ x_2 = -2x_4 + 3x_5; \end{cases}$$

le sue soluzioni sono tutti e soli i vettori del tipo

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ -2\alpha_2 + 3\alpha_3 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad \text{con } \alpha_1, \alpha_2, \alpha_3 \in \mathbf{R} \text{ arbitrari.}$$

Pertanto $U^\perp$ é il sottospazio di dimensione 3 generato dai vettori linearmente indipendenti

$$\mathbf{w}_1 = (1,0,1,0,0)^T, \qquad \mathbf{w}_2 = (0,-2,0,1,0)^T, \qquad \mathbf{w}_3 = (0,3,0,0,1)^T.$$

■

Osservazione 3.2 - Sappiamo che l'insieme delle soluzioni di un sistema lineare omogeneo é un sottospazio vettoriale di $\mathbf{R}^n$; viceversa, dalla Proposizione 1.4 si deduce che qualunque sottospazio vettoriale di $\mathbf{R}^n$ puó essere pensato come l'insieme delle soluzioni di un sistema lineare omogeneo.
Infatti se U é un sottospazio di dimensione k di $\mathbf{R}^n$, allora risulta $U = (U^\perp)^\perp$. Ne segue che U é l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$, dove A é la matrice di tipo $(n-k)\times n$ e caratteristica $n-k$ avente per righe gli elementi di una base di $V = U^\perp$.

■

Osservazione 3.3 - Alla luce del concetto di ortogonalitá possiamo caratterizzare ulteriormente il sottospazio U di $\mathbf{R}^n$ generato dalle righe di una matrice A, il sottospazio V di $\mathbf{R}^k$ generato dalle colonne di A, nonché il nucleo e l'immagine della trasformazione lineare associata ad A.

A tal fine, dopo aver ricordato che V coincide con lo spazio immagine della trasformazione lineare L_A associata ad A, osserviamo quanto segue.

(1) Il nucleo $Ker(L_A)$ di L_A coincide con l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$, e quindi rappresenta il complemento ortogonale allo spazio U generato dalle righe di A. Si ha dunque $Ker(L_A) = U^\perp$ e quindi $U = \big(Ker(L_A)\big)^\perp$.

(2) Analogamente, lo spazio V generato dalle colonne di A, (cioé dalle righe di A^T), sará il complemento ortogonale al nucleo della trasformazione lineare di $\mathbf{R}^k$ in $\mathbf{R}^n$ associata alla trasposta di A.
Pertanto $Im(L_A) = V$ é il complemento ortogonale all'insieme delle soluzioni del sistema lineare omogeneo $A^T\mathbf{y} = \mathbf{0}_n$.

(3) L_A é iniettiva se e solo se $Ker(L_A) = \{\mathbf{0}_n\}$, e quindi se e solo se $U = \big(Ker(L_A)\big)^\perp = \{\mathbf{0}_n\}^\perp = \mathbf{R}^n$.
L_A é suriettiva se e solo se $V = Im(L_A) = \mathbf{R}^k$ e quindi se e solo se $Ker(L_{A^T}) = V^\perp = \{\mathbf{0}_k\}$.

In definitiva, per poter descrivere completamente la trasformazione lineare L_A basta risolvere i sistemi lineari omogenei $A\mathbf{x} = \mathbf{0}_k$ e $A^T\mathbf{y} = \mathbf{0}_n$:

(a) se il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}_k$ ha solo la soluzione nulla $\mathbf{0}_n$, allora risulta

$$Ker(L_A) = \{\mathbf{0}_n\}, \qquad U = \big(Ker(L_A)\big)^\perp = \{\mathbf{0}_n\}^\perp = \mathbf{R}^n, \qquad L_A \text{ é iniettiva;}$$

(b) se invece l'insieme delle soluzioni del sistema $A\mathbf{x} = \mathbf{0}_k$ é generato da s soluzioni non banali linearmente indipendenti, allora, detta N la matrice di tipi $s \times n$ avente per righe tali soluzioni linearmente indipendenti, si ha che U é il complemento ortogonale allo spazio generato da tali soluzioni, cioé l'insieme delle soluzioni del sistema lineare omogeneo di s equazioni in n incognite $N\mathbf{x} = \mathbf{0}_s$;

(c) se il sistema lineare omogeneo $A^T\mathbf{y} = \mathbf{0}_n$ ha solo la soluzione nulla $\mathbf{0}_k$, allora risulta

$$Ker(L_{A^T}) = \{\mathbf{0}_k\}, \qquad V = \big(Ker(L_{A^T})\big)^\perp = \{\mathbf{0}_k\}^\perp = \mathbf{R}^k, \qquad L_A \text{ é suriettiva;}$$

(d) se invece l'insieme delle soluzioni del sistema $A^T\mathbf{y} = \mathbf{0}_n$ é generato da s soluzioni non banali linearmente indipendenti, allora, detta P la matrice di tipo $s \times k$ avente per righe tali soluzioni linearmente indipendenti, si ha che $V = Im(L_A)$ é il complemento ortogonale allo spazio generato da tali soluzioni, cioé l'insieme delle soluzioni del sistema lineare omogeneo di s equazioni in k incognite $P\mathbf{y} = \mathbf{0}_s$.

■

Osservazione 3.4 - Con le notazioni della precedente Osservazione, se L_A non é iniettiva, allora L_A trasforma $Ker(L_A)$ nel vettore nullo $\mathbf{0}_k$ di $\mathbf{R}^k$ e trasforma ogni sottospazio affine $\mathbf{x} + Ker(L_A)$ parallelo a $Ker(L_A)$ in un unico punto di $V = Im(L_A)$ e viceversa ogni punto di V é l'immagine mediante L_A di un sottospazio affine parallelo a $Ker(L_A)$.

D'altra parte, si vede facilmente che la restrizione di L_A ad U é una bigezione tra U e V; in altri termini, per ogni $\mathbf{y} \in V$ esiste uno ed un solo $\mathbf{u} \in U$ tale che $\mathbf{y} = L_A(\mathbf{u})$.

Infatti, l'unicitá discende dal fatto che se $\mathbf{u}'$ e $\mathbf{u}''$ sono elementi di U tali che $L_A(\mathbf{u}') = L_A(\mathbf{u}'') = \mathbf{y}$, si ha che $L_A(\mathbf{u}' - \mathbf{u}'') = L_A(\mathbf{u}') - L_A(\mathbf{u}") = \mathbf{y} - \mathbf{y} = \mathbf{0}_k$, e quindi che $\mathbf{u}' - \mathbf{u}'' \in U \cap Ker(L_A) = U \cap U^\perp = \{\mathbf{0}_n\}$, donde segue che $\mathbf{u}' = \mathbf{u}''$.

L'esistenza deriva invece dal fatto che fissato $\mathbf{y} \in V = Im(L_A)$ esiste $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{y} = L_A(\mathbf{x})$. Ebbene, posto $\mathbf{u} = P_U(\mathbf{x})$ e $\mathbf{v} = \mathbf{x} - \mathbf{u}$, si ha che $\mathbf{u} \in U$, $\mathbf{v} \in U^\perp = Ker(L_A)$ ed $\mathbf{x} = \mathbf{u} + \mathbf{v}$. Se ne deduce che

$$\mathbf{y} = L_A(\mathbf{x}) = L_A(\mathbf{u} + \mathbf{v}) = L_A(\mathbf{u}) + L_A(\mathbf{v}) = L_A(\mathbf{u}) + \mathbf{0} = L_A(\mathbf{u}).$$

Pertanto l'insieme $X = \{\mathbf{x} \in \mathbf{R}^n \mid L_A(\mathbf{x}) = \mathbf{y}\}$ é un sottospazio affine parallelo a $Ker(L_A)$; tale sottospazio affine interseca U in un unico punto $\mathbf{u}$ di U. Tale punto é caratterizzato dal fatto che risulta $\mathbf{u} = P_U(\mathbf{x})$ per ogni $\mathbf{x} \in X$.

Si noti che, per la (3) della Prop. 2.3, si ha $||\mathbf{x}||^2 \geq ||P_U(\mathbf{x})||^2$ e quindi $||P_U(\mathbf{x})|| \leq ||\mathbf{x}||$. Ne segue che l'unico punto di intersezione di X con U é il vettore di minima norma di X.

In altri termini, per ogni $\mathbf{y} \in V$ il sistema $A\mathbf{x} = \mathbf{y}$ ha una ed una sola soluzione $\mathbf{u}$ in U; tale soluzione é (tra tutte le soluzioni del sistema) la soluzione di minima norma.

■

Esempio 2 - Consideriamo la trasformazione lineare L_A associata alla matrice

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 3 & 6+2\alpha & \alpha & \alpha^2 \\ 0 & 0 & -\alpha & 2\alpha \end{pmatrix};$$

al variare del parametro α, vogliamo trovare il nucleo di L_A, il sottospazio U generato dalle righe di A, il sottospazio generato dalle colonne di A, cioé lo spazio immagine di L_A, e dedurne se L_A é iniettiva, suriettiva, ecc.

A tal fine dobbiamo risolvere i sistemi lineari omogenei $A\mathbf{x} = \mathbf{0}_4$ e $A^T\mathbf{y} = \mathbf{0}_4$, e per fare questo sará sufficiente eseguire l'algoritmo di Gauss Jordan sia su A che su A^T.

Ebbene, con due successivi passi di pivot e sostituendo poi l'ultima riga con la somma delle ultime due, la matrice A subisce le seguenti trasformazioni:

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 0 & 2\alpha & \alpha & \alpha^2 \\ 0 & 0 & -\alpha & 2\alpha \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & \alpha & \alpha^2+\alpha \\ 0 & 0 & -\alpha & 2\alpha \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & \alpha & \alpha^2+\alpha \\ 0 & 0 & 0 & \alpha^2+3\alpha \end{pmatrix}.$$

Invece, con due successivi passi di pivot e sostituendo infine l'ultima riga con la stessa riga meno il multiplo secondo il fattore $k+1$ della terza riga, la matrice A^T diventa via via

$$\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & -2 & 2\alpha & 0 \\ 0 & 0 & \alpha & -\alpha \\ 0 & 1 & \alpha^2 & 2\alpha \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -\alpha & 0 \\ 0 & 0 & \alpha & -\alpha \\ 0 & 0 & \alpha^2+\alpha & 2\alpha \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -\alpha & 0 \\ 0 & 0 & \alpha & -\alpha \\ 0 & 0 & 0 & \alpha^2+3\alpha \end{pmatrix}.$$

Se α é diverso da 0 e da -3, le matrici A ed A^T sono diventate matrici triangolari superiori con elementi diagonali diversi da 0; pertanto le colonne di A ed A^T sono linearmente indipendenti.

I sistemi limeari omogenei $A\mathbf{x} = \mathbf{0}_4$ ed $A\mathbf{y} = \mathbf{0}_4$ hanno solo la soluzione nulla; i nuclei di L_A e di L_{A^T} coincidono con $\{\mathbf{0}_4\}$; la matrice A ha caratteristica $r = 4$, gli spazi U e V generati dalle righe e colonne di A coincidono con $\{\mathbf{0}_4\}^\perp = \mathbf{R}^4$.

Pertanto L_A é iniettiva e suriettiva e quindi bigettiva; il sistema non omogeneo $A\mathbf{x} = \mathbf{y}$ ha una ed una sola soluzione per ogni $\mathbf{y} \in \mathbf{R}^4$.

Se $\alpha = -3$, le matrici A e A^T sono diventate rispettivamente

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & -3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ne segue che A ed A^T hanno caratteristica $3 < 4 = n$; i sistemi lineari omogenei $A\mathbf{x} = \mathbf{0}$ ed $A^T\mathbf{y} = \mathbf{0}$ hanno rispettivamente le infinite soluzioni

$$\mathbf{x} = (-\beta,\ \beta/2,\ 2\beta,\ \beta)^T = \beta(-1,\ 1/2,\ 2,\ 1)^T \quad \text{ed} \quad \mathbf{y} = (-3\beta,\ -3\beta,\ \beta,\ \beta)^T = \beta(-3,\ -3,\ 1,\ 1)^T$$

con $\beta \in \mathbf{R}$ arbitrario.

Il nucleo di L_A é la retta generata dal vettore $(-1, 1/2, 2, 1)^T$; lo spazio U generato dalle righe di A é quindi l'iperpiano ortogonale a questa retta, cioé l'iperpiano di equazione $-x_1 + (1/2)x_2 + 2x_3 + x_4 = 0$ o equivalentemente $2x_1 - x_2 - 4x_3 - 2x_4 = 0$.

Analogamente il nucleo di L_{A^T} é la retta generata dal vettore $(-3, -3, 1, 1)^T$; lo spazio immagine di L_A é quindi l'iperpiano ortogonale a tale retta, cioé l'iperpiano di equazione $-3y_1 - 3y_2 + y_3 + y_4 = 0$ o equivalentemente $3y_1 + 3y_2 - y_3 - y_4 = 0$.

L_A é una bigezione tra l'iperpiano U e l'iperpiano V; per ogni $\mathbf{y} \in \mathbf{R}^4$ il sistema $A\mathbf{x} = \mathbf{y}$ é compatibile se e solo se $\mathbf{y} = (y_1, y_2, y_3, y_4)^T$ soddisfa l'equazione $3y_1 + 3y_2 - y_3 - y_4 = 0$. Se tale equazione é soddisfatta, il sistema ha una ed una sola soluzione $\mathbf{x}$ nell'iperpiano U di equazione $2x_1 - x_2 - 4x_3 - 2x_4 = 0$, tutte le altre sono del tipo $\mathbf{x} + \beta(-1, 1/2, 2, 1)^T$ con $\beta \in \mathbf{R}$. L'unica soluzione del sistema $A\mathbf{x} = \mathbf{y}$ che appartiene ad U é la soluzione di minima norma.

Se $\alpha = 0$, le matrici A e A^T sono diventate

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Questa volta le matrici A ed A^T hanno caratteristica $2 < 4 = n$. I sistemi lineari omogenei $A\mathbf{x} = \mathbf{0}$ e $A^T\mathbf{y} = \mathbf{0}$ hanno rispettivamente le infinite soluzioni

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\alpha_2 \\ \alpha_2/2 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 1/2 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} -3\alpha_1 \\ 0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

con α_1, α_2 arbitrari numeri reali.

Dunque il nucleo di L_A é il sottospazio di dimensione $2 = 4 - car(A)$, generato dai vettori $\mathbf{u}_1 = (0,0,1,0)^T$ e $\mathbf{u}_2 = (-1, 1/2, 0, 1)^T$ e il nucleo di L_{A^T} é il sottospazio di dimensione $2 = 4 - car(A)$, generato dai vettori linearmente indipendenti $\mathbf{v}_1 = (-3,0,1,0)^T$ e $\mathbf{v}_2 = ((0,0,0,1)^T$.

Lo spazio U generato dalle righe di A é quindi il sottospazio ortogonale ai vettori $\mathbf{u}_1, \mathbf{u}_2$; l'insieme $Im(L_A)$ é invece il sottospazio ortogonale ai vettori $\mathbf{v}_1, \mathbf{v}_2$. In altri termini, U e V sono rispettivamente gli insiemi delle soluzioni dei sistemi lineari omogenei

$$(1) \qquad \begin{cases} x_3 = 0 \\ 2x_1 - x_2 - 2x_4 = 0 \end{cases} \qquad\qquad (2) \qquad \begin{cases} 3y_1 - y_3 = 0 \\ y_4 = 0 \end{cases},$$

cioé l'intersezione degli iperpiani di equazione $x_3 = 0$, ed $x_1 - x_2 - 2x_4 = 0$ e l'intersezione degli iperpiani di equazione $3y_1 - y_3 = 0$ ed $y_4 = 0$.

L_A é una bigezione tra U e V; per ogni $\mathbf{y} \in \mathbf{R}^4$ il sistema $A\mathbf{x} = \mathbf{y}$ é compatibile se e solo se $\mathbf{y} = (y_1, y_2, y_3, y_4)^T$ soddisfa le equazioni (2); in tal caso il sistema ammette una ed una sola soluzione $\mathbf{x}$ soddisfacente il sistema (1); tutte le altre soluzioni sono del tipo $\mathbf{x} + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$ con α_1 e α_2 arbitrari numeri reali. L'unica soluzione $\mathbf{x}$ appartenente ad U é la soluzione di minima norma del sistema.

■

4. - Basi ortonormali

Se U é un sottospazio vettoriale di $\mathbf{R}^n$, un ruolo di particolare rilievo é giocato dalle basi di U che sono formate da vettori mutuamente ortogonali. A tal fine osserviamo in primo luogo che:

Proposizione 4.1. *Se $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ sono vettori non nulli a due a due ortogonali di $\mathbf{R}^n$, allora essi sono linearmente indipendenti e quindi formano una base dello spazio generato da tali vettori.*

Infatti se per assurdo tali vettori fossero linearmente dipendenti, allora uno di essi, ad esempio $\mathbf{u}_1$, apparterrebbe allo spazio generato dagli altri. Esisterebbero quindi dei numeri reali $\alpha_2, \alpha_3, \ldots, \alpha_k$ tali che $\mathbf{u}_1 = \alpha_2\mathbf{u}_2 + \ldots + \alpha_k\mathbf{u}_k$; ne seguirebbe che

$$(\mathbf{u}_1|\mathbf{u}_1) = (\mathbf{u}_1|\alpha_2\mathbf{u}_2 + \ldots + \alpha_k\mathbf{u}_k) = \alpha_2(\mathbf{u}_1|\mathbf{u}_2) + \alpha_3(\mathbf{u}_1|\mathbf{u}_3) + \ldots + \alpha_k(\mathbf{u}_1|\mathbf{u}_k) = 0 + 0 + \ldots + 0 = 0,$$

e dunque che $\mathbf{u}_1 = \mathbf{0}$, contro l'ipotesi che i vettori sono non nulli.

■

La precedente Proposizione giustifica la seguente

Definizione 4.2. *Se U é un sottospazio vettoriale di $\mathbf{R}^n$, dicesi* **base ortogonale** *di U un insieme finito di vettori che generano U e sono non nulli e a due a due ortogonali.*

Dicesi **base ortonormale** *di U un insieme finito di vettori che generano U, hanno norma unitaria e sono a due a due ortogonali.*

■

Ad esempio, le colonne della matrice identica di ordine n, cioé i vettori

$$E^1 = (1,0,0,\ldots,0)^T, \qquad E^2 = (0,1,0,\ldots,0)^T, \qquad \ldots\ldots, \qquad E^n = (0,0,\ldots,0,1)^T$$

formano una base ortonormale di $\mathbf{R}^n$.

■

La proposizione che segue illustra l'importanza e l'utilitá delle basi ortonormali.

Proposizione 4.3 -. *Se U é un sottospazio di $\mathbf{R}^n$ ed $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ é una base ortonormale di U, allora:*

a) *per ogni $\mathbf{x} \in \mathbf{R}^n$ la proiezione ortogonale di $\mathbf{x}$ su U é il vettore*

$$\mathbf{u} = (\mathbf{x}|\mathbf{u}_1) \cdot \mathbf{u}_1 + (\mathbf{x}|\mathbf{u}_2) \cdot \mathbf{u}_2 + \dots + (\mathbf{x}|\mathbf{u}_k) \cdot \mathbf{u}_k;$$

b) *per ogni $\mathbf{x} \in U$ si ha* $\quad \mathbf{x} = (\mathbf{x}|\mathbf{u}_1) \cdot \mathbf{u}_1 + (\mathbf{x}|\mathbf{u}_2) \cdot \mathbf{u}_2 + \dots + (\mathbf{x}|\mathbf{u}_k) \cdot \mathbf{u}_k$.
In altri termini, le coordinate di $\mathbf{x}$ rispetto alla base ortonormale $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ sono i numeri

$$(\mathbf{x}|\mathbf{u}_1), \qquad (\mathbf{x}|\mathbf{u}_2), \qquad \dots \qquad (\mathbf{x}|\mathbf{u}_k),$$

che sono detti i **coefficienti di Fourier di x rispetto alla base ortonormale** $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$.

Dim. Per provare la a), basta provare che $\mathbf{x} - \mathbf{u} \in U^\perp$, cioé che $\mathbf{x} - \mathbf{u}$ é ortogonale ad $\mathbf{u}_j$ per ogni j.
In effetti per ogni $i = 1, 2, \dots k$ si ha che $(\mathbf{u}_i|\mathbf{u}_j)$ é uguale a 0 se $i \neq j$ ed é uguale ad 1 se $i = j$ e quindi risulta:

$$(\mathbf{x} - \mathbf{u}|\mathbf{u}_j) = (\mathbf{x}|\mathbf{u}_j) - \sum_{i=1}^{k} (\mathbf{x}|\mathbf{u}_i) \cdot (\mathbf{u}_i|\mathbf{u}_j) = (\mathbf{x}|\mathbf{u}_j) - (\mathbf{x}|\mathbf{u}_j) = 0.$$

La b) discende dalla a) e dal fatto che $\mathbf{x} = P_U(\mathbf{x})$ se $\mathbf{x} \in U$.

■

La proposizione seguente dimostra che da una base di un sottospazio U di $\mathbf{R}^n$ si puó sempre ricavare una base ortonormale dello stesso sottospazio.

Proposizione 4.4 -. *Se U é un sottospazio di dimensione k di $\mathbf{R}^n$ e $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ é una base di U, allora esiste una base ortonormale $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ di U tale che:*

(1) $\qquad \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_j) \quad$ *per ogni $j = 1, 2, \dots, k$.*

Dim. Ovviamente, sará sufficiente costruire una base ortogonale $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ di U soddisfacente la (1), perché a quel punto sostituendo ciascun $\mathbf{u}_j$ con $\mathbf{u}_j^* = (1/||\mathbf{u}_j||) \cdot \mathbf{u}_j$ si otterrá una base ortonormale di U che soddisfa ancora la (1).
Ebbene, i vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ vengono costruiti attraverso il seguente procedimento ricorsivo, detto **Procedimento di ortogonalizzazione di Gram Schmidt**.
Come prima cosa si pone

$$\mathbf{u}_1 = \mathbf{a}_1;$$

chiaramente $\mathbf{u}_1$ é non nullo, (poiché $\mathbf{u}_1 = \mathbf{a}_1$ fa parte di una base), e risulta $\mathcal{L}(\mathbf{u}_1) = \mathcal{L}(\mathbf{a}_1)$.
Si costruisce quindi il vettore $\mathbf{u}_2$, ponendo

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{(\mathbf{a}_2|\mathbf{u}_1)}{(\mathbf{u}_1|\mathbf{u}_1)} \cdot \mathbf{u}_1.$$

Si ha chiaramente $(\mathbf{u}_2|\mathbf{u}_1) = 0$; inoltre é $\mathbf{u}_2 \neq \mathbf{0}$, altrimenti $\mathbf{a}_2$ apparterrebbe ad $\mathcal{L}(\mathbf{u}_1) = \mathcal{L}(\mathbf{a}_1)$, contro l'ipotesi che i vettori $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ formano una base e quindi sono linearmente indipendenti; infine risulta chiaramente $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2)$.
A questo punto si costruisce il vettore $\mathbf{u}_3$ ponendo

$$\mathbf{u}_3 = \mathbf{a}_3 - \frac{(\mathbf{a}_3|\mathbf{u}_1)}{(\mathbf{u}_1|\mathbf{u}_1)} \cdot \mathbf{u}_1 - \frac{(\mathbf{a}_3|\mathbf{u}_2)}{(\mathbf{u}_2|\mathbf{u}_2)} \cdot \mathbf{u}_2.$$

Si ha chiaramente $(\mathbf{u}_3|\mathbf{u}_1) = 0$ ed $(\mathbf{u}_3|\mathbf{u}_2) = 0$; inoltre $\mathbf{u}_3 \neq \mathbf{0}$, (altrimenti $\mathbf{a}_3$ apparterrebbe al sottospazio $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2)$, contro l'ipotesi che i vettori $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$ sono linearmente indipendenti); infine risulta chiaramente $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$.
In generale, dopo aver costruito j vettori non nulli e a due a due ortogonali $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j$ tali che

$\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_j)$, si costruisce il vettore $\mathbf{u}_{j+1}$ ponendo:

$$\mathbf{u}_{j+1} = \mathbf{a}_{j+1} - \frac{(\mathbf{a}_{j+1}|\mathbf{u}_1)}{(\mathbf{u}_1|\mathbf{u}_1)} \cdot \mathbf{u}_1 - \frac{(\mathbf{a}_{j+1}|\mathbf{u}_2)}{(\mathbf{u}_2|\mathbf{u}_2)} \cdot \mathbf{u}_2 - \dots\dots - \frac{(\mathbf{a}_{j+1}|\mathbf{u}_j)}{(\mathbf{u}_j|\mathbf{u}_j)} \cdot \mathbf{u}_j.$$

Tale vettore é chiaramente ortogonale ai vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j$ giá costruiti, é non nullo (altrimenti apparterrebbe ad $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_j)$) e risulta $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{j+1}) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{j+1})$.
Dopo k iterazioni, sono stati costruiti k vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ che formano la desiderata base ortogonale; infatti essi sono non nulli, a due a due ortogonali, (perché ciascuno di essi é ortogonale a quelli costruiti prima), e generano U, poiché soddisfano la (1) per $j = k$.

■

Esempio - Consideriamo il sottospazio U di $\mathbf{R}^4$ generato dai vettori

$$\mathbf{a}_1 = (2, 1, -2, 0)^T, \qquad \mathbf{a}_2 = (0, -3, 3, 4)^T, \qquad \mathbf{a}_3 = (5, 1, 1, 4)^T;$$

una base ortogonale di U associata a tale base si costruisce ponendo :

$$\mathbf{u}_1 = \mathbf{a}_1 = (2, 1, -2, 0)^T,$$

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{(\mathbf{a}_2|\mathbf{u}_1)}{(\mathbf{u}_1|\mathbf{u}_1)} \cdot \mathbf{u}_1 = \begin{pmatrix} 0 \\ -3 \\ 3 \\ 4 \end{pmatrix} - \frac{0-3-6+0}{4+1+4+0} \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 1 \\ 4 \end{pmatrix}$$

$$\mathbf{u}_3 = \mathbf{a}_3 - \frac{(\mathbf{a}_3|\mathbf{u}_1)}{(\mathbf{u}_1|\mathbf{u}_1)} \cdot \mathbf{u}_1 - \frac{(\mathbf{a}_3|\mathbf{u}_2)}{(\mathbf{u}_2|\mathbf{u}_2)} \cdot \mathbf{u}_2 =$$

$$= \begin{pmatrix} 5 \\ 1 \\ 1 \\ 4 \end{pmatrix} - \frac{10+1-2+0}{4+1+4+0} \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \\ 0 \end{pmatrix} - \frac{10-2+1+16}{4+4+1+16} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}.$$

I vettori $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ sono non nulli, a due a due ortogonali e generano U. Si ha inoltre

$$||\mathbf{u}_1|| = \sqrt{4+1+4+0} = 3, \qquad ||\mathbf{u}_2|| = \sqrt{4+4+1+16} = 5, \qquad ||\mathbf{u}_3|| = \sqrt{1+4+4+0} = 3;$$

una base ortonormale di U é quindi $\{\mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_3^*\}$, dove

$$\mathbf{u}_1^* = (1/3)\mathbf{u}_1 = \begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2^* = (1/5)\mathbf{u}_5 = \begin{pmatrix} 2/5 \\ -2/5 \\ 1/5 \\ 4/5 \end{pmatrix}, \qquad \mathbf{u}_3^* = (1/3)\mathbf{u}_3 = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \\ 0 \end{pmatrix}.$$

■

5 . - Matrici ortogonali

Definizione 5.1 -. *Dicesi* **matrice ortogonale** *una matrice quadrata A di ordine n, le cui colonne formano una base ortonormale di $\mathbf{R}^n$.*

■

Una fondamentale proprietá delle matrici ortogonali é contenuta nella seguente:

Proposizione 5.2 -. *Se A é una matrice quadrata di ordine n, allora le seguenti proprietá sono equivalenti:*

a) *A é una matrice ortogonale,*
b) $A^T \cdot A = I_n$,
c) *A é invertibile e la sua inversa coincide con la trasposta A^T di A,*
d) $A \cdot A^T = I_n$,
e) *A^T é ortogonale, e quindi le righe di A formano una base ortonormale di $\mathbf{R}^n$.*

Dim. L'equivalenza di a) e b) discende dal fatto che dire che A é ortogonale equivale a dire che il prodotto scalare di due colonne distinte é nullo, mentre il prodotto scalare di una colonna per sé stessa é uguale ad 1, cioé che risulta : $A^T \cdot A = I_n$.

L'equivalenza di e) e d) discende dall' equivalenza di a) e b); infatti A^T é ortogonale se e solo se risulta $(A^T)^T \cdot A^T = I_n$ e dunque se e solo se $A \cdot A^T = I_n$.

Infine da c) discendono chiaramente sia b) che d). Viceversa, da b) o da d) segue che $\det(A) \neq 0$ e quindi che A é invertibile; d'altra parte, per l'unicitá dell'inversa di una matrice quadrata invertibile, da b) o d) si deduce che l'inversa di A é A^T, come vuole la c).

■

Ad esempio si vede immediatamente che sia le righe che le colonne della matrice

$$A = \begin{pmatrix} 2/3 & 1/3 & -2/3 & 0 \\ 2/5 & -2/5 & 1/5 & 4/5 \\ 1/3 & 2/3 & 2/3 & 0 \\ 8/15 & -8/15 & 4/15 & -9/15 \end{pmatrix}$$

formano una base ortonormale di $\mathbf{R}^4$; pertanto A e A^T sono matrici ortogonali ed A^T é l'inversa di A.

■

Un esempio di matrice ortogonale é costituito dalla matrice

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

che descrive la rotazione del piano intorno all'origine di un angolo di ampiezza θ.

E' evidente che una rotazione del piano conserva l'angolo tra due vettori, la loro norma, la loro distanza, il loro prodotto scalare, la loro eventuale ortogonalitá. In realtá questa proprietá sussiste per tutte le matrici ortogonali; sussiste infatti la seguente

Proposizione 5.3. *Se A é una matrice quadrata ortogonale di ordine n, allora risulta:*

(1) $(A\mathbf{x}|A\mathbf{y}) = (\mathbf{x}|\mathbf{y})$ *per ogni* $\mathbf{x},\ \mathbf{y} \in \mathbf{R}^n$.
(2) *se* $\mathbf{x}$ *é ortogonale ad* $\mathbf{y}$, *allora* $A\mathbf{x}$ *é ortogonale ad* $A\mathbf{y}$,
(3) $||A\mathbf{x}|| = ||\mathbf{x}||$ *per ogni* $\mathbf{x} \in \mathbf{R}^n$,
(4) $||A\mathbf{x} - A\mathbf{y}|| = ||\mathbf{x} - \mathbf{y}||$ *per ogni* $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$

e quindi la trasformazione lineare L_A associata ad A conserva la norma, la distanza, l'ortogonalitá e l'angolo tra due vettori.

Dim. Fissati $\mathbf{x}$ ed $\mathbf{y}$, si ha che:

$$(A\mathbf{x}|A\mathbf{y}) = (A\mathbf{x})^T \cdot (A\mathbf{y}) = (\mathbf{x}^T A^T) \cdot (A\mathbf{y}) = \mathbf{x}^T \cdot (A^T \cdot A) \cdot \mathbf{y} = \mathbf{x}^T \cdot I_n \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y} = (\mathbf{x}|\mathbf{y}).$$

La (2) e la (3) discendono chiaramente dalla (1), dal momento che

$$(\mathbf{x}|\mathbf{y}) = 0 \Longrightarrow (A\mathbf{x}|A\mathbf{y}) = (\mathbf{x}|\mathbf{y}) = 0 \qquad \text{e} \qquad ||A\mathbf{x}|| = \sqrt{(A\mathbf{x}|A\mathbf{x})} = \sqrt{(\mathbf{x}|\mathbf{x})} = ||\mathbf{x}||.$$

Infine la (4) discende dalla (3) essendo $||A\mathbf{x} - A\mathbf{y}|| = ||A(\mathbf{x} - \mathbf{y})|| = ||\mathbf{x} - \mathbf{y}||$.

∎

E S E R C I Z I

n. 1 - Dati i vettori $\mathbf{u}_1 = (1, 1, -1, 0)$, $\mathbf{u}_2 = (-1, 3, 0, 2)$ e $\mathbf{u}_3 = (2, 1, 3, 3)$,

(1) trovare lo spazio U generato da tali vettori e il suo complemento ortogonale $U^\perp$, indicando la dimensione ed una base di tali sottospazi;
(2) da tali basi di U ed $U^\perp$ dedurne una base ortonormale con il procedimento di ortogonalizzazione di Gram Schmidt;
(3) trovare il nucleo e l'immagine della trasformazione lineare associata alla matrice A avente per righe tali vettori, e il nucleo e l'immagine della trasformazione lineare associata alla trasposta di A.
(4) dire se il vettore $\mathbf{y} = (1, -2, 3, 5)^T$ appartiene allo spazio immagine della trasformazione lineare associata ad A^T; in caso affermativo, risolvere il sistema lineare $A^T\mathbf{x} = \mathbf{y}$.

Ripetere l'esercizio per i vettori

$\mathbf{u}_1 = (1, 1, 0, -1)^T$, $\mathbf{u}_2 = (-1, 0, 2, 2)^T$, $\mathbf{u}_3 = (0, 1, 3, -2)^T$ ed $\mathbf{y} = (2, 1, -3, 0)^T$.

n. 2 - Date le matrici

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 & 1 & 0 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & -1 & 2 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & -1 & 0 & 1 \\ \alpha & 1 & 0 & 1 \\ 0 & -1 & 2 & 1 \\ 2 & \alpha & -4 & 0 \end{pmatrix},$$

a) trovare la loro caratteristica,
b) trovare la dimensione del nucleo e dell'immagine della trasformazione lineare associata a tali matrici,
c) dire se tale trasformazione lineare é iniettiva, suriettiva, bigettiva,
d) trovare il nucleo e l'immagine delle suddette trasformazioni lineari e il complemento ortogonale a tali sottospazi, indicandone esplicitamente una base.

n. 3 - Date le matrici

$$A = \begin{pmatrix} h & 1 & 2-h \\ 1 & 0 & -1 \\ 2 & 1 & h \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & h \\ 2 & 5 & 1 \\ h-1 & -1 & 1 \end{pmatrix},$$

a) trovare il rango di A al variare del parametro h,
b) scrivere le equazioni della trasformazione lineare L_A associata alla matrice A;
c) trovare il nucleo e l'immagine di tale trasformazione lineare, indicando di tali sottospazi la dimensione ed una base;
d) trovare il complemento ortogonale al nucleo ed all'immagine di L_A, indicando anche di tali sottospazi la dimensione ed una base.
e) dire per quali valori del parametro h si ha che il vettore $\mathbf{y} = (1, -1, 0)^T$ appartiene allo spazio immagine di L_A; per tali valori risolvere il sistema lineare nonomogeneo $A\mathbf{x} = \mathbf{y}$.

Ripetere l'esercizio per la matrice B

n. 4 - Con il procedimento di ortogonalizzazione di Gram Schmidt, costruire una base ortonormale dello spazio $\mathbf{R}^4$ a partire dai vettori

$$\mathbf{u}_1 = (2, 0, 1, -2)^T, \qquad \mathbf{u}_2 = (-4, 4, -1, 0)^T, \qquad \mathbf{u}_3 = (0, -2, 1, -4)^T, \qquad \mathbf{u}_4 = (1, 2, -4 - 1)^T.$$

Ripetere l'esercizio per i vettori

$$\mathbf{u}_1 = (1, 0, 2, -2)^T, \qquad \mathbf{u}_2 = (4, 2, 3, -4)^T, \qquad \mathbf{u}_3 = (-2, 4, 1, 3)^T, \qquad \mathbf{u}_4 = (-1, 3, 3, 3)^T.$$

Capitolo VII - AUTOVALORI - AUTOVETTORI

1. - Definizioni, esempi e prime proprietá.

Sia A una matrice quadrata di ordine n e, per ogni $\lambda \in \mathbf{R}$ consideriamo la matrice $A - \lambda I$. Allora possono presentarsi solo le seguenti due alternative:

a) il sistema lineare omogeneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$ ha solo la soluzione banale, e quindi $A - \lambda I$ é invertibile,
b) il sistema lineare omogeneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$ ha anche soluzioni non banali.

Nel caso a), si dice che λ é un **valore regolare** di A (o della trasformazione lineare L_A associata ad A); nel caso b) si dice che λ é un **valore caratteristico** o **valore proprio** o **valore latente** o **valore spettrale** o **autovalore** di A o di L_A.

L'insieme dei valori regolari di A dicesi l'insieme **risolvente** di A e si denota con il simbolo $\rho(A)$; l'insieme degli autovalori di A dicesi lo **spettro** di A e si denota con il simbolo $\sigma(A)$.

Se λ é un autovalore di A, allora:

(1) una soluzione non banale del sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$ dicesi **autovettore** di A corrispondente all'autovalore λ;
(2) l'insieme $U(\lambda)$ delle soluzioni del sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$, (cioé il nucleo della trasformazione lineare associata alla matrice $A - \lambda I$), é un sottospazio vettoriale di $\mathbf{R}^n$, detto **l'autospazio** di A corrispondente all'autovalore λ;
(3) la dimensione $n - car(A - \lambda I)$ di $U(\lambda)$, (cioé il massimo numero di autovettori linearmente indipendenti corrispondenti all'autovalore λ), dicesi la **molteplicitá geometrica** di λ .

■

In altri termini, un vettore non nullo $\mathbf{u} \in \mathbf{R}^n$ é un autovettore di A se esiste un numero reale λ, (l'autovalore corrispondente ad $\mathbf{u}$), tale che $(A - \lambda I)\mathbf{u} = \mathbf{0}$, cioé tale che $L_A(\mathbf{u}) = A\mathbf{u} = \lambda\mathbf{u}$.

Pertanto la trasformazione lineare associata ad A trasforma l'autovettore $\mathbf{u}$ in $\lambda\mathbf{u}$ e tutti i punti $\alpha\mathbf{u}$ della retta generata da $\mathbf{u}$ in punti $(\lambda\alpha)\mathbf{u}$ della stessa retta; gli autovettori rappresentano quindi le **direzioni invarianti** di L_A, cioé le direzioni delle rette che non vengono modificate dalla trasformazione L_A.

Piú in generale, si ha che se λ é un autovalore di A, allora ogni sottospazio vettoriale U di $U(\lambda)$ é un sottospazio invariante per L_A, nel senso che $\mathbf{u} \in U \implies L_A(\mathbf{u}) = A\mathbf{u} = \lambda\mathbf{u} \in U$.

■

A questo punto sorgono spontanee le domande:

(1) *qualunque matrice quadrata A é dotata di autovalori?*
(2) *quanti sono gli autovalori di A?*
(3) *come si fa a trovarli?*

Alla prima di tali domande é facile dare una risposta negativa: **esistono delle matrici senza autovalori.**

Basti pensare (per un dato $\theta \in]0, 2\pi[$) alla matrice $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ che descrive la rotazione intorno all'origine di un angolo di ampiezza θ. E' evidente che nessuna retta passante per l'origine rimane invariante in seguito a tale trasformazione; pertanto A non ha autovettori, e quindi non ha autovalori.

Alle altre domande si puó dare facilmente risposta in base alla definizione di autovalore; infatti sappiamo che un sistema lineare omogeneo di n equazioni in n incognite ha soluzioni non banali se e solo se il determinante della matrice dei coefficienti del sistema é uguale a 0. Se ne deduce quindi che

Proposizione 1.1 -. *Un numero reale λ é un autovalore di A se e solo se risulta*

$$(1) \qquad \det(A - \lambda I) = 0 \quad \textit{cioé} \quad \det\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix} = 0.$$

■

Gli autovalori di A si trovano dunque semplicemente risolvendo l'equazione (1), che viene pertanto detta **l'equazione caratteristica** di A. Analogamente, il polinomio $p_A(\lambda) = \det(A - \lambda I)$, (le cui radici sono gli autovalori di A), viene detto il **polinomio caratteristico** di A.

■

A conferma di quanto visto prima, una matrice quadrata di ordine n puó anche non avere alcun autovalore, perché l'equazione caratteristica puó non avere alcuna radice reale. D'altra parte, é evidente che

Proposizione 1.2 -. *Una matrice quadrata di ordine n ha al massimo n autovalori.*

Dim. Infatti, il polinomio caratteristico $p_A(\lambda)$ di una matrice quadrata di ordine n é un polinomio di grado n e quindi ha al massimo n radici reali.

■

Esempio 1 - . Cerchiamo gli autovalori e gli autospazi della matrice

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

Si tratta di risolvere l'equazione caratteristica

$$\det(A - \lambda I) = \det\begin{pmatrix} 2-\lambda & -2 & 3 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{pmatrix} = 0.$$

Il determinante puó essere calcolato con la regola di Sarrus o con la regola di Laplace, oppure puó essere calcolato con l'algorimo di Gauss Jordan, scambiando la prima con la terza riga ed effettuando poi due passi di pivot parziali:

$$\det(A - \lambda I) = -\det\begin{pmatrix} 1 & 3 & -1-\lambda \\ 1 & 1-\lambda & 1 \\ 2-\lambda & -2 & 3 \end{pmatrix} = -\det\begin{pmatrix} 1 & 3 & -1-\lambda \\ 0 & -2-\lambda & 2+\lambda \\ 0 & -8+3\lambda & 5+\lambda-\lambda^2 \end{pmatrix} =$$

$$= (\lambda+2)\cdot\det\begin{pmatrix} 1 & 3 & -1-\lambda \\ 0 & 1 & -1 \\ 0 & -8+3\lambda & 5+\lambda-\lambda^2 \end{pmatrix} = (\lambda+2)\cdot\det\begin{pmatrix} 1 & 3 & -1-\lambda \\ 0 & 1 & -1 \\ 0 & 0 & -3+4\lambda-\lambda^2 \end{pmatrix} =$$

$$= -(\lambda+2)\cdot(\lambda^2 - 4\lambda + 3).$$

Il polinomio caratteristico di A é quindi

$$p_A(\lambda) = -(\lambda+2)(\lambda^2 - 4\lambda + 3) = -\lambda^3 + 2\lambda^2 + 5\lambda - 6;$$

una radice di tale polinomio é $\lambda_1 = -2$; le altre due si trovano risolvendo l'equazione $\lambda^2 - 4\lambda + 3 = 0$ e sono quindi $\lambda_2 = 1, \lambda_3 = 3$. Gli autovalori di A sono quindi $\lambda_1 = -2,\ \lambda_2 = 1,\ \lambda_3 = 3$.

Per trovare i corrispondenti autovettori, dobbiamo risolvere i sistemi lineari omogenei $(A - \lambda I)\mathbf{x} = \mathbf{0}$ con $\lambda = -2, \lambda = 1$ e $\lambda = 3$.

Per $\lambda = -2$ si ha il sistema

x_1	x_2	x_3	
4	-2	3	0
1	3	1	0
1	3	1	0

x_1	x_2	x_3	
1	3	1	0
0	-14	-1	0
0	0	0	| 0

x_1	x_3	x_2	
1	0	-11	0
0	1	14	0
0	0	0	0

L'autospazio associato all'autovalore λ_1 é quindi la retta generata dall' autovettore $\mathbf{u}_1 = (11, 1, -14)^T$.

Per $\lambda = 1$ si ha il sistema

x_1	x_2	x_3	
1	-2	3	0
1	0	1	0
1	3	-2	0

x_1	x_2	x_3	
1	-2	3	0
0	2	-2	0
0	5	-5	0

x_1	x_2	x_3	
1	0	1	0
0	1	-1	0
0	0	0	0

L'autospazio associato all'autovalore $\lambda_2 = 1$ é quindi la retta generata dall' autovettore $\mathbf{u}_2 = (-1, 1, 1)^T$.

Infine, per $\lambda_3 = 3$ si ha il sistema

x_1	x_2	x_3	
-1	-2	3	0
1	-2	1	0
1	3	-4	0

x_1	x_2	x_3	
1	2	-3	0
0	-4	4	0
0	1	-1	0

x_1	x_2	x_3	
1	0	-1	0
0	1	-1	0
0	0	0	0

L'autospazio associato all'autovalore $\lambda_3 = 3$ é quindi la retta generata dall' autovettore $\mathbf{u}_3 = (1, 1, 1)^T$.

■

Osservazione 1.3 - . Se A é una matrice triangolare superiore o inferiore, allora gli autovalori di A sono proprio i suoi elementi diagonali.

Infatti $A - \lambda I$ é, al pari di A, una matrice triangolare superiore o inferiore e quindi il suo determinante sará dato dal prodotto degli elementi della diagonale. Pertanto il polinomio caratteristico di A sará

$$p_A(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda),$$

e quindi gli autovalori di A sono $a_{11}, a_{22}, \dots, a_{nn}$.

■

Osservazione 1.4 - Ovviamente dire che 0 é un autovalore di A equivale a dire che A non é invertibile.

■

Osservazione 1.5 - Se λ^* é un autovalore di A ed é una radice di molteplicitá k del polinomio caratteristico $p_A(\lambda)$ di A, (cioé p_A é divisibile per $(\lambda - \lambda^*)^k$ e non é divisibile per $(\lambda - \lambda^*)^{k+1}$), allora si dice che l'autovalore λ^* di A ha **molteplicitá algebrica** k.

Si dimostra che, per un qualunque autovalore di A, la molteplicitá geometrica é minore o uguale della molteplicitá algebrica.

Si dice che λ^* é un **autovalore regolare** di A, se la sua molteplicitá geometrica coincide con la molteplicitá algebrica; si dice invece che λ^* é un **autovalore singolare** di A, se la sua molteplicitá geometrica é minore della molteplicitá algebrica.

■

Esempio 2 - . Cerchiamo gli autovalori e gli autospazi della matrice

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

Si tratta di risolvere l'equazione caratteristica

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = 0.$$

Il determinante puó essere calcolato utilizzando la formula di Laplace rispetto all'ultima riga; si ottiene allora che il polinomio caratteristico di A é

$$p_A(\lambda) = (2-\lambda)[(1-\lambda)^2 - 1] = -\lambda \cdot (\lambda - 2)^2,$$

donde si deduce che gli autovalori di A sono $\lambda_1 = 0$ e $\lambda_2 = 2$. Di essi, λ_1 é un autovalore semplice, cioé di molteplicitá algebrica 1, mentre $\lambda_2 = 2$ é un autovalore doppio, cioé di molteplicitá algebrica 2.

L'autospazio corrispondente all'autovalore $\lambda_1 = 0$ si ottiene risolvendo il sistema lineare omogeneo $(A - 0 \cdot I)\mathbf{x} = \mathbf{0}$, cioé $A\mathbf{x} = \mathbf{0}$, ovvero:

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

L'autospazio associato a $\lambda_1 = 0$ é la retta generata dall'autovettore $\mathbf{u}_1 = (-1, 1, 0)^T$.

Invece gli autovettori corrispondenti all'autovalore doppio $\lambda_2 = 2$ si trovano risolvendo il sistema lineare omogeneo $(A - 2I)\mathbf{x} = \mathbf{0}$, cioé:

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

L'autospazio corrispondente all'autovalore $\lambda_2 = 2$ é il piano in $\mathbf{R}^3$ di equazione $x_1 - x_2 = 0$; esso é generato dagli autovettori $\mathbf{u_2} = (1, 1, 0)^T$ ed $\mathbf{u_3} = (0, 0, 1)^T$. Pertanto $\lambda_2 = 2$ é un autovalore regolare: la sua molteplicitá geometrica é 2, al pari della molteplicitá algebrica.

■

Esempio 3 - . Cerchiamo gli autovalori e gli autospazi della matrice

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Si tratta di risolvere l'equazione caratteristica

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & -1 & -\lambda \end{pmatrix} = 0.$$

Sviluppando il calcolo del determinante con la regola di Laplace, secondo gli elementi della prima riga, si trova che il polinomio caratteristico $p_A(\lambda)$ di A é

$$p_A(\lambda) = (1-\lambda)(\lambda^2 - \lambda + 1) - (1-\lambda) = (1-\lambda)(\lambda^2 - \lambda) = -\lambda(\lambda - 1)^2.$$

Gli autovalori di A sono quindi $\lambda_1 = 0$, (autovalore semplice), e $\lambda_2 = 1$, (autovalore doppio).

Per quanto riguarda l'autospazio corrispondente a λ_1, occorre risolvere il sistema lineare omogeneo $(A - 0 \cdot I)\mathbf{x} = \mathbf{0}$, cioé:

x_1	x_2	x_3	
1	0	1	0
0	1	1	0
1	−1	0	0

x_1	x_2	x_3	
1	0	1	0
0	1	1	0
0	−1	−1	0

x_1	x_2	x_3	
1	0	1	0
0	1	1	0
0	0	0	0

L'autospazio associato all'autovalore $\lambda_1 = 0$ é quindi la retta generata dall' autovettore $\mathbf{u}_1 = (-1, -1, 1)^T$.

Invece gli autovettori corrispondenti all'autovalore doppio $\lambda_2 = 1$ si trovano risolvendo il sistema lineare omogeneo $(A - I)\mathbf{x} = \mathbf{0}$, cioé:

x_1	x_2	x_3	
1	−1	−1	0
0	0	1	0
0	0	1	0

x_1	x_3	x_2	
1	0	−1	0
0	1	0	0
0	0	0	0

L'autospazio corrispondente all'autovalore $\lambda_2 = 1$ é la retta generata dal vettore $\mathbf{u}_2 = (1, 1, 0)^T$ ed ha quindi dimensione 1; pertanto $\lambda_2 = 1$ é un autovalore singolare: la sua molteplicitá geometrica é 1, mentre la molteplicitá algebrica é 2.

■

Osservazione 1.6 - Consideriamo la matrice A studiata nell'Esempio 1 ed osserviamo che abbiamo trovato tre autovettori $\mathbf{u}_1 = (11, 1, -14)^T$, $\mathbf{u}_2 = (-1, 1, 1)^T$ ed $\mathbf{u}_3 = (1, 1, 1)^T$, in corrispondenza degli autovalori semplici $\lambda_1 = -2$, $\lambda_2 = 1$, $\lambda_3 = 3$. E' facile constatare che tali vettori sono linearmente indipendenti, calcolando il determinante della matrice B avente per colonne tali vettori e constatando che tale determinnte é diverso da 0.

Analogamente, nel caso dell'Esempio 2, abbiamo trovato che A ha due autovalori $\lambda_1 = 0$ e $\lambda_2 = 2$; al primo autovalore corrisponde l' autovettore $(-1, 1, 0)^T$, mentre a λ_2 corrispondono due autovettori linearmente indipendenti $(1, 1, 0)^T$ e $(0, 0, 1)^T$. Anche in questo caso i tre autovettori trovati sono linearmente indipendenti, perché il determinante della matrice B avente per colonne tali vettori é chiaramente diverso da 0.

■

Naturalmente tutto questo non é un caso, poiché si dimostra la seguente

Proposizione 1.7 -. *Se $\lambda_1, \lambda_2, \ldots, \lambda_h$ sono gli autovalori di A e B_1, B_2, ... , B_h sono basi rispettivamente degli autospazi di A corrispondenti a tali autovalori, allora risulta $B_i \cap B_j = \emptyset$ per ogni $i \neq j$ e l'insieme $B = B_1 \cup B_2 \cup \ldots \cup B_h$ é un insieme di vettori linearmente indipendenti in $\mathbf{R}^n$.*

Pertanto il massimo numero di autovettori linearmente indipendenti di A é dato dal numero di elementi di B, cioé dalla somma delle molteplicitá geometriche degli autovalori di A.

La dimostrazione viene omessa per semplicitá.

■

Osservazione 1.8 - . Se A ha n autovalori $\lambda_1, \lambda_2, \ldots, \lambda_n$, (ripetendo un autovalore k volte se k é la sua molteplicitá algebrica), allora si vede facilmente che il determinante di A coincide con il prodotto degli autovalori di A, cioé:

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n.$$

Infatti, detto $p_A(\lambda)$ il polinomio caratteristico di A, si ha:

$$\det(A - \lambda I) = p_A(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \ldots a_1\lambda + a_0 = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda);$$

ne segue che $\det(A) = p_A(0) = (\lambda_1 - 0)(\lambda_2 - 0) \cdot \ldots \cdot (\lambda_n - 0) = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n$.

■

Osservazione 1.9 - . Sappiamo che, per il Teorema fondamentale dell'Algebra, un qualunque polinomio ha almeno una radice nel campo $\mathbf{C}$ dei numeri complessi ed infatti l'ambiente naturale per studiare i polinomi é l'ambiente dei numeri complessi.

Di conseguenza, riesce naturale estendere le precedenti nozioni di autovalore ed autovettore alle matrici quadrate A di numeri complessi, nel senso che un numero complesso λ dicesi autovalore di A se e solo se esiste un vettore non nullo $\mathbf{z} \in \mathbf{C}^n$ tale che $A\mathbf{z} = \lambda\mathbf{z}$; in tal caso $\mathbf{z}$ dicesi autovettore di A corrispondente all'autovalore λ.

Si vede allora che $\lambda \in \mathbf{C}$ é autovalore di A se e solo se λ é radice dell'equazione caratteristica $\det(A - \lambda I) = 0$, (ovviamente dopo aver preventivamente esteso in modo ovvio il concetto di determinante di una matrice quadrata di numeri complessi).

Adesso qualunque matrice quadrata ha almeno un autovalore; anzi, qualunque matrice quadrata di ordine n ha esattamente n autovalori complessi, contando ciascun autovalore k volte, se k é la sua molteplicitá algebrica. Il prodotto di tali n autovalori coincide con il determinante di A, con lo stesso ragionamento illustrato nella precedente Osservazione 1.6.

In particolare, se A é una matrice quadrata di numeri reali, allora le radici non reali del polinomio caratteristico di A sono coppie di numeri complessi coniugati, con la stessa molteplicitá, e quindi gli autovalori non reali di A sono coppie di numeri complessi coniugati, con la stessa molteplicitá. Inoltre se $\mathbf{z}$ é un autovettore di A corrispondente all'autovalore non reale $\lambda = \alpha + i\beta$, allora il vettore $\bar{\mathbf{z}}$, avente per componenti i coniugati delle componenti di $\mathbf{z}$, é un autovettore corrispondente all'autovalore $\bar{\lambda} = \alpha - i\beta$, coniugato di λ.

Nel seguito continueremo a considerare autovalori ed autovettori reali, ma quello che diremo potrebbe (in parte) essere esteso al caso degli autovalori ed autovettori complessi.

■

2. - Matrici simili - Matrici diagonalizzabili

Ricordiamo che se A ed A' sono matrici quadrate di ordine n, si dice che A' é **simile** ad A se esiste una matrice quadrata invertibile B tale che $A' = B^{-1}AB$.

Si vede allora facilmente che:

(1) A' é simile ad A se e solo se A é simile ad A',
(2) se A' é simile ad A ed A'' é simile ad A', allora A'' é simile ad A.

Infatti si ha

$$A' = B^{-1}AB \qquad \Longleftrightarrow \qquad A = BA'B^{-1} = (B^{-1})^{-1}A(B^{-1}).$$

D'altra parte, se risulta $A' = B^{-1}AB$ ed $A'' = C^{-1}A'C$, allora si ha:

$$A'' = C^{-1}A'C = C^{-1}(B^{-1}AB)C = (C^{-1}B^{-1})A(BC) = (BC)^{-1}A(BC).$$

Sussiste inoltre la seguente

Proposizione 2.1. *Se A ed A' sono matrici simili, allora A ed A' hanno la stessa caratteristica, lo stesso determinante, lo stesso polinomio caratteristico e quindi gli stessi autovalori.*

Inoltre gli autovalori di A ed A' hanno la stessa molteplicitá algebrica e geometrica.

Dim. Infatti se A' é simile ad A e B é una matrice quadrata invertibile tale che $A' = B^{-1}AB$, allora risulta:

$$\begin{aligned} p_{A'}(\lambda) &= \det(A' - \lambda I) = \det(B^{-1}AB - \lambda B^{-1}B) = \det(B^{-1}(A - \lambda I)B) = \\ &= \det(B^{-1}) \cdot \det(A - \lambda I) \cdot \det(B) = \big(1/\det B\big) \cdot \det(A - \lambda I) \cdot \det B = \det(A - \lambda I) = p_A(\lambda). \end{aligned}$$

In particolare risulta: $\det(A') = p_{A'}(0) = p_A(0) = \det A$.

Inoltre α é autovalore di A di molteplicitá algebrica h se e solo se il polinomio caratteristico $p_A(\lambda) = p_{A'}(\lambda)$ é divisibile per $(\lambda - \alpha)^h$ e non é divisibile per $(\lambda - \alpha)^{h+1}$ e quindi se e solo se α é autovalore di A' di molteplicitá algebrica h.

Per quanto riguarda la molteplicitá geometrica degli autovalori, sia λ un autovalore di A ed A' e siano $U_A(\lambda)$ ed $U_{A'}(\lambda)$ i relativi autospazi; vogliamo dimostrare che $U_A(\lambda)$ ed $U_{A'}(\lambda)$ hanno la stessa dimensione.

Ebbene si ha chiaramente che

$$\mathbf{u} \in U_{A'}(\lambda) \iff A'\mathbf{u} = \lambda\mathbf{u} \iff AB\mathbf{u} = \lambda B\mathbf{u} \iff B\mathbf{u} \in U_A(\lambda).$$

e quindi risulta $U_A(\lambda) = L_B(U_{A'}(\lambda))$, dove L_B é la trasformazione lineare associata alla matrice quadrata B. Tale trasformazione lineare é bigettiva, dal momento che B é invertibile, e quindi per il Teorema 2.1 del Cap. V si ha che $U_{A'}(\lambda)$ ed $L_B(U_{A'}(\lambda)) = U_A(\lambda)$ hanno la stessa dimensione.

Infine, per quanto riguarda la caratteristica di A ed A', sappiamo che, detti V e V' i sottospazi di $\mathbf{R}^n$ generati dalle colonne di A e A' rispettivamente, si ha che

$$car(A) = \dim(V), \quad car(A') = \dim(V'), \quad V = Im(L_A), \quad V' = Im(L_{A'}) \ .$$

Se proviamo che $V = L_B(V')$, ancora per il Teorema 2.1 del Cap. V avremo che V e V' hanno la stessa dimensione e quindi A ed A' hanno la stessa caratteristica.

Ebbene, se $\mathbf{y} \in V' = Im(L_{A'})$, allora esiste $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{y} = L_{A'}(\mathbf{x}) = A'\mathbf{x} = B^{-1}AB\mathbf{x}$ e quindi $B\mathbf{y} = AB\mathbf{x} \in Im(L_A) = V$; questo prova che $L_B(V') \subseteq V$.

Viceversa se $\mathbf{z} \in V = Im(L_A)$, allora esiste $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{z} = L_A(\mathbf{x}) = A\mathbf{x}$; ne segue che

$$\mathbf{y} = B^{-1}\mathbf{z} = B^{-1}A\mathbf{x} = B^{-1}AB(B^{-1}\mathbf{x}) = A'(B^{-1}\mathbf{x}) = L_{A'}(B^{-1}\mathbf{x}) \in Im(L_{A'}) = V'$$

e quindi $\mathbf{z} = B\mathbf{y} = L_B(\mathbf{y}) \in L_B(V')$. Pertanto risulta $V \subseteq L_B(V')$ e dunque $V = L_B(V')$, come volevasi.

■

Osservazione 2.2 - Abbiamo visto nel paragrafo n. 6 del Cap. V che due matrici quadrate di ordine n sono simili se e solo se rappresentano la stessa trasformazione lineare di $\mathbf{R}^n$ in sé rispetto a due diverse basi di $\mathbf{R}^n$.

E' questa la ragione profonda per cui il determinante, il polinomio caratteristico e gli autovalori, con la relativa molteplicitá algebrica e geometrica, non cambiano quando si passa da una matrice ad una matrice simile; il motivo é che tali proprietá sono **intrinseche** alla trasformazione lineare, e non dipendono dalla matrice che la rappresenta rispetto ad una particolare base.

■

Siamo ora in grado di affrontare e risolvere il problema formulato alla fine del Capitolo V di stabilire sotto quali condizioni é possibile trovare una base in $\mathbf{R}^n$ rispetto a cui una trasformazione lineare puó essere rappresentata da una matrice diagonale Λ, o equivalentemente sotto quali condizioni una matrice quadrata é simile ad una matrice diagonale.

A tal fine é utile premettere la seguente

Definizione 2.3. *Si dice che una matrice quadrata A é* **diagonalizzabile***, se é simile ad una matrice diagonale Λ, cioé se esiste una matrice quadrata invertibile B tale che $B^{-1}AB$ sia una matrice diagonale.*

In tal caso si dice che A é diagonalizzata da B o che la matrice B diagonalizza la matrice A.

■

Si ha allora la seguente

Proposizione 2.4. *Una matrice quadrata A é diagonalizzabile se e solo se esistono n autovettori linearmente indipendenti di A e quindi se e solo se in $\mathbf{R}^n$ esiste una base formata da autovettori di A.*

Piú esattamente si ha che:

a) *se B é una matrice quadrata invertibile che diagonalizza A, allora le colonne di B sono n autovettori linearmente indipendenti di A, i corrispondenti autovalori sono gli elementi diagonali della matrice diagonale $\Lambda = B^{-1}AB$. Le colonne di B formano una base di $\mathbf{R}^n$ rispetto a cui la trasformazione lineare associata ad A é rappresentata dalla matrice Λ.*

b) *Viceversa, se esistono n autovettori linearmente indipendenti di A, allora la matrice B avente per colonne tali autovettori é una matrice che diagonalizza A e gli elementi diagonali di $\Lambda = B^{-1}AB$ sono gli n autovalori di A.*

Dim. di a). Siano $\lambda_1, \lambda_2, \ldots, \lambda_n$ gli elementi diagonali della matrice $\Lambda = B^{-1}AB$ e siano $E^1, E^2, \ldots, E^n$ le colonne della matrice identica I. Dall'essere $B^{-1}AB = \Lambda$ segue che $AB = B\Lambda$ e quindi, per ogni $j = 1, 2, \ldots, n$, si ha

$$AB^j = (AB)^j = (B\Lambda)^j = B\Lambda^j = B(\lambda_j E^j) = \lambda_j(BE^j) = \lambda_j B^j.$$

Questo dimostra che $\lambda_1, \lambda_2, \ldots, \lambda_n$ sono autovalori di A e che le colonne di B sono i corrispondenti autovettori. D'altra parte, le colonne di B sono linearmente indipendenti, dal momento che B é invertibile; questo prova completamente la tesi.

Dim. di b). Siano $\lambda_1, \lambda_2, \ldots, \lambda_n$ gli autovalori di A corrispondenti alle colonne di B e sia Λ la matrice diagonale, i cui elementi diagonali sono (nell'ordine) $\lambda_1, \lambda_2, \ldots, \lambda_n$. Allora, per ogni $j = 1, 2, \ldots, n$ si ha:

$$(AB)^j = AB^j = \lambda_j B^j = \lambda_j BE^j = B(\lambda_j E^j) = B\Lambda^j = (B\Lambda)^j.$$

Questo prova che $AB = B\Lambda$, e quindi che $B^{-1}AB = \Lambda$, come volevasi.

■

A questo punto, per sapere se una matrice é diagonalizzabile oppure no, occorre sapere sotto quali condizioni una matrice quadrata di ordine n possiede n autovettori linearmente indipendenti.

Dalle Proposizioni 2.4 e 1.7 si deduce immediatamente la seguente

Proposizione 2.5 -. *Una matrice A é diagonalizzabile se e solo se la somma delle molteplicitá geometriche degli autovalori $\lambda_1, \lambda_2, \ldots, \lambda_h$ di A é uguale ad n, cioé se e solo se risulta:*

(1) $$dim(U(\lambda_1)) + \dim(U(\lambda_2)) + \ldots + dim(U(\lambda_h)) = n.$$

Pertanto A é diagonalizzabile se il polinomio caratteristico p_A di A ha n radici reali e distinte oppure se p_A ha solo radici reali ed ogni radice di p_A é un autovalore regolare.

Invece A non é diagonalizzabile se A ha almeno un autovalore singolare oppure se il polinomio caratteristico di A ha radici complesse non reali.

Dim. Infatti, detta B l'unione delle basi degli autospazi di A, per la prop. 1.7 si ha che il numero degli elementi di B, (cioé la somma delle molteplicitá geometriche degli autovalori di A), rappresenta il massimo numero di autovettori linearmente indipendenti di A; ne segue che A é diagonalizzabile o meno a seconda che il numero di elementi di B é uguale o minore di n.

Ebbene, il numero di elementi di B é minore di n se A ha almeno un autovalore singolare oppure se il polinomio caratteristico di A ha radici complesse non reali, ed é invece uguale ad n, se il polinomio caratteristico di A ha n radici reali e distinte oppure se le radici del polinomio caratteristico sono tutte reali, (e quindi la somma delle molteplicitá algebriche degli autovalori di A é uguale ad n), e la molteplicitá algebrica di ogni autovalore coincide con la sua molteplicitá geometrica. Questo dimostra la tesi.

■

Ad esempio, delle matrici

$$\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

studiate rispettivamente negli Esempi 1, 2, 3 si ha che la prima é diagonalizzabile perché ha 3 autovalori reali e distinti, la seconda é diagonalizzabile perché la somma delle molteplicitá geometriche dei suoi due autovalori é 3; invece la terza non é diagonalizzabile, perché uno dei suoi autovalori aveva molteplicitá geometrica 1 e molteplicitá algebrica 2.

■

Esempio 4 - . Consideriamo la matrice quadrata

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ \alpha & 0 & -1 \end{pmatrix}.$$

Ci chiediamo, (al variare del parametro α) se la matrice A é diagonalizzabile oppure no.

A tal fine consideriamo il polinomio caratteristico:

$$p_A(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} 3-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ \alpha & 0 & -1-\lambda \end{pmatrix} = (2-\lambda)\cdot(\lambda^2 - 2\lambda - 3 - \alpha).$$

Le radici di p_A sono quindi $\lambda_1 = 2$, $\lambda_2 = 1 - \sqrt{\alpha+4}$, $\lambda_3 = 1 + \sqrt{\alpha+4}$. Ne segue che se $\alpha < -4$, allora p_A ha una radice reale e due radici complesse coniugate e quindi A non é diagonalizzabile.

Invece A é diagonalizzabile se le radici di p_A sono reali e distinte: questo accade se risulta $\alpha + 4 > 0$ e $2 \neq 1 \pm \sqrt{\alpha+4}$, cioé se α é maggiore di -4 e diverso da -3.

Ad esempio, per $\alpha = 0$, gli autovalori sono $\lambda_1 = 2,\ \lambda_2 = 1 - \sqrt{4} = -1,\ \lambda_3 = 1 + \sqrt{4} = 3.$ I corrispondenti autovettori si trovano risolvendo rispettivamente i sistemi

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 4 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}, \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -4 & 0. \end{array}$$

Si trova cosí che gli autovettori corrispondenti agli autovalori $\lambda_1 = 2,\ \lambda_2 = -1,\ \lambda_3 = 3$ sono rispettivamente $\mathbf{u}_1 = (0,1,0)^T,\ \mathbf{u}_2 = (-1,0,4)^T,\ \mathbf{u}_3 = (1,0,0)^T$.

La matrice B avente per colonne tali vettori é la matrice che diagonalizza A; si ha infatti

$$AB = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 3 \\ 2 & 0 & 0 \\ 0 & -4 & 0 \end{pmatrix},$$

$$B\Lambda = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 3 \\ 2 & 0 & 0 \\ 0 & -4 & 0 \end{pmatrix}.$$

Si ha quindi $AB = B\Lambda$; questo prova che $B^{-1}AB = \Lambda$, cioé che A é diagonalizzata da B.

Se $\alpha = -4$, si ha che A ha un autovalore semplice, $(\lambda = 2)$, ed uno doppio, $(\lambda = 1)$. Gli autovettori relativi all'autovalore $\lambda = 2$ si trovano risolvendo il sistema

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -4 & 0 & -3 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

L'autospazio relativo all'autovalore $\lambda_1 = 2$ é quindi la retta generata dal vettore $\mathbf{u}_1 = (0,1,0)^T$.
Gli autovettori relativi all'autovalore doppio $\lambda = 1$ si trovano risolvendo il sistema

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -4 & 0 & -2 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

L'autospazio relativo all'autovalore doppio $\lambda = 1$ é quindi la retta generata dal vettore $\mathbf{u}_1 = (1,0,-2)^T$; perció l'autovalore $\lambda = 1$ é un autovalore singolare di A. Se ne deduce che per $\alpha = -4$ la matrice A ha al massimo due autovettori linearmente indipendenti e quindi NON é diagonalizzabile.

Se $\alpha = -3$, si ha che A ha un autovalore semplice , $(\lambda = 1 - \sqrt{1} = 0)$, ed uno doppio, $(\lambda = 2)$. Gli autovettori relativi all'autovalore $\lambda = 0$ si trovano risolvendo il sistema

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 3 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ -3 & 0 & -1 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

L'autospazio relativo all'autovalore $\lambda = 0$ é quindi la retta generata dal vettore $\mathbf{u}_1 = (1,0,-3)^T$.
Invece, gli autovettori relativi all'autovalore doppio $\lambda = 2$ si trovano risolvendo il sistema

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & 0 & -3 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

Il sistema $(A - 2I)\mathbf{x} = \mathbf{0}$ si riduce quindi all'equazione $x_1 + x_3 = 0$ o equivalentemente $x_1 = 0x_2 - x_3$ che ha due soluzioni linearmente indipendenti $(0,1,0)^T$ e $(-1,0,1)^T$. L'autospazio relativo all'autovalore doppio $\lambda = 2$ é quindi il piano generato dai vettori $(0,1,0)^T$ e $(-1,0,1)^T$; pertanto gli autovalori di A sono regolari, la somma delle loro molteplicitá geometriche é uguale a 3 e quindi A é diagonalizzabile.
I vettori $(1,0,-3)^T$, $(0,1,0)^T$ e $(-1,0,1)^T$ costituiscono una base di $\mathbf{R}^3$ formata da autovettori di A; la matrice B avente per colonne tali vettori diagonalizza A nel senso che $B^{-1}AB$ é la matrice diagonale avente per elementi diagonali i numeri $\lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 2$. In effetti si ha

$$A \cdot B = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ -3 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$B \cdot \Lambda = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

■

3. - Il caso delle matrici simmetriche

Un caso particolarmente interessante é costituito dalle matrici quadrate simmetriche. In tal caso, non solo si ha che A é diagonalizzabile, ma si puó dimostrare che esiste una base di $\mathbf{R}^n$ formata da autovettori a due a due ortogonali. Si ha infatti il seguente

Teorema 3.1 -. *Se A é una matrice quadrata simmetrica di ordine n, allora:*

(1) *le radici del polinomio caratteristico di A sono tutte reali,*
(2) *ogni autovalore λ di A é regolare, cioé tale che la molteplicitá geometrica di λ coincide con la sua molteplicitá algebrica,*
(3) *due autovettori corrispondenti ad autovalori distinti di A sono ortogonali,*
(4) *esiste una base ortonormale di $\mathbf{R}^n$ formata da autovettori di A.*

Ne segue che ogni matrice quadrata simmetrica é diagonalizzabile, ed anzi puó essere diagonalizzata da una matrice ortogonale B; in altri termini, esiste una matrice ortogonale B tale che $B^T AB = B^{-1}AB$ coincide con la matrice diagonale Λ, i cui elementi diagonali sono gli autovalori di A.

Dim. La dimostrazione di (1) e (2) viene omessa per semplicitá.

Dim. di (3). Siano $\mathbf{u}_1$ ed $\mathbf{u}_2$ due autovettori relativi agli autovalori distinti λ_1 e λ_2, si ha allora: $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$, $A\mathbf{u}_2 = \lambda_2\mathbf{u}_2$. D'altra parte, il fatto che A é simmetrica significa che $A = A^T$; ne segue che

$$\lambda_1(\mathbf{u}_1|\mathbf{u}_2) = (\lambda_1\mathbf{u}_1|\mathbf{u}_2) = (A\mathbf{u}_1)|\mathbf{u}_2) = \mathbf{u}_2^T A\mathbf{u}_1 = \mathbf{u}_2^T A^T\mathbf{u}_1 = (A\mathbf{u}_2)^T\mathbf{u}_1 = (\lambda_2\mathbf{u}_2)_1^{\mathbf{u}} = \lambda_2(\mathbf{u}_2|\mathbf{u}_1),$$

e dunque che $(\lambda_1 - \lambda_2)(\mathbf{u}_1|\mathbf{u}_2) = 0$. Se ne deduce, (essendo $\lambda_1 - \lambda_2 \neq 0$), che $(\mathbf{u}_1|\mathbf{u}_2) = 0$, come volevasi.

Dim. di (4). Per dimostrare la (4), procediamo come segue. Per ogni autovalore λ di A scegliamo una base dell'autospazio $U(\lambda)$; adoperando l'algoritmo di Gram Schmidt, é possibile modificare tale base in modo da ottenere una base ortonormale dell'autospazio. Per la (3), l'unione di tali basi é un insieme ortonormale di vettori di $\mathbf{R}^n$, e quindi tali vettori sono linearmente indipendenti per la Prop. 4.1 del Cap. VI. Infine, per la (1) e la (2), il numero complessivo di tali vettori é n e quindi essi formano una base di $\mathbf{R}^n$.
Da (4) segue che, detta B la matrice avente per colonne i vettori di tale base, da una parte si ha che tale matrice é una matrice ortogonale, e dall'altra (per la Prop. 2.4) si ha che essa diagonalizza A; se ne deduce che $B^T AB = B^{-1}AB$ é la matrice diagonale i cui elementi diagonali sono gli autovalori di A.

■

Esempio 1 - Consideriamo la matrice quadrata simmetrica

$$A = \begin{pmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{pmatrix}$$

Calcolando il determinante con la regola di Sarrus o con il metodo di Laplace, si trova (con un po' di calcoli) che il polinomio caratteristico di A é

$$\det(A - \lambda I) = \det \begin{pmatrix} 5-\lambda & 2 & -2 \\ 2 & 5-\lambda & -2 \\ -2 & -2 & 5-\lambda \end{pmatrix} = -\lambda^3 + 15\lambda^2 - 63\lambda + 81.$$

Eventuali radici intere del polinomio caratteristico vanno cercate tra i divisori di 81; procedendo per tentativi si vede che una radice del polinomio caratteristico é $\lambda_1 = 3$.
Le altre radici di p_A si trovano dividendo il polinomio $p_A(\lambda)$ per il polinomio $\lambda - 3$; si ottiene cosí il polinomio $-\lambda^2 + 12\lambda - 27$. Risolvendo l'equazione $\lambda^2 - 12\lambda + 27 = 0$, si trovano le radici $\lambda_2 = 3$ e $\lambda_3 = 9$.
In definitiva gli autovalori di A sono $\lambda_1 = \lambda_2 = 3$, autovalore doppio, e $\lambda_3 = 9$, autovalore semplice.
Gli autovettori di A corrispondenti all'autovalore doppio $\lambda_1 = \lambda_2 = 3$ si trovano risolvendo il sistema lineare omogeneo $(A - 3I)\mathbf{x} = \mathbf{0}$, cioé :

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 2 & 2 & -2 & 0 \\ 2 & 2 & -2 & 0 \\ -2 & -2 & 2 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

L'autospazio relativo all'autovalore doppio $\lambda_1 = 3$ é quindi il piano (in $\mathbf{R}^3$) di equazione $x_1 + x_2 - x_3 = 0$; una base di tale autospazio é costituita dai vettori $\mathbf{u}_1 = (1, 0, 1)^T$ e $\mathbf{u}_2 = (0, 1, 1)^T$.

Invece gli autovettori di A corrispondenti all'autovalore $\lambda_3 = 9$ si trovano risolvendo il sistema lineare omogeneo $(A - 9I)\mathbf{x} = \mathbf{0}$, cioé :

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline -4 & 2 & -2 & 0 \\ 2 & -4 & -2 & 0 \\ -2 & -2 & -4 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & -2 & -1 & 0 \\ 0 & -6 & -6 & 0 \\ 0 & -6 & -6 & 0 \end{array} \qquad \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0. \end{array}$$

L'autospazio relativo all'autovalore $\lambda_3 = 9$ é quindi la retta generata dal vettore $\mathbf{u}_3 = (-1, -1, 1)^T$.

Si noti che gli autovalori di A sono tutti reali e per entrambi si ha che la molteplicitá geometrica coincide con la molteplicitá algebrica, come prevedono la (1) e la (2) del Teorema 3.1.

Inoltre, i vettori $\mathbf{u}_1$, $\mathbf{u}_2$ ed $\mathbf{u}_3$ sono linearmente indipendenti, poiché il determinante della matrice avente per colonne tali vettori é diverso da 0 . Pertanto la matrice A é diagonalizzabile.

Inoltre, $\mathbf{u}_3$ é ortogonale sia ad $\mathbf{u}_1$ che ad $\mathbf{u}_2$, come previsto dalla (3).

Per ottenere una base ortonormale occorre innanzitutto sostituire la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ dell'autospazio relativo all'autovalore $\lambda_1 = 3$ con una base ortogonale. A tal fine, basta porre :

$$\mathbf{v}_1 = \mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_2 = \mathbf{u}_2 - \frac{(\mathbf{u}_2|\mathbf{v}_1)}{(\mathbf{v}_1|\mathbf{v}_1)}\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix}.$$

Una base ortogonale di autovettori di A é costituita dai vettori $\mathbf{v}_1 = (1,\ 0,\ 1)^T$, $\mathbf{v}_2 = (-1/2,\ 1,\ 1/2)^T$ e $\mathbf{v}_3 = \mathbf{u}_3 = (-1, -1, 1)^T$; per avere una base ortonormale occorre dividere ciascuno di tali vettori per la sua norma. La base ortonormale di autovettori di A é quindi $\{\mathbf{v}_1^*, \mathbf{v}_2^*, \mathbf{v}_3^*\}$, dove

$$\mathbf{v}_1^* = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \qquad \mathbf{v}_2^* = \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \qquad \mathbf{v}_3^* = \begin{pmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

La matrice

$$B = (\,\mathbf{v}_1^* \quad \mathbf{v}_2^* \quad \mathbf{v}_3^*\,) = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

é una matrice ortogonale. La sua inversa é

$$B^{-1} = B^T = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}.$$

e risulta:

$$B^TAB = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} =$$

$$= \begin{pmatrix} 3/\sqrt{2} & 0 & 3/\sqrt{2} \\ -3/\sqrt{6} & 6/\sqrt{6} & 3/\sqrt{6} \\ -9/\sqrt{3} & -9/\sqrt{3} & 9/\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \Lambda.$$

■

E S E R C I Z I

1 - Trovare gli autovalori, gli autovettori e gli autospazi delle matrici

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 5 & -1 & 0 \\ -9 & 3 & -6 \\ 3 & 1 & 8 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 1 & \alpha \\ 1 & 0 & \alpha \\ -1 & -1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -4 & 1 & -6 \\ 0 & 1 & 0 \\ 3 & -1 & 5 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & -3 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -2 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 3 & 2 & -2 \\ -2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & -2 \\ -2 & 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 2 \\ -2 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix},$$

Dire quindi se tali matrici sono diagonalizzabili; in caso affermativo, trovare una base di $\mathbf{R}^3$ formata di autovettori. e indicare una matrice che le diagonalizza.

2 - Date le matrici

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & 0 \\ -h & 2 & 3 \end{pmatrix} \qquad \begin{pmatrix} -2 & 1 & -h \\ -1 & 2 & -h \\ -1 & -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & h \\ 3 & 2 & h \end{pmatrix}$$

a) trovare gli autovalori e i corrispondenti autospazi di tali matrici;
b) dire per quali valori di h tali matrici sono diagonalizzabili; per tali valori del parametro h trovare una base di $\mathbf{R}^3$ formata di autovettori.

3 - Trovare gli autovalori e gli autospazi delle matrici

$$\begin{pmatrix} -2 & 2 & -2 \\ 2 & -5 & 4 \\ -2 & 4 & -5 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \qquad \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0 & -3 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 2 & 0 & -1 & 1 \end{pmatrix}.$$

Verificare che tali matrici sono diagonalizzabili e che gli autovettori corrispondenti ad autovalori distinti sono a due a due ortogonali; trovare quindi una base ortonormale di autovettori.
Indicare una matrice B che le diagonalizza e verificare che B e B^T sono matrici ortogonali e che $B^{-1} = B^T$.

Capitolo VIII - FORME QUADRATICHE

1 . - Definizioni e prime proprietá

Definizione 1.1. *Sia A una matrice quadrata simmetrica di ordine n; si dice allora* **forma quadratica** *associata alla matrice A la funzione $\varphi_A : \mathbf{R}^n \mapsto \mathbf{R}$ che ad ogni $\mathbf{x} \in \mathbf{R}^n$ associa il numero reale*

$$\varphi_A(\mathbf{x}) = (\mathbf{x}|A\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = (x_1, x_2, \dots, x_n) \cdot \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \dots\dots\dots\dots \\ a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,n}x_n \end{pmatrix} =$$

$$\begin{aligned} &= a_{1,1}x_1^2 + a_{1,2}x_1x_2 + \dots + a_{1,n}x_1x_n + a_{2,1}x_2x_1 + a_{2,2}x_2^2 + \dots + a_{2,n}x_2x_n + \\ &\quad + \dots\dots\dots\dots + a_{n,1}x_nx_1 + a_{n,2}x_nx_2 + \dots + a_{n,n}x_n^2 = \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_ix_j = \sum_{i=1}^{n} a_{i,i}x_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_{i,j}x_ix_j. \end{aligned}$$

■

Ad esempio la forma quadratica associata alla matrice

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

é la funzione che associa ad ogni $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbf{R}^3$ il numero reale

$$\varphi_A(\mathbf{x}) = (x_1, x_2, x_3) \begin{pmatrix} 2 & 3 & -1 \\ 3 & 4 & -2 \\ -1 & -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (x_1, x_2, x_3) \begin{pmatrix} 2x_1 + 3x_2 - x_3 \\ 3x_1 + 4x_2 - 2x_3 \\ -x_1 - 2x_2 + 5x_3 \end{pmatrix} =$$

$$\begin{aligned} &= 2x_1^2 + 3x_1x_2 - x_1x_3 + 3x_2x_1 + 4x_2^2 - 2x_2x_3 - x_3x_1 - 2x_3x_2 + 5x_3^2 = \\ &= 2x_1^2 + 4x_2^2 + 5x_3^2 + 6x_1x_2 - 2x_1x_3 - 4x_2x_3. \end{aligned}$$

Chiaramente φ_A é un polinomio omogeneo di secondo grado nelle variabili $x_1, x_2, \dots, x_n$; inoltre, é evidente che risulta

$$\varphi_A(\alpha\mathbf{x}) = \alpha^2\varphi_A(\mathbf{x}), \qquad \text{per ogni } \alpha \in \mathbf{R},\ \mathbf{x} \in \mathbf{R}^n .$$

■

Le forme quadratiche possono essere classificate in base al segno dei valori che φ_A assume. Si ha infatti la seguente

Definizione 1.2. *Si dice che la matrice A (o la forma quadratica φ_A associata ad A) é:*

a) **definita positiva** *se risulta $\varphi_A(\mathbf{x}) > 0$ per ogni $\mathbf{x} \neq \mathbf{0}$,*
b) **definita negativa** *se risulta $\varphi_A(\mathbf{x}) < 0$ per ogni $\mathbf{x} \neq \mathbf{0}$,*
c) **semidefinita positiva** *se risulta $\varphi_A(\mathbf{x}) \geq 0$ per ogni $\mathbf{x}$ ed esiste $\mathbf{u} \neq \mathbf{0}$ tale che $\varphi_A(\mathbf{u}) = 0$,*
d) **semidefinita negativa** *se risulta $\varphi_A(\mathbf{x}) \leq 0$ per ogni $\mathbf{x}$ ed esiste $\mathbf{u} \neq \mathbf{0}$ tale che $\varphi_A(\mathbf{u}) = 0$,*
e) **indefinita** *se esistono $\mathbf{u} \neq \mathbf{0}$ e $\mathbf{v} \neq \mathbf{0}$ tali che $\varphi_A(\mathbf{u}) > 0$ e $\varphi_A(\mathbf{v}) < 0$.*

Studiare il **segno** *di A, o della forma quadratica associata ad A, significa stabilire se A é definita positiva o definita negativa, o semidefinita o indefinita.*

■

Osservazione 1.3 - Il fatto che la forma quadratica sia definita positiva (rispett. negativa) significa che φ_A é illimitata superiormente (rispett. inferiormente), che il valore minimo (rispett. massimo) di φ é 0 e che $\mathbf{0}$ é l'unico suo punto di minimo (rispett. massimo); invece il fatto che φ_A sia semidefinita positiva (rispett. negativa) significa che φ_A é illimitata superiormente (rispett. inferiormente), che il valore minimo (massimo) di φ é 0, ma che φ_A ha infiniti punti di minimo (rispett. massimo). Infine φ_A é indefinita se é illimitata sia inferiormente che superiormente e $\mathbf{0}$ é un punto di sella per φ_A, cioé é punto di minimo su una retta passante per l'origine e punto di massimo su un'altra retta.

■

Al fine di trovare il segno di una forma quadratica é opportuno trasformare A in una matrice che sia in un certo senso equivalente ad A, ma di cui sia piú facile studiare il segno. Questo ci porta ad introdurre il concetto di **congruenza** di due matrici quadrate simmetriche.

2 - Matrici congruenti

Definizione 2.1. *Se A ed A' sono matrici quadrate simmetriche di ordine n, si dice che A é* **congruente** *ad A' se esiste una matrice quadrata invertibile B tale che $A' = B^TAB$.*

■

Osservazione 2.2 - Si vede facilmente che la relazione di congruenza tra matrici gode delle proprietá riflessiva, simmetrica e transitiva, nel senso che

(1) ogni matrice quadrata A é congruente a sé stessa,
(2) se A é congruente ad A', allora A' é congruente ad A,
(3) se A é congruente ad A' ed A' é congruente ad A'', allora A é congruente ad A''.

Infatti, la (1) deriva dal fatto che risulta $A = I_nAI_n = I_n^TAI_n$.

La (2) deriva dal fatto che dall'essere $A' = B^TAB$ si deduce che $A = (B^T)^{-1}A'B^{-1} = (B^{-1})^TA'B^{-1}$.

Infine la (3) discende dal fatto che, se esistono due matrici quadrate invertibili B e C tali che $A' = B^TAB$ e $A'' = C^TA'C$, allora si ha $A'' = C^TA'C = C^T(B^TAB)C = (C^TB^T)A(BC) = (BC)^TA(BC)$.

■

Osservazione 2.3 - Possiamo renderci conto facilmente di cosa significa passare da una matrice ad una congruente. Infatti, se A ed A' sono matrici quadrate simmetriche congruenti e B é una matrice invertibile tale che $A' = B^TAB$, allora per ogni $\mathbf{x} \in \mathbf{R}^n$, posto $\mathbf{x} = B\mathbf{u} = u_1B^1 + u_2B^2 + \ldots + u_nB^n$, si ha che

(1) $u_1, u_2, \ldots , u_n$ sono le coordinate di un generico punto $\mathbf{x}$ di $\mathbf{R}^n$ rispetto alla base $\{B^1, B^2, \ldots , B^n\}$, mentre $x_1, x_2, \ldots , x_n$ sono le coordinate dello stesso punto rispetto alla base canonica $E^1, E^2, \ldots , E^n$.
(2) risulta $\varphi_A(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = (B\mathbf{u})^TA(B\mathbf{u}) = \mathbf{u}^TB^TAB\mathbf{u} = \mathbf{u}^TA'\mathbf{u} = \varphi_{A'}(\mathbf{u})$.

In altri termini, la forma quadratica é una funzione di $\mathbf{R}^n$ in $\mathbf{R}$; tale funzione viene descritta dalla matrice A se si adopera la base canonica per descrivere i punti di $\mathbf{R}^n$, viene descritta da A' se si adopera la base $\{B^1, B^2, \ldots , B^n\}$ per descrivere gli stessi punti.

Trasformare una matrice quadrata simmetrica in una congruente significa quindi cercare una nuova base rispetto a cui sia piú semplice rappresentare la forma quadratica.

L'ideale sarebbe trasformare A in una matrice congruente diagonale D, perché in tal caso si avrebbe

$$\varphi_A(\mathbf{x}) = \varphi_D(\mathbf{u}) = d_1u_1^2 + d_2u_2^2 + \ldots + d_nu_n^2,$$

dove $d_1, d_2, \ldots , d_n$ sono gli elementi diagonali di D.

Ebbene, si dimostra che ogni matrice quadrata simmetrica A é congruente ad una matrice diagonale.

Il processo di costruzione di una matrice invertibile B e di una matrice diagonale D tali che $B^T AB = D$, dicesi **riduzione in forma canonica** di A (o della forma quadratica φ_A associata ad A); nel prossimo paragrafo descriveremo tre diversi metodi per ridurre una matrice simmetrica in forma canonica.

■

3 . - Riduzione in forma canonica

Affrontiamo dunque il problema di costruire una matrice diagonale D ed una matrice invertibile B tali che $B^T AB = D$, e quindi tali che per ogni $\mathbf{x} \in \mathbf{R}^n$ risulti:

$$\varphi_A(\mathbf{x}) = \varphi_D(\mathbf{u}) = d_1 u_1^2 + d_2 u_2^2 + \ldots + d_n u_n^2, \tag{2}$$

dove $d_1, d_2, \ldots, d_n$ sono gli elementi diagonali di D e si é posto $\mathbf{x} = B\mathbf{u}$ o equivalentemente $\mathbf{u} = B^{-1}\mathbf{x}$.

Ovviamente questo puó essere realizzato in diversi modi; di seguito illustreremo tre diversi approcci.

■

Metodo degli autovalori.

Il metodo ripercorre la dimostrazione del Teorema 4.1 del Cap. VII e si articola come segue:

(1) si trovano gli autovalori e i rispettivi autospazi di A;
(2) di ogni autospazio si trova una base e da questa (mediante il procedimento di ortogonalizzazione di Gram Schmidt) si ottiene una base ortonormale;
(3) si ottengono cosí complessivamente n autovettori che formano una base ortonormale di $\mathbf{R}^n$;
(4) la matrice B avente per colonne tali autovettori é una matrice ortogonale che diagonalizza A;
(5) detta Λ la matrice diagonale avente per elementi diagonali gli autovalori di A nello stesso ordine in cui sono disposte le colonne di B, si ha che $B^T AB = B^{-1}AB = \Lambda$.

Pertanto A é congruente alla matrice diagonale Λ i cui elementi diagonali sono gli autovalori di A; posto $\mathbf{x} = B\mathbf{u}$ si ha

$$\varphi_A(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = (B\mathbf{u})^T A(B\mathbf{u}) = \mathbf{u}^T(B^T AB)\mathbf{u} = \mathbf{u}^T \Lambda \mathbf{u} = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \ldots + \lambda_n u_n^2.$$

■

Esempio 1 - Riprendiamo l'esempio 1 del n. 4 del Cap. VII; si é visto che gli autovalori della matrice

$$\begin{pmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{pmatrix}$$

sono i numeri $\lambda_1 = \lambda_2 = 3$ e $\lambda_3 = 9$. Una base dell'autospazio relativo all'autovalore doppio $\lambda_1 = \lambda_2 = 3$ é costituita dai vettori $\mathbf{u}_1 = (1, 0, 1)^T$ e $\mathbf{u}_2 = (0, 1, 1)^T$; invece l'autospazio corrispondente all'autovalore $\lambda_3 = 9$ é la retta generata dal vettore $\mathbf{u}_3 = (-1, -1, 1)^T$.

Con il procedimento di Gram Schmidt la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ del primo autospazio é stata sostituita con la base ortonormale $\{\mathbf{v}_1^*, \mathbf{v}_2^*\}$, dove $\mathbf{v}_1^* = (1/\sqrt{2},\, 0,\, 1/\sqrt{2})^T$ $\mathbf{v}_2^* = (-1/\sqrt{6},\, 2/\sqrt{6},\, 1/\sqrt{6})^T$; infine l'autovettore $\mathbf{u}_3$ é stato sostituito dal vettore normalizzato $v_3^* = (-1/\sqrt{3},\ -1/\sqrt{3},\ 1/\sqrt{3})^T$.

Si é visto che, considerate le matrici

$$B = (\,\mathbf{v}_1^* \quad \mathbf{v}_2^* \quad \mathbf{v}_3^*\,) = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{pmatrix},$$

si ha che $B^T AB = \Lambda$ Si ha dunque che A é congruente alla matrice Λ dei suoi autovalori; la forma quadratica φ_A é stata ridotta in forma canonica ponendo $\mathbf{x} = B\mathbf{y}$ e quindi

$$\varphi_A(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T B^T AB\mathbf{y} = \mathbf{y}^T \Lambda \mathbf{y} = 3y_1^2 + 3y_2^2 + 9y_3^2.$$

In effetti ponendo $\mathbf{x} = B\mathbf{y}$, cioé

$$x_1 = \frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{6}}y_2 - \frac{1}{\sqrt{3}}y_3, \qquad x_2 = \frac{2}{\sqrt{6}}y_2 - \frac{1}{\sqrt{3}}y_3, \qquad x_3 = \frac{1}{\sqrt{2}}y_1 + \frac{1}{\sqrt{6}}y_2 + \frac{1}{\sqrt{3}}y_3,$$

si verifica che

$$\varphi_A(\mathbf{x} = 5x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 4x_2x_3 = 3y_1^2 + 3y_2^2 + 9y_3^2.$$

I calcoli vengono lasciati per esercizio

■

Metodo di Lagrange.

In sostanza ridurre la forma quadratica in forma canonica equivale a cercare di rappresentare il polinomio omogeneo di secondo grado $\varphi_A(\mathbf{x})$ come somma di prodotti di opportuni coefficienti $d_1, d_2, \dots, d_n$ per i quadrati di opportuni polinomi omogenei di primo grado.

A tal fine, si sceglie una variabile e si considera il polinomio formato da tutti i monomi di $\varphi_A(\mathbf{x})$ che contengono quella variabile; a questo punto, (aggiungendo e sottraendo opportuni monomi), si fa in modo da rappresentare questo polinomio come somma di un multiplo del quadrato di un polinomio di primo grado con un polinomio in cui la variabile scelta non é piú presente.

Pertanto $\varphi_A(\mathbf{x})$ é stato rappresentato come somma di un multiplo del quadrato di un polinomio di primo grado e di un polinomio omogeneo di secondo grado in $n-1$ variabili. A questo ultimo polinomio di applica lo stesso procedimento e lo si rappresenta come somma di un multiplo del quadrato di un polinomio di primo grado e di un polinomio omogeneo di secondo grado in $n-2$ variabili.

Ripetendo il procedimento $n-1$ volte si raggiunge l'obiettivo cercato di rappresentare $\varphi_A(\mathbf{x})$ in forma canonica.

■

Esempio 2 - Consideriamo la forma quadratica associata alla matrice

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

cioé il polinomio

$$\varphi_A(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 6x_2x_3.$$

Consideriamo i termini che contengono x_1 ed (aggiungendo e sottraendo opportuni monomi) facciamo in modo da ottenere il multiplo del quadrato di un polinomio di primo grado:

$$\begin{aligned} 2x_1^2 + 4x_1x_2 - 8x_1x_3 &= 2(x_1^2 + 2x_1x_2 - 4x_1x_3) = \\ &= 2\big[x_1^2 + 2x_1x_2 + 2x_1(-2x_3) + x_2^2 + 4x_3^2 + 2x_2(-2x_3)\big] - 2x_2^2 - 8x_3^2 + 8x_2x_3 = \\ &= 2[x_1 + x_2 - 2x_3]^2 - 2x_2^2 - 8x_3^2 + 8x_2x_3. \end{aligned}$$

Ne segue che risulta

$$\begin{aligned} \varphi_A(\mathbf{x}) &= 2[x_1 + x_2 - 2x_3]^2 - 2x_2^2 - 8x_3^2 + 8x_2x_3 + 3x_2^2 + 5x_3^2 - 6x_2x_3 = \\ &= 2[x_1 + x_2 - 2x_3]^2 + x_2^2 - 3x_3^2 + 2x_2x_3. \end{aligned}$$

A questo punto nel polinomio $x_2^3 + 3x_3^2 + 2x_2x_3$ consideriamo i termini che contengono x_2 ed (aggiungendo e sottraendo opportuni monomi) facciamo in modo da ottenere il multiplo del quadrato di un polinomio di primo grado:

$$x_2^2 + 2x_2x_3 = (x_2^2 + 2x_2x_3 + x_3^2) - x_3^2 = [x_2 + x_3]^2 - x_3^2.$$

In definitiva, risulta:

$$\begin{aligned}\varphi_A(\mathbf{x}) &= 2[x_1+x_2-2x_3]^2+[x_2+x_3]^2-x_3^2-3x_3^2 = \\ &= 2[x_1+x_2-2x_3]^2+[x_2+x_3]^2-4x_3^2 = 2y_1^2+y_2^2-4y_3^2,\end{aligned}$$

dove si é posto $\quad y_1 = x_1+x_2-2x_3, \quad y_2 = x_2+x_3, \quad y_3 = x_3.$
Pertanto φ_A é stata ridotta in forma canonica; le matrici D e C tali che

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

soddisfano la relazione $C^T DC = A$ e quindi A é congruente a D.

■

Osservazione 3.1 - Il procedimento descritto sopra entra in crisi se nel polinomio $\varphi_A(\mathbf{x})$ mancano i termini quadratici, cioé se gli elementi diagonali di A sono tutti nulli. L'impasse viene superata scegliendo un elemento non nullo $a_{h,k} \neq 0$ ed operando le sostituzioni

$$x_h = y_h + y_k, \qquad x_k = y_h - y_k, \qquad x_i = y_i \quad \text{per ogni } i \text{ diverso da } h \text{ e da } k.$$

Infatti, con queste sostituzioni l'addendo $2a_{h,k}x_hx_k$ diventa $2a_{h,k}y_h^2 - 2a_{h,k}y_k^2$ e il procedimento puó partire.
Ad esempio se

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

la forma quadratica φ_A assume la forma $\quad \varphi_A(\mathbf{x}) = 4x_1x_2 + 6x_1x_3 + 2x_2x_3. \quad$ Posto

(*) $$x_1 = y_1 + y_2, \qquad x_2 = y_1 - y_2, \qquad x_3 = y_3,$$

tale forma quadratica diventa

$$\begin{aligned}\varphi_A(\mathbf{x}) &= 4(y_1+y_2)(y_1-y_2)+6(y_1+y_2)y_3+2(y_1-y_2)y_3 = \\ &= 4y_1^2-4y_2^2+6y_1y_3+6y_2y_3+2y_1y_3-2y_2y_3 = 4y_1^2-4y_2^2+8y_1y_3+4y_2y_3 = \\ &= (4y_1^2+8y_1y_3)+(-4y_2^2+4y_2y_3) = 4(y_1+y_3)^2-4y_3^3-4y_2^2+4y_2y_3 = \\ &= 4(y_1+y_3)^2-(2y_2-y_3)^2-3y_3^2.\end{aligned}$$

Da (*) si deduce che $\quad y_1 = (x_1+x_2)/2, \quad y_2 = (x_1-x_2)/2, \quad y_3 = x_3, \quad$ e quindi

$$\begin{aligned}\varphi_A(\mathbf{x}) &= 4\,[\,(x_1+x_2)/2 - x_3\,]^2 - (x_1-x_2-x_3)^2 - 3x_3^2 = \\ &= (x_1+x_2-2x_3)^2-(x_1-x_2-x_3)^2-3x_3^2 = u_1^2-u_2^2-3u_3^2,\end{aligned}$$

dove si é posto $\quad u_1 = x_1+x_2-2x_3, \quad u_2 = x_1-x_2-x_3, \quad u_3 = x_3.$
La forma quadratica φ_A é quindi ridotta in forma canonica; posto

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

si ha che $A = C^T DC$ e quindi A é congruente alla matrice diagonale D.

■

Metodo alla Gauss Jordan. - Il metodo di Lagrange é concettualmente semplice, ma alquanto macchinoso da un punto di vista pratico; il metodo degli autovalori é semplice e profondo da un punto di vista concettuale, ma puó risultare di difficile applicazione se non riusciamo a calcolare materialmente le radici del polinomio caratteristico di A.

Il metodo che segue riesce a trovare una matrice diagonale congruente ad A in maniera semplice ed efficiente utilizzando un approccio simile all'algoritmo di Gauss Jordan che abbiamo utilizzato per calcolare il rango di una matrice, il determinante di una matrice quadrata, ecc.

Ricordiamo che le operazioni elementari sulle righe di una matrice A, (cioé scambiare di posto due righe, sostituire una riga con un suo multiplo, sostituire una riga con la stessa riga piú il multiplo di un'altra riga), equivalgono a moltiplicare a sinistra A per la matrice invertibile B ottenuta effettuando le suddette operazioni sulle righe della matrice identica.

Ne segue che effettuare le stesse operazioni sulle colonne di A, (cioé sulle righe di A^T), equivale a moltiplicare a destra A per la matrice ottenuta effettuando le suddette operazioni sulle colonne della matrice identica, cioé B^T.

Di conseguenza, se in una matrice quadrata simmetrica A, si effettua una delle seguenti operazioni:

(1) si scambia la riga i-sima con la riga j-sima e subito dopo si scambia la (attuale) colonna i-sima con la colonna j-sima ,
(2) si sostituisce la riga i-sima A_i con αA_i, (con $\alpha \neq 0$), e subito dopo si sostituisce la (attuale) colonna i-sima A^i con αA^i,
(3) si sostituisce la riga i-sima A_i con $A_i + \alpha A_j$, e subito dopo si sostituisce la (attuale) colonna i-sima A^i con $A^i + \alpha A^j$,

la matrice A sará stata moltiplicata a sinistra per una matrice invertibile B e a destra per la sua trasposta, e quindi é stata trasformata in una matrice congruente ad A.

Ebbene, adoperando in maniera sistematica le suddette operazioni, con un procedimento ricorsivo simile all'algoritmo di Gauss Jordan, possiamo trasformare A in una matrice diagonale congruente ad A.

Tale procedimento ricorsivo consiste nell'effettuare per $h = 1, 2, \dots, n-1$ le seguenti operazioni.

(1) Finché l'attuale elemento $a_{h,h}$ é diverso da 0, si effettua il passo di pivot parziale su $a_{h,h}$ e subito dopo si ripetono le stesse operazioni sulle colonne, cioé per ogni $j > h$ si sostituisce la riga A_j con $A_j - (a_{j,h}/a_{h,h}) \cdot A_h$, e subito dopo si sostituisce la (attuale) colonna A^j con $A^j - (a_{j,h}/a_{h,h}) \cdot A^h$;
(2) Se $h = n-1$ oppure le ultime $n-h+1$ righe e colonne sono nulle: STOP; A é stata trasformata in una matrice diagonale; in caso contrario si va al passo successivo.
(3) Se $a_{h,h} = 0$, ma esiste $j > h$ tale che $a_{j,j} \neq 0$, allora si scambia la riga A_h con A_j e subito dopo si scambia la attuale colonna h-sima A^h con A^j; con la matrice cosí modificata si torna al passo (1).
(4) Se $a_{h,h}$ e tutti i successivi elementi diagonali sono nulli, ma esiste $j > h$ tale che $a_{h,j} \neq 0$, allora si sostituisce la riga A_h con $A_h + A_j$ e subito dopo si sostituisce la attuale colonna A^h con $A^h + A^j$; con la matrice cosí modificata si torna al passo (1).
(5) Se la riga e la colonna h-sima sono nulle, tale riga e tale colonna vengono spostate all'ultimo posto; con la matrice cosí modificata si torna al passo (4).

■

Ad ogni iterazione la matrice viene moltiplicata a sinistra per una matrice invertibile e a destra per la sua trasposta, e quindi viene trasformata in una matrice congruente alla precedente, e dunque congruente alla matrice iniziale A. Alla fine della esecuzione dell'algoritmo la matrice A é stata trasformata in una matrice diagonale D congruente ad A.

Anzi, se consideriamo la matrice (A, I_n) e su tale matrice effettuiamo le operazioni descritte sopra, alla fine dell'algoritmo la matrice A sará stata trasformata in una matrice diagonale D e la matrice identica I_n sará stata trasformata nella matrice B tale che $BAB^T = D$. Posto $\mathbf{x} = B^T\mathbf{u}$, si ha

$$\varphi_A(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{u}^T BAB^T\mathbf{u} = \mathbf{u}^T D\mathbf{u} = d_1u_1^2 + d_2u_2^2 + \ldots + d_nu_n^2,$$

e quindi la forma quadratica φ_A é stata ridotta in forma canonica.

■

Esempio 3 - Consideriamo la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 0 & 2 \\ 3 & 0 & 5 & 0 \\ 0 & 2 & 0 & 6 \end{pmatrix} \qquad \text{e quindi} \qquad (\,A \quad I\,) = \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 0 & 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & 2 & 0 & 1 & 0 & 0 \\ 3 & 0 & 5 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 6 & 0 & 0 & 0 & 1 \end{array}\right);$$

effettuando un passo di pivot parziale sull'elemento $a_{1,1} = 1$ della matrice (A, I), si ottiene la matrice

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -6 & 2 & -2 & 1 & 0 & 0 \\ 0 & -6 & -4 & 0 & -3 & 0 & 1 & 0 \\ 0 & 2 & 0 & 6 & 0 & 0 & 0 & 1 \end{array}\right).$$

Sostituendo poi le attuali colonne A^2 ed A^3 rispettivamente con $A^2 - 2A^1$ ed A^3 con $A^3 - 3A^1$, si ottiene la matrice

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -6 & 2 & -2 & 1 & 0 & 0 \\ 0 & -6 & -4 & 0 & -3 & 0 & 1 & 0 \\ 0 & 2 & 0 & 6 & 0 & 0 & 0 & 1 \end{array}\right).$$

Poiché l'attuale elemento $a_{2,2}$ é nullo ma $a_{3,3} = -4$ é diverso da 0, scambiamo la seconda con la terza riga e subito dopo scambiamo la seconda con la terza colonna; otteniamo cosí le matrici

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -6 & -4 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & -6 & 2 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 6 & 0 & 0 & 0 & 1 \end{array}\right), \qquad \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & -6 & 0 & -3 & 0 & 1 & 0 \\ 0 & -6 & 0 & 2 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 6 & 0 & 0 & 0 & 1 \end{array}\right).$$

A questo punto effettuiamo un secondo passo di pivot parziale sull'attuale elemento $a_{2,2} = -4$ e poi effettuiamo le stesse operazioni sulle colonne; otteniamo le matrici

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & -6 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 9 & 2 & 5/2 & 1 & -3/2 & 0 \\ 0 & 0 & 2 & 6 & 0 & 0 & 0 & 1 \end{array}\right), \qquad \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 9 & 2 & 5/2 & 1 & -3/2 & 0 \\ 0 & 0 & 2 & 6 & 0 & 0 & 0 & 1 \end{array}\right).$$

L'attuale elemento $a_{3,3} = 9$ é diverso da 0; effettuando un passo di pivot parziale su tale elemento e ripetendo le stesse operazioni sulle relative colonne, otteniamo le matrici:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 9 & 2 & 5/2 & 1 & -3/2 & 0 \\ 0 & 0 & 0 & 50/9 & -5/9 & -2/9 & 1/3 & 1 \end{array}\right), \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 9 & 0 & 5/2 & 1 & -3/2 & 0 \\ 0 & 0 & 0 & 50/9 & -5/9 & -2/9 & 1/3 & 1 \end{array}\right).$$

A questo punto le matrici A ed I sono state trasformate nelle matrici

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 50/9 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 5/2 & 1 & -3/2 & 0 \\ -5/9 & -2/9 & 1/3 & 1 \end{pmatrix},$$

e risulta

$$BAB^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 5/2 & 1 & -3/2 & 0 \\ -5/9 & -2/9 & 1/3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 0 & 2 \\ 3 & 0 & 5 & 0 \\ 0 & 2 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 & 5/2 & -5/9 \\ 0 & 0 & 1 & -2/9 \\ 0 & 1 & -3/2 & 1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -6 & -4 & 0 \\ 0 & 9 & 0 & 2 \\ 0 & 0 & 0 & 50/9 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 & 5/2 & -5/9 \\ 0 & 0 & 1 & -2/9 \\ 0 & 1 & -3/2 & 1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 50/9 \end{pmatrix} = D.$$

Pertanto la matrice A é congruente alla matrice D; posto $\mathbf{x} = B^T\mathbf{y}$, cioé:

$$x_1 = y_1 - 3y_2 + (5/2)y_3 - (5/9)y_4, \quad x_2 = y_3 - (2/9)y_4, \quad x_3 = y_2 - (3/2)y_3 + (1/3)y_4, \quad x_4 = y_4,$$

si ottiene lo sviluppo in forma canonica di φ_A :

$$\begin{aligned} \varphi_A(\mathbf{x}) &= x_1^2 + 4x_2^2 + 5x_3^2 + 6x_4^2 + 4x_1x_2 + 6x_1x_3 + 4x_2x_4 = \\ &= [y_1 - 3y_2 + (5/2)y_3 - (5/9)y_4]^2 + 4[y_3 - (2/9)y_4]^2 + 5[y_2 - (3/2)y_3 + (1/3)y_4]^2 + 6y_4^2 + \\ &\quad + 4[y_1 - 3y_2 + (5/2)y_3 - (5/9)y_4][y_3 - (2/9)y_4] + \\ &\quad + 6[y_1 - 3y_2 + (5/2)y_3 - (5/9)y_4][y_2 - (3/2)y_3 + (1/3)y_4] + \\ &\quad + 4[y_3 - (2/9)y_4][y_2 - (3/2)y_3 + (1/3)y_4] = \\ &= \ldots\ldots\ldots\ldots = y_1^2 - 4y_2^2 + 9y_3^2 + (50/9)\, y_4^2; \end{aligned}$$

(lasciamo come esercizio il compito di sviluppare i calcoli).

■

Esempio 4 - Riduciamo in forma canonica la matrice

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

Dal momento che gli elementi diagonali di A sono tutti nulli, sostituiamo la prima riga con la somma delle prime due e subito dopo sostituiamo la prima colonna con la somma delle prime due; la matrice (A, I) diventa

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right), \qquad \left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right).$$

A questo punto si effettua il passo di pivot parziale sul primo elemento della prima colonna, e poi si ripetono le stesse operazioni sulle colonne di A; si ottengono quindi le matrici

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 1 & 0 \\ 0 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 1/2 & -1/2 & -1/2 & -1/2 & 1 \end{array}\right), \qquad \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 1/2 & -1/2 & -1/2 & -1/2 & 1 \end{array}\right).$$

A questo punto si effettua un passo di pivot parziale sul secondo elemento diagonale e poi si ripetono le stesse operazioni sulle colonne di A; si ottengono le matrici

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{array}\right), \qquad \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1/2 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{array}\right).$$

Le matrici A ed I sono state trasformate rispettivamente nelle matrici

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 1 & 1 & 0 \\ -1/2 & 1/2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Pertanto A é stata ridotta in forma canonica; posto $\mathbf{x} = B^T\mathbf{y}$, cioé

$$x_1 = y_1 - y_2/2 - y_3, \quad x_2 = y_1 + y_2/2, \quad x_3 = y_3,$$

si ha:

$$\varphi_A(\mathbf{x}) = 2x_1x_2 + 2x_2x_3 = 2x_2(x_1 + x_3) = 2(y_1 + y_2/2)(y_1 - y_2/2) = 2(y_1^2 - y_2^2/4) = 2y_1^2 - (1/2)y_2^2 = \mathbf{y}^T D\mathbf{y}.$$

■

Se non ci interessa conoscere esplicitamente la matrice B che diagonalizza A per congruenza, ma solo la matrice D congruente ad A, possiamo evitare di eseguire per intero il procedimento che abbiamo detto "alla Gauss Jordan", ma limitarci a calcolare semplicemente dei determinanti.

A tal fine, é utile premettere la seguente

Definizione 3.2. *Per ogni $k = 1, 2, \ldots, n$, definiamo* **minore principale di ordine** k *della matrice quadrata A, (in simboli $m_k(A)$), il determinante della sottomatrice $A(k)$ formata dalle prime k righe e k colonne di A.*

■

Ció posto, supponiamo di stare eseguendo la riduzione in forma canonica di A usando il metodo che abbiamo chiamato "alla Gauss Jordan", di essere partiti da $a_{1,1} \neq 0$ e di avere eseguito il passo (1) nelle prime $k-1$ iterazioni. Pertanto la matrice A é stata trasformata in una matrice congruente A' della forma $A' = \begin{pmatrix} D & O \\ O & C \end{pmatrix}$, dove D é una matrice diagonale non singolare di ordine $k-1$.

La matrice A' é stata ottenuta attraverso una sistematica sostituzione di una riga o una colonna con la somma della stessa riga (o colonna) con un multiplo opportuno di una riga o colonna precedente; queste operazioni non alterano i determinanti, sicché, detti $d_1, d_2, \ldots, d_{k-1}$ gli elementi diagonali di D, si ha:

$$\begin{aligned} m_j(A) &= \det A(j) = \det A'(j) = d_1 d_2 \ldots d_j, \qquad \text{per ogni } j = 1, 2, \ldots, k-1\ , \\ m_k(A) &= \det(A(k) = \det A'(k) = d_1 d_2 \ldots d_{k-1} \cdot a'_{k,k} = m_{k-1}(A) \cdot a'_{k,k}. \end{aligned}$$

Risulta dunque $a'_{k,k} = m_k(A)/m_{k-1}(A)$, e quindi si potrá eseguire ancora una volta il passo (1) se risulta $m_k(A) \neq 0$, mentre si dovrá eseguire uno dei passi (3) o (4) nel caso $m_k(A) = 0$.
Se ne deduce che

Proposizione 3.3 -. *Se i primi $n-1$ minori principali di A sono diversi da 0, allora la matrice A é congruente alla matrice D i cui elementi diagonali $d_1, d_2, \ldots, d_n$ sono descritti dalle relazioni*

$$(*) \qquad d_1 = m_1(A), \quad d_2 = \frac{m_2(A)}{m_1(A)}, \quad d_3 = \frac{m_3(A)}{m_2(A)}, \qquad \ldots\ldots, \qquad d_n = \frac{m_n(A)}{m_{n-1}(A)}.$$

Dim. Infatti per quanto appena visto, si ha che se i primi $n-1$ minori principali di A sono diversi da 0, allora potremo eseguire $n-1$ volte il passo (1) e trasformare A nella matrice diagonale D i cui elementi diagonali sono dati dalle relazioni (*).

■

Osservazione 3.4 - Se invece risulta $m_k(A) = 0$, allora per proseguire l'esecuzione dell'algoritmo alla Gauss Jordan, occorrerebbe passare al passo (3) e poi eventualmente al passo (4) che consistono nello scambiare la riga e la colonna k-sima con una successiva o nella sostituzione della riga e colonna k-sima con la somma di tale riga e colonna con una riga e colonna successiva nella speranza che il nuovo minore k-simo sia diverso da 0 e si possa proseguire con l'esecuzione dell'algoritmo.

Questo peró ci espone al rischio di effettuare innumerevoli tentativi a vuoto se l'operazione tentata produce un determinante nullo. Il metodo é quindi poco pratico tenuto conto del fatto che il calcolo di un determinante é oneroso se l'ordine della matrice non é piccolo. Si fa prima ad eseguire l'algoritmo alla Gauss Jordan che permette di trovare la matrice diagonale congruente ad A in maniera mirata, poiché individua subito le righe e colonne da scambiare o le righe e colonne da sommare.

■

Osservazione 3.5 - Scambiando di posto due righe e le corrispondenti colonne di una matrice quadrata simmetrica A si ottiene una matrice congruente ad A. Perció, scambiando opportunamente le righe e le colonne di una matrice diagonale congruente ad A, possiamo sempre fare in modo da disporre in D dapprima tutti gli eventuali elementi diagonali positivi, poi gli eventuali elementi diagonali negativi, infine gli eventuali elementi diagonali nulli.

■

Osservazione 3.6 - Se D é una matrice diagonale congruente ad A e $d_i \neq 0$ é l'i-simo elemento diagonale di D, allora, dividendo per $\sqrt{|d_i|}$ la riga e la colonna i-sima di D, si ottiene ancora una matrice diagonale congruente ad A il cui elemento diagonale i-simo é uguale ad 1 se $d_i > 0$ a -1 se $d_i < 0$.

Ripetendo questo ragionamento per tutti gli elementi diagonali non nulli di D, si ha che A é congruente ad una matrice diagonale i cui elementi diagonali sono tutti uguali a $1,\ -1$ o 0.

Se poi le righe e le colonne vengono opportunamente scambiate di posto, si puó fare in modo che A sia congruente ad una matrice diagonale i cui primi s elementi diagonali sono uguali ad 1, i successivi t elementi diagonali sono uguali a -1, gli ultimi $n-s-t$ elementi diagonali sono nulli.

Pertanto una forma quadratica puó essere ridotta in forma canonica nella forma

$$\varphi_A(\mathbf{x}) = y_1^2 + y_2^2 + \ldots + y_s^2 - y_{s+1}^2 - y_{s+2}^2 - y_{s+t}^2, \tag{1}$$

La (1) dicesi **forma normale di Sylvester** della forma quadratica associata ad A.

■

Osservazione 3.7 - Se D é una matrice diagonale congruente alla matrice simmetrica A, e tutti gli elementi diagonali di D sono maggiori di 0, allora per ogni $\mathbf{x} \neq \mathbf{0}$ risulta

$$\varphi_A(\mathbf{x}) = d_1 u_1^2 + d_2 u_2^2 + \ldots + d_n u_n^2 > 0\,, \tag{2}$$

dal momento che tutti gli addendi sono maggiori o uguali a 0 ed uno almeno é diverso da 0.

Ne segue che ogni altra matrice diagonale D' congruente ad A avrá gli elementi diagonali maggiori di 0. Infatti se esistesse un elemento diagonale di D', ad esempio d'_1, che é minore o uguale a 0, allora, detta C la matrice quadrata invertibile tale che $C^T AC = D'$ e detta C^1 la prima colonna si C, si avrebbe

$$\varphi_A(C^1) = (C^1)^T AC^1 = (CE^1)^T ACE^1 = (E^1)^T C^T ACE^1 = (E^1)^T D'E^1 = d'_1 \leq 0,$$

contraddicendo la (2).

In realtá sussiste un risultato molto piú generale, come vedremo nel prossimo paragrafo.

■

4 . - Segnatura di una matrice quadrata simmetrica

Sappiamo che se A é una matrice quadrata simmetrica di ordine n, allora il polinomio caratteristico di A ha solo radici reali, che la molteplicitá algebrica di ogni autovalore di A coincide con la sua molteplicitá geometrica e quindi che la somma delle molteplicitá degli autovalori di A é uguale ad n,

A questo punto é possibile introdurre il concetto di **segnatura** di una matrice quadrata simmetrica.

Definizione 4.1. *Se A é una matrice quadrata simmetrica di ordine n, dicesi:*

- **indice di positivitá** *di A, (in simboli $i_+(A)$), la somma delle molteplicitá degli autovalori strettamente positivi di A,*
- **indice di negativitá** *di A, (in simboli $i_-(A)$), la somma delle molteplicitá degli autovalori strettamente negativi di A,*
- **indice di nullitá** *di A, (in simboli $i_0(A)$), la molteplicitá dell'autovalore nullo di A,*

con la convenzione che $i_+(A) = 0$ se A non ha autovalori strettamente positivi, che $i_-(A) = 0$ se A non ha autovalori strettamente negativi, che $i_0(A) = 0$ se 0 non é un autovalore di A.

La terna ordinata $\Sigma(A) = (i_+(A), i_-(A), i_0(A))$ dicesi **segnatura** *di A.*

■

Osservazione 4.2 - É evidente che risulta $i_+(A) + i_-(A) + i_0(A) = n$; ne segue che

$$\Sigma(A) = (i_+(A), i_-(A), n - i_+(A) - i_-(A)).$$

Per questo motivo, alcuni autori definiscono **segnatura di** A semplicemente la coppia ordinata

$$(i_+(A), i_-(A)).$$

D'altra parte, é evidente che $i_0(A)$ é la dimensione del nucleo della trasformazione lineare associata ad A, cioé dell'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$; si ha quindi che $i_0(A) = n - car(A)$. e dunque che $i_+(A) + i_-(A) = n - i_0(A) = car(A)$.

■

Osservazione 4.3 - Se ciascun autovalore viene ripetuto un numero di volte uguale alla sua molteplicitá (algebrica o geometrica), allora ovviamente $i_+(A)$, $i_-(A)$ ed $i_0(A)$ rappresentano rispettivamente il numero di autovalori di A che sono maggiori di 0, minori di 0 o uguali a 0.

Per trovare la segnatura di A, basta quindi risolvere l'equazione caratteristica di A e contare il numero di radici di tale equazione che sono maggiori di 0, minori di 0 ed uguali a 0.

Ad esempio, nel caso dell'esempio 1 del precedente paragrafo, gli autovalori di A sono 3 (doppio) e 9 (semplice) e quindi la segnatura di A é $\Sigma(A) = (3, 0, 0)$.

In realtá non é nemmeno necessario trovare esplicitamente le radici dell'equazione caratteristica di A, (cosa tutt'altro che facile), ma basta analizzare i segni dei coefficienti del polinomio caratteristico

$$p_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots\ldots + a_1\lambda + a_0.$$

■

Osservazione 4.4 - In primo luogo osserviamo che 0 é radice del polinomio

$$p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots\ldots + a_1\lambda + a_0$$

se e solo se risulta $a_0 = 0$ ed anzi che 0 é radice di molteplicitá h di p se e solo se a_h é l'ultimo coefficiente diverso da 0 del polinomio p.

In tal caso le radici non nulle di p coincidono con le radici di

$$q(\lambda) = a_n\lambda^{n-h} + a_{n-1}\lambda^{n-h-1} + a_{n-2}\lambda^{n-h-2} + \ldots\ldots + a_{h+1}\lambda + a_h.$$

■

Per quanto riguarda le radici positive di un polinomio, si ha che sussiste il seguente Teorema di cui si omette la dimostrazione.

Teorema 4.5 - (Regola dei segni di Cartesio).
Se p é un polinomio di grado n a coefficienti reali avente solo radici reali, allora il numero delle radici positive di p é uguale al numero delle **variazioni** *di segno nella sequenza dei coefficienti non nulli di p.*

■

Osservazione 4.6 - Per quanto riguarda le radici negative di un polinomio,

$$p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots\ldots + a_1\lambda + a_0$$

possiamo osservare che α é una radice negativa di p se e solo se $-\alpha$ é radice positiva del polinomio

$$\begin{aligned}\hat{p}(\lambda) = p(-\lambda) = a_n(-\lambda)^n + a_{n-1}(-\lambda)^{n-1} + a_{n-2}(-\lambda)^{n-2} + \ldots\ldots + a_1(-\lambda) + a_0 = \\ = (-1)^n a_n\lambda^n + (-1)^{n-1}a_{n-1}\lambda^{n-1} + (-1)^{n-2}a_{n-2}\lambda^{n-2} + \ldots\ldots - a_1\lambda + a_0.\end{aligned}$$

Pertanto il numero di radici negative di un polinomio di grado n avente solo radici reali é uguale al numero di variazioni nella sequenza dei segni dei coefficienti non nulli del polinomio $\hat{p}$, ottenuto da p conservando i segni dei coefficienti delle potenze pari e cambiando i segni dei coefficienti delle potenze dispari.

■

Osservazione 4.7 - In particolare, le radici di p sono tutte positive se e solo se la sequenza dei segni dei coefficienti di p presenta n variazioni e quindi se e solo se i coefficienti di p sono tutti diversi da 0 e di segno alternato.

Invece le radici di p sono tutte negative se e solo se le radici di $\hat{p}$ sono tutte positive e quindi se e solo se i coefficienti di $\hat{p}$ sono tutti diversi da 0 e di segno alterno; questo accade se e solo se i coefficienti di p sono tutti diversi da 0 e dello stesso segno.

■

Osservazione 4.8 - Pertanto se $p_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ é il polinomio caratteristico di una matrice quadrata simmetrica A di ordine n e nella sequenza dei segni dei coefficienti non nulli del polinomio p_A ci sono s variazioni, si ha quanto segue:

- se $a_0 \neq 0$, allora p_A ha s radici positive ed $n-s$ radici negative, e quindi la segnatura di A é $\Sigma(A) = (s, n-s, 0)$;
- se gli ultimi h coefficienti di p_A sono nulli, allora p_A ha s radici positive, h radici nulle ed $n-s-h$ radici negative, e quindi la segnatura di A é $\Sigma(A) = (s, n-s-h, h)$.

In particolare si ha che:

a) $\Sigma(A) = (n, 0, 0)$ se e solo se p_A ha n radici positive e quindi se e solo se i coefficienti di p_A sono diversi da 0 e di segno alterno;
b) $\Sigma(A) = (0, n, 0)$ se e solo se p_A ha n radici negative e quindi se e solo se i coefficienti di p_A sono diversi da 0 e dello stesso segno;
c) $\Sigma(A) = (n-h, 0, h)$ se e solo se i primi $n-h+1$ coefficienti di p_A sono non nulli e di segno alterno e gli ultimi h coefficienti di p_A sono nulli;
d) $\Sigma(A) = (0, n-h, h)$ se e solo se i primi $n-h+1$ coefficienti di p_A sono non nulli e dello stesso segno e gli ultimi h coefficienti di p_A sono nulli.

■

Esempio 5 - Ad esempio se il polinomio caratteristico di una matrice A é $p_A(\lambda) = \lambda^4 - 5\lambda^3 + 4\lambda^2 + 7\lambda - 3$, la sequenza dei segni dei coefficienti di p_A, (cioé $(+, -, +, +, -)$), presenta tre variazioni e quindi p_A ha tre radici positive.

Dal momento che risulta $p_A(0) = -3$, si ha che 0 non é radice di p_A e quindi l'ultima radice di p_A é negativa. D'altra parte, questa conclusione é confermata dal fatto che la sequenza dei segni dei coefficienti del polinomio $\hat{p}_A(\lambda) = p_A(-\lambda) = \lambda^4 + 5\lambda^3 + 4\lambda^2 - 7\lambda - 3$ é $(+, +, +, -, -)$ e quindi presenta una sola variazione; pertanto $\hat{p}_A$ ha una sola radice positiva e quindi p_A ha una sola radice negativa.

La segnatura di A é quindi $\Sigma(A) = (3, 1, 0)$.

■

Esempio 6 - Supponiamo che il polinomio caratteristico di una matrice A sia $p_A(\lambda) = \lambda^7 + 2\lambda^6 - 5\lambda^5 + 2\lambda^3$. Evidentemente $\lambda = 0$ é una radice di molteplicitá 3 di p_A; la sequenza dei segni dei coefficienti non nulli di p_A é $(+, +, -, +)$ e quindi presenta due variazioni. Pertanto p_A ha tre radici nulle, due radici positive e le restanti due radici sono negative.

D'altra parte,a conferma del fatto che p_A ha due radici negative, possiamo osservare che la sequenza dei segni dei coefficienti non nulli del polinomio $\hat{p}_A(\lambda) = p_A(-\lambda) = -\lambda^7 + 2\lambda^6 + 5\lambda^5 - 2\lambda^3$ presenta due variazioni.

La segnatura di A é dunque $\Sigma(A) = (2, 2, 3)$.

■

Esempio 7 - Sia $p_A(\lambda) = -\lambda^5 + 8\lambda^4 - 23\lambda^3 + 30\lambda^2 - 18\lambda + 4$ il polinomio caratteristico di A; i coefficienti di p_A sono diversi da 0 e di segno alternato e quindi la sequenza dei segni dei coefficienti di p_A presenta 5 variazioni e quindi p_A ha 5 radici positive. La segnatura di A é $\Sigma(A) = (5, 0, 0)$.

■

Esempio 8 - Sia $p_A(\lambda) = \lambda^4 + 6\lambda^3 + 10\lambda^2 + 4\lambda$ il polinomio caratteristico di A. La sequenza dei segni dei coefficienti non nulli di A non presenta nessuna variazione e quindi p_A non ha alcuna radice positva. D'altra parte é evidente che $\lambda = 0$ é una radice di molteplicitá 1 di p_A e quindi le altre tre radici di p_A sono negative.

Quest'ultima conclusione é confermata dal fatto che la sequenza dei segni dei coefficienti del polinomio $\hat{p}_A(\lambda) = p_A(-\lambda) = \lambda^4 - 6\lambda^3 + 10\lambda^2 - 4\lambda$ presenta tre variazioni.

La segnatura di A é quindi $\Sigma(A) = (0, 3, 1)$.

■

Il principale risultato sulla segnatura di una matrice quadrata simmetrica é contenuto nel seguente teorema che chiarisce il significato profondo degli indici di positivitá e di negativitá.

Teorema 4.9. -. *Se A é una matrice quadrata simmetrica di ordine n, allora $\mathbf{R}^n$ puó essere decomposto in tre sottospazi V_+, V_- e V_0 a due a due ortogonali tali che:*

(1) $i_+(A) = \dim(V_+)$, $\quad i_-(A) = \dim(V_-)$, $\quad i_0(A) = \dim(V_0)$,

(2) *per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha $\mathbf{x} = \mathbf{x}_+ + \mathbf{x}_- + \mathbf{x}_0$, dove $\mathbf{x}_+, \mathbf{x}_-$ ed $\mathbf{x}_0$ sono le proiezioni ortogonali di $\mathbf{x}$ rispettivamente su V_+, V_-, e V_0;*

(3) *per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha $\varphi_A(\mathbf{x}) = \varphi_A(\mathbf{x}_+) + \varphi_A(\mathbf{x}_-)$, con $\varphi_A(\mathbf{x}_+) \geq 0$ e $\varphi_A(\mathbf{x}_-) \leq 0$;*

(4) $\varphi_A(\mathbf{x}) > 0$ *per ogni $\mathbf{x} \in V_+ - \{\mathbf{0}\}$,*

(5) $\varphi_A(\mathbf{x}) < 0$ *per ogni $\mathbf{x} \in V_- - \{\mathbf{0}\}$,*

(6) $\varphi_A(\mathbf{x}) = 0$ *per ogni $\mathbf{x} \in V_0$,*

(7) *se H é un sottospazio di $\mathbf{R}^n$ e $\dim(H) > i_+(A)$, allora esiste $\mathbf{x} \in H - \{\mathbf{0}\}$ tale che $\varphi_A(\mathbf{x}) \leq 0$,*

(8) *se H é un sottospazio di $\mathbf{R}^n$ e $\dim(H) > i_-(A)$, allora esiste $\mathbf{x} \in H - \{\mathbf{0}\}$ tale che $\varphi_A(\mathbf{x}) \geq 0$,*

Pertanto, se $i_+(A) > 0$, allora $i_+(A)$ rappresenta la massima dimensione dei sottospazi di $\mathbf{R}^n$ su cui φ_A é definita positiva; analogamente, se $i_-(A) > 0$, allora $i_-(A)$ rappresenta la massima dimensione dei sottospazi di $\mathbf{R}^n$ su cui φ_A é definita negativa.

Dim. Sia $r = car(A)$, $s = i_+(A)$, $r - s = i_-(A)$, $n - r = i_0(A)$; sappiamo che esiste una base ortonormale $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ di autovettori di A corrispondenti (nell'ordine) agli autovalori $\lambda_1, \lambda_2, \ldots, \lambda_n$, dove ciascun autovalore é ripetuto un numero di volte uguale alla sua molteplicitá.

Cambiando eventualmente l'ordine degli autovalori e dei corrispondenti autovettori, possiamo supporre che gli eventuali autovalori positivi precedano gli eventuali autovalori negativi e questi ultimi precedano l'eventuale autovalore nullo. Sia dunque V_+ lo spazio generato dai primi $s = i_+(A)$ autovettori, V_- lo spazio generato dai successivi $r - s = i_-(A)$ autovettori e sia infine V_0 lo spazio generato dagli autovettori corrispondenti all'eventuale autovalore 0, con la ovvia convenzione che

$$V_+ = \{\mathbf{0}\} \quad \text{se } \ i_+(A) = 0, \qquad\qquad V_- = \{\mathbf{0}\} \quad \text{se } \ i_-(A) = 0, \qquad\qquad V_0 = \{\mathbf{0}\} \quad \text{se } \ i_0(A) = 0.$$

Evidentemente i sottospazi V_+, V_- e V_0 soddisfano la (1). Inoltre essi sono a due a due ortogonali. Sappiamo infatti che gli autovettori di A sono a due a due ortogonali e quindi ogni vettore di V_+ é ortogonale a V_- e V_0; analogamente ogni vettore di V_- é ortogonale a V_+ e V_0 ed ogni vettore di V_0 é ortogonale a V_+ e V_-.
D'altra parte, per la Proposizione 4.3 del Cap. VI, per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha che le proiezioni ortogonali di $\mathbf{x}$ su V_+, V_- e V_0 sono i vettori

$$\mathbf{x}_+ = \sum_{i=1}^{s} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i \in V_+, \qquad \mathbf{x}_- = \sum_{i=s+1}^{r} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i \in V_-, \qquad \mathbf{x}_0 = \sum_{i=r+1}^{n} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i \in V_0,$$

con la ovvia convenzione che $\mathbf{x}_+ = \mathbf{0}$ se $i_+(A) = 0$, che $\mathbf{x}_- = \mathbf{0}$ se $i_-(A) = 0$ e che $\mathbf{x}_0 = \mathbf{0}$ se $i_0(A) = 0$.
Inoltre dalla stessa Proposizione 4.3 del Cap. VI segue che

$$\mathbf{x} = \sum_{i=1}^{n} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i = \Big(\sum_{i=1}^{s} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i\Big) + \Big(\sum_{i=s+1}^{r} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i\Big) + \Big(\sum_{i=r+1}^{n} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i\Big) = \mathbf{x}_+ + \mathbf{x}_- + \mathbf{x}_0.$$

Inoltre, per ogni $\mathbf{x} \in \mathbf{R}^n$ si ha

$$\varphi_A(\mathbf{x}) = \Big(A\mathbf{x}\Big|\mathbf{x}\Big) = \Big(A\Big(\sum_{i=1}^{n} (\mathbf{x}|\mathbf{u}_i) \cdot \mathbf{u}_i\Big)\Big|\mathbf{x}\Big) = \Big(\sum_{i=1}^{n} (\mathbf{x}|\mathbf{u}_i) \cdot A\mathbf{u}_i\Big|\mathbf{x}\Big) = \sum_{i=1}^{n} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2.$$

Ne segue che

$$\varphi_A(\mathbf{x}_+) = \sum_{i=1}^{n} \lambda_i (\mathbf{x}_+|\mathbf{u}_i)^2 = \sum_{i=1}^{s} \lambda_i (\mathbf{x}_+|\mathbf{u}_i)^2, \qquad \varphi_A(\mathbf{x}_-) = \sum_{i=1}^{n} \lambda_i (\mathbf{x}_-|\mathbf{u}_i)^2 = \sum_{i=s+1}^{r} \lambda_i (\mathbf{x}_-|\mathbf{u}_i)^2$$

dal momento che $(\mathbf{x}_+|\mathbf{u}_i) = 0$ per ogni $i > s$ e $(\mathbf{x}_-|\mathbf{u}_i) = 0$ per ogni $i \leq s$ e per ogni $i > r$.
D'altra parte risulta

$$(\mathbf{x}|\mathbf{u}_j) = (\mathbf{x}_+|\mathbf{u}_j) + (\mathbf{x}_-|\mathbf{u}_j) + (\mathbf{x}_0|\mathbf{u}_j) = \begin{cases} (\mathbf{x}_+|\mathbf{u}_j) & \text{se } \ j \leq s, \\ (\mathbf{x}_-|\mathbf{u}_j) & \text{se } \ s < j \leq r, \\ (\mathbf{x}_0|\mathbf{u}_j) & \text{se } \ r < j \leq n, \end{cases}$$

e quindi si ha

$$\varphi_A(\mathbf{x}_+) = \sum_{i=1}^{s} \lambda_i (\mathbf{x}_+|\mathbf{u}_i)^2 = \sum_{i=1}^{s} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 \geq 0, \qquad \varphi_A(\mathbf{x}_-) = \sum_{i=s+1}^{r} \lambda_i (\mathbf{x}_-|\mathbf{u}_i)^2 = \sum_{i=s+1}^{r} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 \leq 0.$$

Ne segue (essendo $\lambda_i = 0$ per ogni $i > r$) che

$$\varphi_A(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 = \sum_{i=1}^{s} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 + \sum_{i=s+1}^{r} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 + \sum_{i=r+1}^{n} \lambda_i (\mathbf{x}|\mathbf{u}_i)^2 = \varphi_A(\mathbf{x}_+) + \varphi_A(\mathbf{x}-).$$

Se ora é $i_+(A) > 0$ ed $\mathbf{x} \in V_+ - \{\mathbf{0}\}$, allora risulta $\mathbf{x} = \mathbf{x}_+$ e quindi

$$\varphi_A(\mathbf{x}) = \varphi_A(\mathbf{x}_+) = \sum_{i=1}^{s} \lambda_i (x|u_i)^2 > 0,$$

dal momento che tutti gli addendi sono ≥ 0 ed uno almeno di essi é diverso da 0.
Analogamente, se $i_-(A) > 0$ ed $\mathbf{x} \in V_- - \{\mathbf{0}\}$, allora risulta $\mathbf{x} = \mathbf{x}_-$, e quindi

$$\varphi_A(\mathbf{x}) = \varphi_A(\mathbf{x}_-) = \sum_{i=s+1}^{r} \lambda_i (x|u_i)^2 < 0,$$

dal momento che tutti gli addendi sono ≤ 0 ed uno almeno di essi é diverso da 0.
Infine per ogni $\mathbf{x} \in V_0$ si ha che $A\mathbf{x} = \mathbf{0}$ e quindi $\varphi_A(\mathbf{x}) = (A\mathbf{x}|\mathbf{x}) = (\mathbf{0}|\mathbf{x}) = 0$.

Proviamo infine la (7); la (8) si prova in maniera analoga.
Sia H un sottospazio di $\mathbf{R}^n$ tale che $\dim(H) = h > s$; se proviamo che esiste $\mathbf{x} \in H, \mathbf{x} \neq \mathbf{0}$ tale che $\mathbf{x}_+ = \mathbf{0}$, ne seguirá che $\varphi_A(\mathbf{x}) = \varphi_A(\mathbf{x}_-) \leq 0$ e la tesi sará provata..
Ebbene, se $i_+(A) = 0$, allora qualunque vettore $\mathbf{x} \in H - \{\mathbf{0}\}$ soddisfa la tesi. Se invece é $s = i_+(A) > 0$, allora occorre cercare $\mathbf{x} \in H - \{\mathbf{0}\}$ tale che

(1) $$(\mathbf{x}|\mathbf{u}_i) = 0 \qquad \text{per ogni } i = 1, 2, \ldots, s.$$

Detta $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_h\}$ una base di H, occorre cercare h numeri reali non tutti nulli $\alpha_1, \alpha_2, \ldots, \alpha_h$ tali che $\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_h\mathbf{v}_h$ soddisfi la (1). Occorre dunque cercare una soluzione non nulla del sistema lineare omogeneo di s equazioni in h incognite

$$\begin{cases} \alpha_1(\mathbf{v}_1|\mathbf{u}_1) + \alpha_2(\mathbf{v}_2|\mathbf{u}_1) + \ldots + \alpha_h(\mathbf{v}_h|\mathbf{u}_1) & = 0 \\ \alpha_1(\mathbf{v}_1|\mathbf{u}_2) + \alpha_2(\mathbf{v}_2|\mathbf{u}_2) + \ldots + \alpha_h(\mathbf{v}_h|\mathbf{u}_2) & = 0 \\ \cdots\cdots\cdots \\ \alpha_1(\mathbf{v}_1|\mathbf{u}_k) + \alpha_2(\mathbf{v}_2|\mathbf{u}_k) + \ldots + \alpha_h(\mathbf{v}_h|\mathbf{u}_k) & = 0\,. \end{cases}$$

Ebbene questo sistema ha un numero s di equazioni minore del numero h delle incognite e quindi ammette certamente una soluzione non nulla $(\alpha_1, \alpha_2, \ldots, \alpha_h)$; posto $\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_h\mathbf{v}_h$, si ha $\mathbf{x}_+ = \mathbf{0}$ e quindi $\varphi_A(\mathbf{x}) \leq 0$.

■

Dal precedente Teorema discende il seguente fondamentale

Teorema 4.10. - (Legge d'inerzia di Sylvester). - *Due matrici simmetriche sono congruenti se e solo se hanno la stessa segnatura.*

Dim. Se le matrici quadrate simmetriche A ed A' hanno la stessa segnatura, $(s, t, n-s-t)$, allora le forme quadratiche associate ad A ed A' hanno la stessa forma normale di Sylvester,

$$y_1^2 + y_2^2 + \ldots + y_s^2 - y_{s+1}^2 - y_{s+2}^2 - \ldots - y_{s+t}^2;$$

ne segue che A ed A' sono congruenti, dal momento che esse sono entrambe congruenti alla stessa matrice diagonale D, avente s elementi diagonali uguali ad 1, seguiti da t elementi diagonali uguali ad -1 ed infine $n-s-t$ elementi diagonali uguali a 0.

Viceversa sia A' congruente ad A e sia B una matrice quadrata invertibile tale che $B^TAB = A'$. Vogliamo dimostrare che A ed A' hanno la stessa segnatura, cioé che

$$i_+(A') = i_+(A), \qquad i_-(A') = i_-(A), \qquad i_0(A') = i_0(A).$$

A tal fine siano L_B ed $L_{B^{-1}}$ le trasformazioni lineari associate alle matrici B e B^{-1}, cioé le funzioni tali che $L_B(\mathbf{x}) = B\mathbf{x}$ ed $L_{B^{-1}}(\mathbf{x}) = B^{-1}\mathbf{x}$ per ogni $\mathbf{x} \in \mathbf{R}^n$. Essendo B invertibile tali trasformazioni lineari sono iniettive e quindi trasformano sottospazi di $\mathbf{R}^n$ in sottospazi aventi la stessa dimensione.

D'altra parte, sappiamo che esistono tre sottospazi a due a due ortogonali V_+, V_- e V_0 soddisfacenti le condizioni (1)-(8) del Teorema 4.9 ed esistono altri tre sottospazi V'_+, V'_- e V'_0 soddisfacenti condizioni analoghe alle (1)-(8) per la matrice A'. Si ha allora

$$x \in V'_0 \iff A'\mathbf{x} = \mathbf{0} \iff B^T AB\mathbf{x} = \mathbf{0} \iff AB\mathbf{x} = \mathbf{0} \iff B\mathbf{x} \in V^0 \iff L_B(\mathbf{x}) \in V^0$$

e quindi $V_0 = L_B(V'_0)$; ne segue che V_0 e V'_0 hanno la stessa dimensione, cioé $i_0(A') = i_0(A)$.

D'altra parte, per ogni $\mathbf{x} \in V'_+ - \{\mathbf{0}\}$ si ha $0 < \varphi_{A'}(\mathbf{x}) = \mathbf{x}^T A'\mathbf{x} = \mathbf{x}^T B^T AB\mathbf{x} = \varphi_A(B\mathbf{x})$; pertanto φ_A é definita positiva sul sottospazio $L_B(V'_+)$, e quindi per la (7) del Teorema 4.9 si ha

$$(2) \qquad i_+(A') = \dim(V'_+) = \dim(L_B(V'_+)) \le \dim(V_+) = i_+(A).$$

Viceversa, per ogni $\mathbf{x} \in V_+ - \{\mathbf{0}\}$ si ha $0 < \varphi_A(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \left((B^T)^{-1} A' B^{-1}\right)\mathbf{x} = \varphi_{A'}(B^{-1}\mathbf{x})$; ne segue che $\varphi_{A'}$ é definita positiva sul sottospazio $L_{B^{-1}}(V_+)$, e quindi (ancora per la (7) del Teorema 4.9) si ha

$$(3) \qquad i_+(A) = \dim(V_+) = \dim(L_{B^{-1}}(V_+)) \le \dim(V'_+) = i_+(A').$$

Da (2) e (3) segue che $i_+(A') = i_+(A)$. In maniera analoga si vede che $i_-(A') = i_-(A)$ e dunque A ed A' hanno la stessa segnatura.

■

Osservazione 4.11 - Il fatto che due matrici congruenti hanno la stessa segnatura non é affatto sorprendente. Infatti esse rappresentano la stessa funzione di $\mathbf{R}^n$ in $\mathbf{R}$ rispetto a due diverse basi ma i sottospazi di $\mathbf{R}^n$ su cui tale funzione assume valori maggiori di 0 o minori di 0 non dipendono dalla base con cui tale funzione viene rappresentata.

■

Dal precedente Teorema 4.10 discende immediatamente il seguente

Corollario 4.12. *Se A é una matrice quadrata simmetrica congruente ad una matrice diagonale D, allora l'indice di positivitá $i_+(A)$, l'indice di negativitá $i_-(A)$ e l'indice di nullitá $i_0(A)$ di A coincidono rispettivamente con il numero di elementi positivi, negativi e nulli della diagonale di D.*

Dim. Infatti risulta $\Sigma(A) = \Sigma(D)$ e quindi $i_+(A) = i_+(D)$, $i_-(A) = i_-(D)$, $i_o(A) = i_0(D)$. La tesi discende allora dal fatto che gli autovalori di una matrice diagonale sono proprio i suoi elementi diagonali.

■

Corollario 4.13 -. *Se D e D' sono due matrici diagonali congruenti ad A, allora D e D' hanno lo stesso numero di elementi diagonali positivi, lo stesso numero di elementi diagonali negativi e lo stesso numero di elementi diagonali nulli.*

Dim. Infatti, per il precedente corollario, il numero degli elementi positivi, negativi e nulli della diagonale di D o di D' coincidono rispettivamente con $i_+(A)$, $i_-(A)$ e $i_0(A)$.

■

Ad esempio, nel caso dell'esempio 2 del precedente paragrafo, si ha che A é congruente alla matrice diagonale i cui elementi diagonali sono $2, 1, -4$, e quindi la segnatura di A é $\Sigma(A) = (2, 1, 0)$.

Nel caso dell'esempio 3 si é visto che A é congruente alla matrice i cui elementi diagonali sono $1, -4, 9, 50/9$, e quindi la segnatura di A é $\Sigma(A) = (3, 1, 0)$.

Infine, la segnatura della matrice A dell'esempio 4 é $\Sigma(A) = (1,1,1)$, poiché A é congruente alla matrice diagonale D i cui elementi diagonali sono 2, 0 e $-1/2$.

■

Osservazione 4.14 - Si noti che la sequenza dei coefficienti della forma normale di Sylvester puó essere anche rappresentata come

$$\underbrace{+\ +\ \ldots +}_{i_+(A)\ \text{volte}}\ \underbrace{-\ -\ \ldots -}_{i_-(A)\ \text{volte}}$$

ed é proprio da questo che deriva il nome di **segnatura**.

■

5. - Studio del segno di una matrice quadrata simmetrica

Siamo ora in grado di studiare facilmente il segno di una forma quadratica. Infatti dal Teorema 4.6 discende immediatamente il seguente

Corollario 5.1.

a) *A é definita positiva se e solo se* $\Sigma(A) = (n, 0, 0)$,
b) *A é definita negativa se e solo se* $\Sigma(A) = (0, n, 0)$,
c) *A é semidefinita positiva se e solo se risulta* $car(A) < n$ *e* $\Sigma(A) = (car(A), 0, n - car(A))$,
d) *A é semidefinita negativa se e solo se risulta* $car(A) < n$ *e* $\Sigma(A) = (0, car(A), n - car(A))$,
e) *A é indefinita se e solo se* $i_+(A) \cdot i_-(A) \neq 0$.

Dim. La tesi discende dal fatto che evidentemente A é definita positiva se e solo se risulta $V_+ = \mathbf{R}^n$ (e quindi $V_- = V^0 = \{\mathbf{0}_n\}$), é definita negativa se e solo se risulta $V^- = \mathbf{R}^n$ (e quindi $V^+ = V_0 = \{\mathbf{0}_n\}$) , é semidefinita positiva, (rispett. negativa), se e solo se risulta $V_0 \neq \{\mathbf{0}_n\}$) e $V_- = \{\mathbf{0}\}$, (rispett. $V^+ = \{\mathbf{0}\}$), é indefinita se V_+ e V_- sono diversi dallo spazio nullo $\{\mathbf{0}\}$.

■

Il precedente Corollario puó essere riformulato come segue:

Corollario 5.2 - (Criterio del segno degli autovalori).

a) *A é definita positiva se e solo se gli autovalori di A sono tutti maggiori di* 0.
b) *A é definita negativa se e solo se gli autovalori di A sono tutti minori di* 0.
c) *A é semidefinita positiva se e solo se se e solo se* 0 *é un autovalore di A e gli altri autovalori di A sono tutti maggiori di* 0,
d) *A é semidefinita negativa se e solo se* 0 *é un autovalore di A e gli altri autovalori di A sono tutti minori di* 0,
e) *A é indefinita se e solo se A ha due autovalori discordi.*

■

Infine, dalla Osservazione 4.8 discende la seguente ulteriore riformulazione dei Corollari 5.1 e 5.2:

Corollario 5.3 - (Criterio del segno dei coefficienti del polinomio caratteristico).

a) *A é definita positiva se e solo se i coefficienti del polinomio caratteristico sono diversi da* 0 *e di segno alternato;*
b) *A é definita negativa se e solo se i coefficienti del polinomio caratteristico sono diversi da* 0 *e dello stesso segno;*
c) *A é semidefinita positiva se e solo se gli ultimi h coefficienti di p_A sono nulli e i precedenti $n+1-h$ coefficienti sono tutti diversi da* 0 *e di segno alternato;*
d) *A é semidefinita negativa se e solo se gli ultimi h coefficienti di p_A sono nulli e i precedenti $n+1-h$ sono tutti diversi da* 0 *e dello stesso segno;*
e) *A é indefinita nei rimanenti casi.*

Esempio 9 - Ad esempio, nel caso delle matrici

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

il polinomio caratteristico é rispettivamente:

$$p_A(\lambda) = -\lambda^3 + 15\lambda^2 - 63\lambda + 81 \qquad e \qquad p_{A'}(\lambda) = -\lambda^3 + 11\lambda^2 - 24\lambda - 5.$$

i segni dei coefficienti del polinomio p_A sono alternati, e quindi A é definita positiva. (In effetti sappiamo che gli autovalori di A sono 3, (doppio), e 9, (semplice), e quindi sono tutti positivi).

Invece, nella sequenza dei coefficienti di $p_{A'}$ ci sono 2 variazioni e quindi $p_{A'}$ ha due radici positive ed una negativa; pertanto A' é indefinita.

■

Il segno di una forma quadratica puó anche essere indagato senza trovare il polinomio caratteristico, ma attraverso il segno dei minori principali. A tal fine é opportuno premettere alcune considerazioni.

Osservazione 5.4 - Ricordiamo che, per la Osservazione 1.8 del Cap. VII, il determinante di una matrice quadrata di ordine n avente n autovalori é uguale al prodotto di tali autovalori. Ne segue che per una matrice quadrata simmetrica A si ha quanto segue:

(1) se A é semidefinita (positiva o negativa), allora il determinante di A é nullo;
(2) se A é definita positiva, allora il determinante di A é positivo,
(3) se A é definita negativa, allora il determinante di A é positivo se n é pari, negativo se n é dispari,
(4) se n é pari e $\det(A) < 0$, allora A é indefinita.

■

Osservazione 5.5 - Si vede immediatamente che se A é definita positiva, (rispett. negativa), allora per ogni $k = 1, 2, \dots, n$ si ha che $A(k)$ é definita positiva, (rispett. negativa).

Infatti, per ogni $\mathbf{x} \in \mathbf{R}^k$, $\mathbf{x} \neq \mathbf{0}$, denotato $\hat{\mathbf{x}}$ il vettore di $\mathbf{R}^n$, le cui prime k componenti sono le componenti di $\mathbf{x}$ e le restanti $n-k$ componenti sono nulle, si ha evidentemente:

$$\varphi_{A(k)}(\mathbf{x}) = \mathbf{x}^T A(k)\, \mathbf{x} = \hat{\mathbf{x}}^T A \hat{\mathbf{x}} = \varphi_A(\hat{\mathbf{x}}) > 0, \qquad \text{(rispett. } < 0\text{)}.$$

■

Osservazione 5.6 - Se esiste $k = 1, 2, \dots, n$ tale che $A(k)$ é indefinita, allora A é indefinita.

Infatti, se esistono $\mathbf{u}, \mathbf{v} \in \mathbf{R}^k$ tali che $\varphi_{A(k)}(\mathbf{u}) < 0$ e $\varphi_{A(k)}(\mathbf{v}) > 0$, allora, denotati $\hat{\mathbf{u}}$ e $\hat{\mathbf{v}}$ i vettori di $\mathbf{R}^n$ ottenuti aggiungendo $n-k$ coordinate nulle alle coordinate di $\mathbf{u}$ e $\mathbf{v}$, risulta $\varphi_A(\hat{\mathbf{u}}) = \varphi_{A(k)}(\mathbf{u}) < 0$ e $\varphi_A(\hat{\mathbf{v}}) = \varphi_{A(k)}(\mathbf{v}) > 0$.

■

Tutto ció premesso si ha il seguente

Teorema 5.7 - (Criterio del segno dei minori principali) -. *La forma quadratica associata ad una matrice quadrata simmetrica A é*

a) *definita positiva se e solo se tutti i minori principali di A sono maggiori di 0;*
b) *definita negativa se e solo se i minori principali di ordine pari di A sono maggiori di 0 e quelli di ordine dispari sono minori di 0, e quindi se e solo se i minori principali di A sono di segno alternato e il primo minore principale a_{11} é minore di 0;*
c) *semidefinita positiva se i primi $n-1$ minori principali sono positivi e l'ultimo é nullo;*
d) *semidefinita negativa se i primi $n-1$ minori principali sono alternati in segno, il primo minore principale é negativo e l'ultimo é nullo;*
e) *indefinita se esiste $h \leq n$ tale che $m_h(A) \neq 0$ e i primi h minori principali non sono né tutti strettamente positivi e nemmeno sono alternati in segno con il primo minore principale negativo.*

Dim. Se A é definita positiva, allora, per ogni $k = 1, 2, \dots, n$, la matrice $A(k)$ é definita positiva per la Osservazione 5.5. Ne segue, per la Osservazione 5.4, che $m_k(A) = \det(A(k)) > 0$. Se invece A é definita positiva, allora, per ogni $k = 1, 2, \dots, n$, la matrice $A(k)$ é definita negativa per la Osservazione 5.5 e quindi, per la Osservazione 5.4, si ha che

$$m_k(A) = \det(A(k))\begin{cases} > 0 & \text{se } k \text{ é pari} \\ < 0 & \text{se } k \text{ é dispari} \end{cases}.$$

Viceversa, sappiamo che se i primi $n-1$ minori principali di A sono tutti diversi da 0, allora A é congruente alla matrice diagonale D avente per elementi diagonali i numeri

$$d_1 = m_1(A), \qquad d_2 = m_2(A)/m_1(A), \qquad d_3 = m_3(A)/m_2(A), \qquad \dots \qquad d_n = m_n(A)/m_{n-1}(A).$$

Ne segue che:

- se i minori principali di A sono tutti strettamente positivi, allora $d_1, d_2, \dots, d_n$ sono tutti strettamente positivi; la segnatura di A é $\Sigma(A) = (n, 0, 0)$ e quindi A é definita positiva;
- se i minori principali di A sono alternati in segno e il primo é negativo, allora $d_1, d_2, \dots, d_n$ sono tutti strettamente negativi; la segnatura di A é $\Sigma(A) = (0, n, 0)$ e quindi A é definita negativa;
- se i primi $n-1$ minori principali di A sono strettamente positivi e l'ultimo é nullo, allora i numeri reali $d_1, d_2, \dots, d_{n-1}$ sono strettamente positivi e $d_n = 0$; la segnatura di A é $\Sigma(A) = (n-1, 0, 1)$ e quindi A é semidefinita positiva;
- se i primi $n-1$ minori principali di A sono alternati in segno, il primo é negativo e l'ultimo é nullo, allora $d_1, d_2, \dots, d_{n-1}$ sono strettamente negativi e $d_n = 0$; la segnatura di A é $\Sigma(A) = (0, n-1, 1)$ e quindi A é semidefinita negativa.

Questo prova completamente a), b), c), d).

Infine la e) discende dalla Osservazione 5.6 e dal fatto che $A(h)$ é indefinita.
Infatti $A(h)$ non puó essere definita positiva, (altrimenti i minori principali di ordine $1, 2, \dots, h$ sarebbero tutti positivi), non puó essere definita negativa, (altrimenti i suddetti minori principali sarebbero alternati in segno con il primo negativo), non puó essere semidefinita positiva o semidefinita negativa, (altrimenti, per la Osservazione 5.4, sarebbe $m_h(A) = \det(A(h)) = 0$).

■

Osservazione 5.8 - Dalla e) del precedente Teorema 5.7 si deduce che in particolare A é indefinita se esiste un minore principale di ordine pari che é negativo oppure esistono due minori principali di ordine dispari che sono discordi oppure esiste un minore principale nullo ed un successivo minore principale diverso da 0.

Ad esempio, nel caso delle matrici dell'esempio 5,

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

si ha che

$$\begin{cases} m_1(A) = 5 > 0, \\ m_2(A) = 21 > 0, \\ m_3(A) = \det A = 81 > 0, \end{cases} \qquad \begin{cases} m_1(A') = 2 > 0, \\ m_2(A') = -1 < 0, \\ m_3(A') = \det A' = -5 < 0; \end{cases}$$

ne segue che A é definita positiva (poiché ha i minori principali tutti > 0) ed A' é indefinita, (poiché ha un minore di ordine pari negativo o perché ha due minori di ordine dispari discordi).
Del resto A e A' sono congruenti rispettivamente alle matrici D e D' i cui elementi diagonali sono

$$d_1 = 5,\ d_2 = 21/5\,,\ d_3 = 81/21 \qquad \text{e} \qquad d'_1 = 2,\ d'_2 = -1/2,\ d'_3 = 5.$$

Ne segue che $\Sigma(A) = (3,0,0)$ e $\Sigma(A) = (2,1,0)$ e questo conferma che A é definita positiva ed A' indefinita.

■

Esempio 10 - Consideriamo la matrice quadrata simmetrica

$$A = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 5 & 0 \\ -1 & 0 & a \end{pmatrix}.$$

I minori principali di A sono

$$m_1(A) = 1 > 0, \qquad m_2(A) = 5 - 4 = 1 > 0, \qquad m_3(A) = 5a - 5 - 4a = a - 5.$$

Ne segue che A é definita positiva se $a > 5$, semidefinita positiva se $a = 5$, indefinita se $a < 5$.

Del resto A é congruente alla matrice diagonale con elementi diagonali $1, 1, a - 5$ e quindi si ha

$$\Sigma(A) = \begin{cases} (3,0,0) & \text{se } a > 5 \ , \\ (2,0,1) & \text{se } a = 5 \ , \\ (2,1,0) & \text{se } a < 5 \end{cases}$$

Pertanto A é definita positiva se $a > 5$, semidefinita positiva se $a = 5$, indefinita se $a < 5$.

■

Osservazione 5.9 - Il criterio dei minori principali entra in crisi se i primi $k - 1$ minori principali (con $k < n$) sono strettamente positivi oppure sono alternati in segno e il primo é negativo, ma i successivi minori principali sono tutti nulli.

Infatti in questo caso c'é la possibilitá ma non la certezza che la matrice sia semidefinita positiva o semidefinita negativa.

Per sciogliere il dubbio occorrerebbe scambiare la k-sima riga e colonna con una delle successive o sommare la k-sima riga e colonna con una delle successive nella speranza che il nuovo minore principale di ordine k sia diverso da 0. Questo ci espone al rischio di eseguire a vuoto il calcolo di numerosi determinanti, (calcolo che é ovviamente oneroso se l'ordine della matrice non é piccolo); forse é preferibile eseguire l'algoritmo alla Gauss Jordan che permette di trovare la segnatura di A in maniera mirata o usare il metodo dei segni dei coefficienti del polinomio caratteristico, che trova la segnatura di A con il calcolo di un solo determinante.

■

Osservazione 5.10 - Talvolta si riesce a trovare la segnatura di una matrice quadrata simmetrica anche se i minori principali non sono tutti diversi da 0.

Ad esempio se i minori principali di una matrice del quarto ordine sono $m_1 = 2$, $m_2 = 0$, $m_3 = 5$, ed $m_4 = 0$, allora si vede facilmente che la segnatura di A é $\Sigma(A) = (1,2,1)$.

Infatti, il prodotto degli autovalori di $A(3)$ é $m_3 > 0$ e quindi tali autovalori sono tutti positivi oppure uno é positivo e gli altri due sono negativi; la prima possibilitá é esclusa, (altrimenti $\varphi_{A(3)}$ sarebbe definita positiva e quindi anche $m_2(A)$ sarebbe stato maggiore di 0), e quindi certamente un autovalore di $A(3)$ é positivo e gli altri due sono negativi. C'é quindi un sottospazio di dimensione 1 in $\mathbf{R}^3$ su cui $\varphi_{A(3)}$ é definita positiva ed un sottospazio di dimensione 2 su cui essa é definita negativa; aggiungendo la quarta coordinata nulla, si ottiene un sottospazio di dimensione 1 in $\mathbf{R}^4$ su cui φ_A é definita positiva ed un sottospazio di dimensione 2 su cui essa é definita negativa. Pertanto un autovalore di A é positivo e due autovalori sono negativi; infine l'ultimo autovalore é certamente 0, dal momento che $\det(A) = 0$, e la tesi é provata.

■

E S E R C I Z I

1 - Date le matrici quadrate simmetriche:

$$\begin{pmatrix} 3 & -2 & -3 \\ -2 & 5 & -1 \\ -3 & -1 & 6 \end{pmatrix} \qquad \begin{pmatrix} -1 & -2 & 0 \\ -2 & -3 & -1 \\ 0 & -1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & -2 & 1 \\ -2 & 8 & 2 \\ 1 & 2 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -2 & a \end{pmatrix},$$

a) scrivere la forma quadratica associata a tali matrici, trovarne gli autovalori e dedurne la segnatura e il segno della relativa forma quadratica;
b) trovare gli autovettori corrispondenti agli autovalori di tali matrici e verificare che gli autovettori corrispondenti ad autovalori distinti sono a due a due ortogonali; in presenza di autovalori multipli, trovare una base ortonormale di autovettori;
c) trovare quindi una matrice ortogonale che le diagonalizza;
d) ridurre in forma canonica la forma quadratica ad esse associata usando il metodo degli autovalori, il metodo di Lagrange e quello di Gauss Jordan;
e) studiare il segno della forma quadratica associata a tali matrici, utilizzando il criterio del segno dei coefficienti del polinomio caratteristico e il criterio di Sylvester.

2. - Date le matrici quadrate simmetriche:

$$\begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 4 \end{pmatrix} \qquad \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 4 \end{pmatrix} \qquad \begin{pmatrix} -3 & -2 & -1 \\ -2 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} -3 & -2 & -1 \\ -2 & -2 & 0 \\ -1 & 0 & -2 \end{pmatrix}.$$

$$\begin{pmatrix} -4 & 2 & -2 \\ 2 & -2 & 0 \\ -2 & 0 & \alpha \end{pmatrix} \qquad \begin{pmatrix} a & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 3 & -3 & 0 \\ -3 & 4 & a \\ 0 & a & 4 \end{pmatrix} \qquad \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 4 \end{pmatrix},$$

$$\begin{pmatrix} -2 & 2 & -2 \\ 2 & -5 & 4 \\ -2 & 4 & -5 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0 & -3 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 2 & 0 & -1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 1 & 0 & 1 & -1 \\ -1 & 2 & -1 & a \end{pmatrix}.$$

a) scrivere la forma quadratica associata a tali matrici e ridurla in forma canonica e in forma normale di Sylvester, usando il metodo di Lagrange e quello di Gauss Jordan;
b) indicare la segnatura e il segno di tali matrici,
c) studiarne il segno, utilizzando il criterio del segno dei coefficienti del polinomio caratteristico e il criterio del segno dei minori principali.

3. - Trovare le matrici A e B che definiscono le forme quadratiche

$$\varphi_A(x_1,x_2,x_3) = -x_1^2 - 5x_2^2 - x_3^2 + 4x_1x_2 + 2x_2x_3, \qquad \varphi_B(x_1,x_2,x_3,x_4) = 4x_1x_3 - 2x_2x_4 + 2x_4^2$$

e indicarne segnatura e segno, utilizzando rispettivamente il criterio del segno dei coefficienti del polinomio caratteristico e il criterio dei minori principali.

Ridurre poi in forma canonica e in forma normale di Sylvester tali forme quadratiche, usando rispettivamente il metodo di Lagrange e quello di Gauss Jordan.

4. - Usando il metodo di Lagrange e quello di Gauss Jordan, ridurre in forma canonica e in forma normale di Sylvester le forme quadratiche:

$$\varphi_A(x_1,x_2,x_3) = 2x_1^2 + 4x_2^2 + 6x_3^2 + 2x_1x_2 + 6x_1x_3, \qquad \varphi_B(x_1,x_2,x_3,x_4) = x_2^2 - x_3^2 + x_1x_2 + 2x_2x_4.$$

Trovare le matrici A e B che definiscono tali forme quadratiche e indicarne segnatura e segno, utilizzando il criterio del segno dei coefficienti del polinomio caratteristico e il criterio del segno dei minori principali.

5. - Trovare il segno e la segnatura delle matrici quadrate simmetriche A, B, e C sapendo che il loro polinomio caratteristico é rispettivamente

$$p_A(\lambda) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2, \qquad p_B(\lambda) = -\lambda^3 + 11\lambda - 6; \qquad p_C(\lambda) = -\lambda^3 + 3\lambda^2 - 2\lambda.$$

6. - Trovare (al variare del parametro a) il segno e la segnatura della matrice quadrata simmetrica di ordine 4 il cui polinomio caratteristico é

$$p_A(\lambda) = \lambda^4 - 3\lambda^3 + (1+a)\lambda^2 + (1-a)\lambda + a^2 - 1.$$

7. - I minori principali di una matrice quadrata simmetrica A di ordine 4 sono rispettivamente i numeri $3,\ 2,\ a,\ a^2-1$; trovare una matrice diagonale congruente ad A e dedurne (al variare del parametro a) il segno e la segnatura di A.

8. - I minori principali di una matrice quadrata simmetrica A sono rispettivamente: $-2,\ 6,\ a,\ -a-1$; dedurne (al variare del parametro a) il segno e la segnatura di A ed una matrice diagonale congruente ad A.

Capitolo IX - COMPLEMENTI

1. Spazi vettoriali astratti

Ricordiamo che nella Osservazione 1.5 del Cap.I, abbiamo accennato alla possibilitá di estendere il concetto di vettori e spazio vettoriale. Possiamo ora precisare tale concetto secondo la seguente

Definizione 1.1. *Dicesi* **spazio vettoriale reale**, *(rispett.* **complesso***), un insieme non vuoto V in cui sono definite una operazione di addizione (che associa ad ogni coppia ordinata di elementi $\mathbf{x}, \mathbf{y}$ di V uno ed un solo elemento $\mathbf{x} + \mathbf{y}$ di V), ed una operazione esterna, (che associa ad ogni numero reale (o complesso) α e ad ogni $\mathbf{x} \in V$ uno ed un solo elemento $\alpha\mathbf{x} \in V$), in modo che siano verificate le seguenti proprietá:*

(A1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, *per ogni* $\mathbf{x}, \mathbf{y} \in V$,
(A2) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$, *per ogni* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$,
(A3) *esiste un elemento* $\mathbf{0}$ *di* V *tale che* $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$, *per ogni* $\mathbf{x} \in V$,
(A4) *per ogni* $\mathbf{x} \in V$ *esiste un elemento* $-\mathbf{x}$ *di* V, *tale che* $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$,
(M1) $\alpha \cdot (\mathbf{x} + \mathbf{y}) = \alpha \cdot \mathbf{x} + \alpha \cdot \mathbf{y}$, *per ogni* $\alpha \in \mathbf{R}$ *e per ogni* $\mathbf{x}, \mathbf{y} \in V$,
(M2) $(\alpha + \beta) \cdot \mathbf{x} = \alpha \cdot \mathbf{x} + \beta \cdot \mathbf{x}$, *per ogni* $\alpha, \beta \in \mathbf{R}$ *e per ogni* $\mathbf{x} \in V$,
(M3) $\alpha \cdot (\beta \cdot \mathbf{x}) = (\alpha\beta) \cdot \mathbf{x}$, *per ogni* $\alpha, \beta \in \mathbf{R}$ *e per ogni* $\mathbf{x} \in V$,
(M4) $1 \cdot \mathbf{x} = \mathbf{x}$, *per ogni* $\mathbf{x} \in V$.

Gli elementi di V diconsi **vettori**, *mentre i numeri reali (o complessi) diconsi* **scalari** .

■

Gli insiemi $\mathbf{R}^n$ (e $\mathbf{C}^n$) sono esempi di spazi vettoriali reali (o complessi); altri esempi di spazi vettoriali sono:

(1) gli spazi delle matrici di numeri reali (o complessi) di tipo $k \times n$, con le consuete operazioni di addizione tra matrici e di moltiplicazione di un numero (reale o complesso) per una matrice,
(2) gli spazi delle funzioni reali (o complesse) definite in un insieme X con le usuali operazioni di addizione di due funzioni e di moltiplicazione di una funzione costante per una funzione;
(3) l'insieme dei polinomi a coefficienti reali o complessi con le consuete operazioni tra polinomi.

■

Dagli assiomi (A1-A4), (M1-M4) di spazio vettoriale si deduce facilmente che:

Proposizione 1.2. *Se V é uno spazio vettoriale, allora risulta:*

(1) *il vettore* $\mathbf{0}$ *soddisfacente la (A3) é unico e dicesi il* **vettore nullo**;
(2) *per ogni* $\mathbf{x} \in V$ *il vettore* $-\mathbf{x}$ *soddisfacente la (A4) é unico ed é detto il* **vettore opposto** *di* $\mathbf{x}$;
(3) *risulta* $\mathbf{x} + \mathbf{y} = \mathbf{z}$ *se e solo se* $\mathbf{x} = \mathbf{z} + (-\mathbf{y})$;
(4) *se risulta* $\mathbf{x} + \mathbf{x} = \mathbf{x}$, *allora é* $\mathbf{x} = \mathbf{0}$;
(5) $0 \cdot \mathbf{x} = \mathbf{0}$ *ed* $\alpha \cdot \mathbf{0} = \mathbf{0}$, *per ogni* $\mathbf{x} \in V$ *e per ogni* $\alpha \in \mathbf{R}$;
(6) $-(\mathbf{x} + \mathbf{y}) = (-\mathbf{x}) + (-\mathbf{y})$, *per ogni* $\mathbf{x}, \mathbf{y} \in V$;
(7) $(-\alpha) \cdot \mathbf{x} = \alpha \cdot (-\mathbf{x}) = -(\alpha \cdot \mathbf{x})$, *per ogni* $\alpha \in \mathbf{R}, \mathbf{x} \in V$.

■

2. - Sottospazi vettoriali

Il concetto di sottospazio vettoriale puó essere esteso agli spazi vettoriali astratti, nel senso che

Definizione 2.1 -. *Dicesi sottospazio vettoriale di uno spazio vettoriale V ogni insieme non vuoto $H \subseteq \mathbf{V}$ tale che:*

(1) *per ogni* $\mathbf{u}, \mathbf{v} \in H$ *risulta* $\mathbf{u} + \mathbf{v} \in H$,
(2) *per ogni* $\alpha \in \mathbf{R}$ *e per ogni* $\mathbf{v} \in H$ *risulta* $\alpha \cdot \mathbf{v} \in H$.

■

Osservazione 2.2 - La giustificazione di tale definizione consiste nel fatto che se sussistono le proprietá (1) e (2), allora H, munito della restrizione ad $H \times H$ dell'operazione di addizione e della restrizione ad $\mathbf{R} \times H$ della moltiplicazione esterna, é esso stesso uno spazio vettoriale, dal momento che le proprietá (A1-A4) ed (M1-M4) sono verificate in V e quindi in H.

■

Oltre gli esempi giá noti di sottospazi di $\mathbf{R}^n$, altri esempi di sottospazi di uno spazio vettoriale sono i seguenti:

(1) l'insieme delle funzioni continue (o derivabili) in un intervallo I di $\mathbf{R}$ é un sottospazio vettoriale dello spazio vettoriale delle funzioni reali definite in I, con le usuali operazioni di addizione di due funzioni e di moltiplicazione di una funzione costante per una funzione;

(2) per ogni $n \in \mathbf{N}$, l'insieme $\mathcal{P}_n$ dei polinomi di grado minore o uguale ad n é un sottospazio vettoriale dello spazio vettoriale di tutti i polinomi.

Infatti, la somma di due funzioni continue (rispett. derivabili) in I é ancora una funzione continua, (rispett. derivabile), e il prodotto di una funzione costante per una funzione continua (rispett. derivabile) in I é ancora una funzione continua (rispett. derivabile) in I.

Analogamente, la somma di due polinomi (di grado minore o uguale ad n é un polinomio di grado minore o uguale ad n e il prodotto di una costante per un polinomio di grado minore o uguale ad n é un polinomio di grado minore o uguale ad n.

■

Agli spazi vettoriali astratti puó essere esteso il concetto di combinazione lineare di due o piú vettori e di **sottospazio generato** da un insieme finito $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ di vettori di V: esso é l'insieme di tutte le combinazioni lineari dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, cioé l'insieme

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) = \{\mathbf{x} \in \mathbf{R}^n \quad | \quad \exists \alpha_1, \alpha_2, \ldots \alpha_k \in \mathbf{R} \text{ tali che } \mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_k \mathbf{v}_k\}.$$

Tale insieme é un sottospazio di V, contiene $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ ed é contenuto in ogni sottospazio che contiene $\{\mathbf{v}_1 \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

■

3 . - Spazi di dimensione finita o infinita

Agli spazi vettoriali astratti possono essere estese le definizioni di **vettori linearmente dipendenti** e di **vettori linearmente indipendenti** date nel n. 2 del Cap. III .

Ad esempio, per ogni $n \in \mathbf{N}$, i polinomi $p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2, \quad \ldots, \quad p_n(x) = x^n$ sono linearmente indipendenti.

Si puó inoltre estendere il concetto di insieme di generatori nel senso che un insieme finito $X \subset V$ dicesi **insieme di generatori** dello spazio vettoriale V, se e solo se V coincide con il sottospazio generato da X.

Si dá allora la seguente

Definizione 3.1 -. *Si dice che V é uno* **spazio vettoriale di dimensione finita**, *se possiede un insieme finito di generatori, cioé se e solo se esiste un insieme finito $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ tali che:*

$$\text{per ogni } \mathbf{v} \in V \text{ esistono } \alpha_1, \alpha_2, \ldots \alpha_k \in \mathbf{R} \text{ tali che } \quad \mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_k \mathbf{v}_k.$$

Si dice invece che V é uno **spazio di dimensione infinita**, *se non é di dimensione finita, cioé se per ogni insieme finito $X \subset V$ si ha che lo spazio generato da X é diverso da V.*

Se V é uno spazio di dimensione finita, il minimo numero di generatori di V dicesi **dimensione di V.**

■

Ad esempio, per ogni $n \in \mathbf{N}$ l'insieme $\mathcal{P}_n$ dei polinomi di grado minore o uguale ad n é un sottospazio vettoriale di dimensione finita, poiché, (ad esempio), un insieme di generatori di $\mathcal{P}_n$ é l'insieme dei polinomi $\{p_0, p_1, \dots, p_n\}$, dove $p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2, \quad \dots, \quad p_n(x) = x^n$.
Infatti per ogni polinomio $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \in \mathcal{P}_n$ risulta $p = a_0p_0 + a_1p_1 + \dots + a_np_n$.

Invece l'insieme di tutti i polinomi $\mathcal{P}$ é uno spazio di dimensione infinita.
Infatti, se $q_1, q_2, \dots, q_k$ sono k polinomi, detto n il massimo grado di tali polinomi, si ha chiaramente che lo spazio generato da $q_1, q_2, \dots, q_k$ é contenuto in $\mathcal{P}_n$ e quindi é diverso da $\mathcal{P}$.

■

Definizione 3.2 -. *Se V é uno spazio vettoriale di dimensione finita, dicesi base di V un insieme finito di elementi di V che siano linearmente indipendenti e generino lo spazio V.*

■

Ad esempio i polinomi $p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2, \quad \dots, \quad p_n(x) = x^n$ formano una base dello spazio $\mathcal{P}_n$ dei polinomi di grado minore o uguale ad n.

Si dimostra che agli spazi astratti possono essere estese tutte le proprietá enunciate nel Cap.III , tranne la Prop. 4.4 che é riservata agli spazi vettoriali di dimensione finita.
In particolare, si dimostra che se V é uno spazio di dimensione n, allora n rappresenta il numero di elementi di una qualunque base e il massimo numero di vettori linearmente indipendenti dello spazio.
Inoltre, se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sono n vettori di V, allora tali vettori formano una base di V se e solo se per ogni $\mathbf{v} \in V$ esiste uno ed un solo $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$ tale che $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n$; i numeri reali $x_1, x_2, \dots x_n$ diconsi le **coordinate di v rispetto alla base** $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$.

■

4 . - Trasformazioni lineari ed isomorfismi

Agli spazi vettoriali astratti puó essere esteso il concetto di trasformazione lineare secondo la seguente

Definizione 4.1 -. *Se V e W sono due spazi vettoriali reali (o complessi), dicesi* **trasformazione o operatore lineare di V in W** *una qualunque funzione $L : V \mapsto W$ tale che*

a) $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$ *per ogni $\mathbf{u}, \mathbf{v} \in V$,*
b) $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u})$ *per ogni $\mathbf{u} \in V$ e per ogni $\alpha \in \mathbf{R}$ (rispett. $\alpha \in \mathbf{C}$).*

Una trasformazione lineare bigettiva di V in W dicesi **isomorfismo** *tra V e W; in tal caso si dice anche che V e W sono spazi* **isomorfi .**
Una trasformazione lineare di V in V dicesi anche **endomorfismo** *su V.*

■

Ad esempio, per ogni $n \in \mathbf{N}$, l'applicazione L da $\mathbf{R}^{n+1}$ nello spazio $\mathcal{P}_n$ dei polinomi di grado minore o uguale ad n , che associa ad ogni vettore $(a_0, a_1, \dots, a_n)$ di $\mathbf{R}^{n+1}$ il polinomio $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, é un isomorfismo tra $\mathbf{R}^{n+1}$ e $\mathcal{P}_n$. La funzione inversa di L associa ad ogni polinomio p il vettore in $\mathbf{R}^{n+1}$ dei suoi coefficienti, ed é un isomorfismo tra $\mathcal{P}_n$ ed $\mathbf{R}^{n+1}$.

Un altro esempio di operatore lineare é fornito dagli operatori di derivazione $D, D^2, \dots D^k$ che associano ad ogni $x = x(t)$ nello spazio V delle funzioni infinitamente derivabili di $\mathbf{R}$ in sé rispettivamente la sua derivata prima, la sua derivata seconda, ... , la sua derivata k-esima.

■

Osservazione 4.2 - Alle trasformazioni lineari tra spazi vettoriali astratti si possono estendere le proprietá delle trasformazioni lineari dimostrate nel Teorema 2.1 del Cap. V; in particolare una trasformazione lineare iniettiva trasforma un sottospazio di dimensione finita di V in un sottospazio avente la stessa dimensione. Ne segue che due spazi vettoriali isomorfi di dimensione finita hanno la stessa dimensione.

∎

Osservazione 4.3 - Si vede inoltre facilmente che la somma di due trasformazioni lineari di V in W e il multiplo di una trasformazione lineare di di V in W sono trasformazioni lineari di V in W; inoltre la composta di due trasformazioni lineari é una trasformazione lineare e la funzione inversa di un isomorfismo di V in W é un isomorfismo di W in V.

∎

Un risultato fondamentale sugli spazi di dimensione finita é contenuto nella seguente

Proposizione 4.4. *Si ha quanto segue.*

(1) *Ogni spazio vettoriale V di dimensione n é isomorfo ad $\mathbf{R}^n$.*
(2) *Due spazi vettoriali V e W di dimensione n sono isomorfi.*

Infatti, se V é uno spazio vettoriale reale di dimensione n e $\{\mathbf{v}_1, \mathbf{v}_2, \ldots ., \mathbf{v}_n\}$ é una sua base , allora l'applicazione L che associa ad ogni vettore $(\alpha_1, \alpha_2, \ldots \alpha_n)$ di $\mathbf{R}^n$ il vettore $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$ di V é un isomorfismo di $\mathbf{R}^n$ in V.

La seconda parte della tesi discende dalla prima; in effetti, esiste un isomorfismo $L : \mathbf{R}^n \longrightarrow V$ ed un isomorfismo $L' : \mathbf{R}^n \longrightarrow W$, e quindi la funzione composta $L'L^{-1}$ é un isomorfismo tra V e W.

∎

Osservazione 4.5 - La precedente Prop. 4.4 consente di trasferire agli spazi vettoriali astratti V di dimensione finita quanto studiato per i vettori di $\mathbf{R}^n$.

Ad esempio k vettori $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ di V sono linearmente indipendenti (o dipendenti) se e solo se i vettori delle coordinate di tali vettori rispetto ad una fissata base di V sono linearmente indipendenti (o dipendenti) in $\mathbf{R}^n$, e quindi se e solo se la matrice avente per righe tali vettori ha caratteristica k (rispett. minore di k).

Piú in generale il sottospazio generato da $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ é isomorfo al sottospazio di R^n generato dai vettori delle coordinate di $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$, e quindi ha la sua stessa dimensione.

∎

Osservazione 4.6 - Se V é uno spazio vettoriale di dimensione finita e $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ é una sua base, allora i valori assunti da una trasformazione lineare L di V in un altro spazio vettoriale W dipendono esclusivamente dai valori che L assume nei vettori della base.

Infatti, per ogni $\mathbf{v} \in V$ esiste uno ed un solo vettore $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbf{R}^n$ tale che

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n \qquad \text{e quindi} \qquad L(\mathbf{v}) = x_1L(\mathbf{v}_1) + x_2L(\mathbf{v}_2) + \ldots + x_nL(\mathbf{v}_n).$$

∎

Osservazione 4.7 - Siano V e W due spazi vettoriali di dimensione finita e siano $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\}$ una base di V e W rispettivamente; allora, una trasformazione lineare L di V in W puó essere completamente descritta dalla matrice A avente per colonne i vettori delle coordinale dei vettori $L(\mathbf{v_1}), L(\mathbf{v_2}), \ldots, L(\mathbf{v_n})$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\}$.

La matrice A dicesi la **matrice della trasformazione lineare** L rispetto alle basi $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ e $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\}$. Essa descrive il legame tra il vettore $\mathbf{x}$ delle coordinate del generico vettore $\mathbf{v} \in V$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ e il vettore $\mathbf{y}$ delle coordinate del vettore $L(\mathbf{v})$ rispetto alla base $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_k}\}$, nel senso che risulta $\mathbf{y} = A\mathbf{x}$.

Ad esempio, siano V lo spazio dei polinomi di grado minore o uguale a 3 e W lo spazio dei polinomi di grado minore o uguale a 2 e siano

$$p_1(t) = 1, \quad p_2(t) = t, \quad p_3(t) = t^2, \quad p_4(t) = t^3 \qquad \text{e} \qquad p_1(t) = 1, \quad p_2(t) = t, \quad p_3(t) = t^2$$

le basi canoniche di V e W; sia poi L la trasformazione lineare di V in W che associa ad ogni $p \in V$ il polinomio $L(p) = p'' - 2p'$, (dove p' e p'' sono le derivate prime e seconde di p).

La matrice della trasformazione L di V in W rispetto alle basi $\{p_1, p_2, p_3, p_4\}$ e $\{p_1, p_2, p_3\}$ é la matrice A avente per colonne i vettori delle coordinate di $L(p_1), L(p_2), L(p_3), L(p_4)$ rispetto alla base $\{p_1, p_2, p_3\}$ di W.

Ebbene, si ha:

$$p_1' = 0, \quad p_1'' = 0, \qquad p_2' = p_1, \quad p_2'' = 0, \qquad p_3' = 2p_2, \quad p_3'' = 2p_1, \qquad p_4' = 3p_3, \quad p_4'' = 6p_2,$$

e quindi

$$L(p_1) = 0, \qquad L(p_2) = -2p_1, \qquad L(p_3) = 2p_1 - 4p_2, \qquad L(p_4) = 6p_2 - 6p_3.$$

I vettori delle coordinate di $L(p_1)$, $L(p_2)$, $L(p_3)$, $L(p_4)$ rispetto alla base $\{p_1, p_2, p_3\}$ di W sono dunque rispettivamente $(0, 0, 0)^T$, $(-2, 0, 0)^T$, $(2, -4, 0)^T$, $(0, 6, -6)^T$, e quindi la matrice A della trasformazione L di V in W rispetto alle basi $\{p_1, p_2, p_3, p_4\}$ e $\{p_1, p_2, p_3\}$ é

$$A = \begin{pmatrix} 0 & -2 & 2 & 0 \\ 0 & 0 & -4 & 6 \\ 0 & 0 & 0 & -6 \end{pmatrix}.$$

∎

Osservazione 4.8 - Alle trasformazioni lineari $L : V \mapsto W$ tra spazi vettoriali di dimensione finita puó essere esteso il teorema della dimensione

$$\dim V = \dim(Ker(L)) + \dim(Im(L)).$$

Infatti, se A é la matrice di L rispetto ad assegnate basi, allora $Im(L)$ é isomorfo allo spazio generato dalle colonne di A e $Ker(L)$ é isomorfo all'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$.

Ne segue che risulta $dim(Im(L)) = car(A)$ e $dim(Ker(L)) = n - car(A)$ e quindi

$$\dim(Ker(L)) + \dim(Im(L)) = n - car(A) + car(A) = n = dim(V).$$

∎

5 . - Il caso degli endomorfismi

Consideriamo in particolare il caso di un endomorfismo, cioé di una trasformazione lineare L di uno spazio vettoriale V in sé. A tali trasformazioni possono essere estese le nozioni di autovalore ed autovettore di una matrice quadrata.

Definizione 5.1. *Un numero reale (o complesso) λ dicesi* **autovalore** *di L, se esiste $\mathbf{v} \neq \mathbf{0}$ tale che $L(\mathbf{v}) = \lambda\mathbf{v}$; in tal caso $\mathbf{v}$ dicesi* **autovettore di** *L corrispondente all'autovalore λ.*

Inoltre, si dice autospazio relativo all'autovalore λ l'insieme $V(\lambda)$ dei vettori $\mathbf{v}$ tali che $L(\mathbf{v}) = \lambda\mathbf{v}$, cioé il nucleo di $L - \lambda I$, (dove I rappresenta la trasforrmazione identica di V, cioé l'applicazione che associa ad ogni vettore $\mathbf{v}$ sé stesso).

Infine, si dice **molteplicitá geometrica** *di λ la dimensione dell'autospazio relativo a λ.*

∎

Chiaramente gli autovettori di L rappresentano i vettori che vengono trasformati mediante L in vettori proporzionali, e quindi rappresentano le direzioni delle rette invarianti nella trasformazione L.

Esempio Sia V lo spazio delle funzioni infinitamente derivabili di $\mathbf{R}$ in sé e sia L la trasformazione lineare di V in sé che associa ad ogni funzione $x = x(t) \in V$ la sua derivata $L(x) = x'$.

Si vede allora facilmente che ogni numero reale λ é un autovalore di L e la funzione $x_\lambda = e^{\lambda t}$ é un corrispondente autovettore, poiché risulta $x'_\lambda = \lambda x_\lambda$; inoltre per ogni funzione infinitamente derivabile $x = x(t)$ si ha che $x' = \lambda x$ se e solo se esiste $c \in \mathbf{R}$ tale che $x = ce^{\lambda t} = cx_\lambda$. Pertanto l'autospazio relativo all'autovalore λ é lo spazio di dimensione 1 generato da x_λ e quindi la molteplicitá geometrica dell'autovalore λ é 1.

■

Osservazione 5.2 - Se V é uno spazio vettoriale di dimensione finita ed L é una trasformazione lineare di V in sé, allora, scegliendo nello spazio di partenza V e nello spazio di arrivo V la stessa base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$, si ha che L é rappresentata da una matrice quadrata A, (precisamente la matrice avente per colonne i vettori delle coordinate di $L(\mathbf{v_1}), L(\mathbf{v_2}), \dots, L(\mathbf{v_n})$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$.

Tale matrice dicesi la **matrice della trasformazione lineare L rispetto alla base** $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e trasforma il generico punto $\mathbf{v} \in V$ di coordinate $\mathbf{x}$ rispetto alla base $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$) nel punto $L(\mathbf{v})$ di coordinate $\mathbf{y} = A\mathbf{x}$ rispetto alla stessa base.

Ne segue che λ é un autovalore di L se e solo se λ é autovalore della matrice A e la molteplicitá di λ come autovalore di L coincide con la moleplicitá geometrica di λ come autovalore di A.

■

Osservazione 5.3 - Sia L un endomorfismo su V e sia A la matrice di L rispetto ad una base di V; si vede facilmente che, cambiando base in V, la matrice A viene trasformata in una matrice simile e viceversa ogni matrice simile ad A rappresenta la matrice della trasformazione lineare L rispetto ad un'altra base.

Ne segue che l'immagine e il nucleo di L, gli autovalori di L, gli autospazi relativi agli autovalori di L e le dimensioni di tutti questi sottospazi sono chiaramente indipendenti dalla base scelta in V, e dunque comuni a tutte le matrici (simili tra loro) che rappresentano L, ed infatti sappiamo che due matrici simili, hanno la stessa caratteristica, lo stesso determinante, lo stesso polinomio caratteristico, gli stessi autovalori con la stessa molteplicitá algebrica e geometrica.

D'altra parte é naturale cercare di scegliere la base di V in modo che L sia rappresentata da una matrice quanto piú semplice possibile; l'ideale sarebbe riuscire a rappresentare L con una matrice diagonale, perché in tal caso, detti $\lambda_1, \lambda_2, \dots, \lambda_n$ gli elementi diagonali di tale matrice, la trasformazione lineare L si limita a trasformare il generico vettore $\mathbf{v}$ di coordinate $(x_1, x_2, \dots, x_n)^T$ nel vettore $L(\mathbf{v})$ di coordinate $(\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n)^T$.

Ebbene, questo é possibile se e solo se la matrice A é simile ad una matrice diagonale Λ, cioé se e solo se A é una matrice **diagonalizzabile**, e sappiamo che questo accade se e solo se la somma delle molteplicitá geometriche degli autovalori reali di A é uguale ad n.

Se invece il polinomio caratteristico di A ha radici complesse non reali oppure se A ha un autovalore singolare, allora A non é diagonalizzabile, e quindi non esiste una base di V rispetto a cui L é rappresentata da una matrice diagonale.

■

6. - Spazi euclidei

Definizione 6.1 -. *Dicesi* **spazio euclideo** *uno spazio vettoriale reale V in cui é definito un* **prodotto scalare**, *cioé una funzione $\varphi : V \times V \mapsto \mathbf{R}$ tale che per ogni $\alpha \in \mathbf{R}$ e per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ si ha:*

(1) $\varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x})$,
(2) $\varphi(\mathbf{x} + \mathbf{y}, \mathbf{z}) = \varphi(\mathbf{x}, \mathbf{z}) + \varphi(\mathbf{y}, \mathbf{z})$, $\quad \varphi(\mathbf{z}, \mathbf{x} + \mathbf{y}) = \varphi(\mathbf{z}, \mathbf{x}) + \varphi(\mathbf{z}, \mathbf{y})$,
(3) $\varphi(\alpha\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{x}, \alpha\mathbf{y}) = \alpha\varphi(\mathbf{x}, \mathbf{y})$,
(4) $\varphi(\mathbf{x}, \mathbf{x}) > 0$ *per ogni* $\mathbf{x} \in V - \{\mathbf{0}\}$.

In tal caso si dice anche che φ é una **forma bilineare simmetrica definita positiva**.

■

Esempio 1 - Lo spazio $\mathbf{R}^n$ con l'usuale prodotto scalare é uno spazio euclideo. In generale, qualunque sottospazio V di R^n, con la restrizione a $V \times V$ del prodotto scalare di $\mathbf{R}^n$, é uno spazio euclideo.

■

Esempio 2 - Se A é una matrice quadrata simmetrica definita positiva di ordine n, allora la funzione $\varphi : \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ tale che $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$, verifica le suddette proprietá e quindi é un prodotto scalare in $\mathbf{R}^n$.

■

Esempio 3 - Sia V lo spazio delle funzioni continue in un intervallo $[a, b]$, e consideriamo la funzione $\varphi : V \times V \mapsto \mathbf{R}$ tale che

$$\varphi(u, v) = \int_a^b u(x)v(x)\,dx, \qquad \text{per ogni} \quad u, v \in V = \mathcal{C}([a, b]).$$

Si vede facilmente che φ soddisfa tutte le proprietá dei prodotti scalari e quindi lo spazio V delle funzioni continue in $[a, b]$, munito di tale prodotto scalare, é un esempio di spazio euclideo di dimensione infinita.

Se φ é un prodotto scalare in V, allora, per ogni $\mathbf{x} \in V$, il numero reale $\sqrt{\varphi(\mathbf{x}, \mathbf{x})} \geq 0$ dicesi norma (o meglio **φ-norma)** euclidea di $\mathbf{x}$. Due vettori $\mathbf{x}$ e $\mathbf{y}$ di V diconsi **ortogonali o φ-ortogonali,** se risulta $\varphi(\mathbf{x}, \mathbf{y}) = 0$.
Si vede facilmente che k vettori non nulli a due a due ortogonali, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$, sono sempre linearmente indipendenti, e quindi formano una base dello spazio da essi generato. La base é una base **ortonormale** se ciascun vettore ha norma unitaria, cioé se risulta $\varphi(\mathbf{u}_j, \mathbf{u}_j) = 1$ per ogni j.

■

Osservazione 6.2 - Il procedimento di ortonormalizzazione di Gram-Schmidt puó essere esteso agli spazi euclidei astratti, nel senso che se U é un sottospazio di dimensione finita di uno spazio euclideo V ed $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$ é una base di U, allora attraverso il procedimento ricorsivo:

$$\begin{cases} \mathbf{u}_1 & = \mathbf{a}_1, \\ \mathbf{u}_j & = \mathbf{a}_j - \dfrac{\varphi(\mathbf{a}_j, \mathbf{u}_1)}{\varphi(\mathbf{u}_1, \mathbf{u}_1)} \cdot \mathbf{u}_1 - \dfrac{\varphi(\mathbf{a}_j, \mathbf{u}_2)}{\varphi(\mathbf{u}_2, \mathbf{u}_2)} \cdot \mathbf{u}_2 - \ldots \dfrac{\varphi(\mathbf{a}_j, \mathbf{u}_{j-1})}{\varphi(\mathbf{u}_{j-1}, \mathbf{u}_{j-1})} \cdot \mathbf{u}_{j-1}), \qquad \text{per } j = 2, 3, \ldots, k\ , \end{cases}$$

si costruisce una base di U formata da vettori a due a due φ-ortogonali. Per avere una base ortonormale, basta **normalizzare** tali vettori, cioé sostituire ciascun $\mathbf{u}_j$ con $\hat{\mathbf{u}}_j = (1/||\mathbf{u}_j||)\mathbf{u}_j$.

■

Osservazione 6.3 - Se U é un sottospazio di V, l'insieme $U^\perp = \{\mathbf{x} \in V | \varphi(\mathbf{x}, \mathbf{u}) = 0 \quad \text{per ogni } \mathbf{u} \in U\}$ é un sottospazio vettoriale di V che prende il nome di **complemento φ-ortogonale di U.**
Se V ha dimensione finita, allora risulta $\dim(U^\perp) = \dim V - \dim(U)$.

Osservazione 6.4 - Le proprietá delle matrici ortogonali possono essere estese agli spazi euclidei astratti, nel senso che se $(V+, \cdot, \varphi)$ é uno spazio euclideo ed $L : V \mapsto V$ é una trasformazione lineare tale che

$$\varphi(L(\mathbf{x}), L(\mathbf{y})) = \varphi(\mathbf{x}, \mathbf{y}) \qquad \text{per ogni} \quad \mathbf{x}, \mathbf{y} \in V,$$

allora L conserva la norma, la distanza, l'ortogonalitá e l'angolo tra due vettori di V.
Per questo motivo si dice che L é un operatore **unitario** in V.

■

Osservazione 6.5 - Se A é una matrice quadrata simmetrica di ordine n, allora si ha:

$$(A\mathbf{x}|\mathbf{y}) = \mathbf{y}^T A\mathbf{x} = (\mathbf{y}^T A\mathbf{x})^T = \mathbf{x}^T A^T \mathbf{y} = \mathbf{x}^T A\mathbf{y} = (\mathbf{x}|A\mathbf{y}) \qquad \text{per ogni } \mathbf{x}, \mathbf{y} \in \mathbf{R}^n.$$

In generale, se V é uno spazio euclideo ed L é una trasformazione lineare di V in sé tale che

$$\varphi(L(\mathbf{x}), \mathbf{y}) = \varphi(\mathbf{x}, L(\mathbf{y})) \qquad \text{per ogni } \mathbf{x}, \mathbf{y} \in V,$$

allora si dice che L é un **operatore simmetrico**.

Se L é un operatore lineare simmetrico su uno spazio euclideo V di dimensione n, allora L ha n autovalori reali $\lambda_1, \lambda_2, \dots, \lambda_n$, (contando ciascun autovalore h volte se h é la sua molteplicitá geometrica), ed n autovettori a due a due ortogonali, che formano una base ortonormale di V; rispetto a tale base L é rappresentata dalla matrice diagonale i cui elementi diagonali sono gli autovalori di L.

■