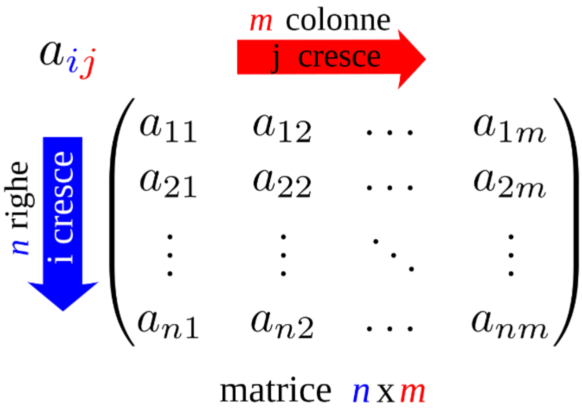
Matrici

Una matrice a coefficienti in un campo numerico $\mathbb{K}$ è una **tabella di numeri**:



Gli elementi che compongono la matrice vengono detti **coefficienti**, i quali appartengono ad un dato **campo numerico**.

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

La matrice A è costituita da 2 righe e 3 colonne, cioè si tratta di una matrice 2×3.

Una matrice costituita da una sola riga viene detta **matrice riga** (o semplicemente **riga**), mentre una matrice costituita da una sola colonna viene detta **matrice colonna** (o semplicemente **colonna**).

La riga m della matrice A si indica con la notazione $A^{(m)}$, mentre la colonna n della matrice A si indica con la notazione $A_{(n)}$.

Operazioni con le matrici

Moltiplicazione di un numero per una matrice

Dato un numero α ed una matrice *A* di dimensioni $m \times n$, si indica con αA la matrice ottenuta moltiplicando tutti i coefficienti di *A* per il numero α.

$$A = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & -1 & \sqrt{2} \end{pmatrix} \quad \alpha = 6$$

$$\alpha A = \begin{pmatrix} 6 \cdot 1 & 6 \cdot 0 & 6 \cdot \frac{1}{2} \\ 6 \cdot 0 & 6 \cdot (-1) & 6 \cdot \sqrt{2} \end{pmatrix} = \begin{pmatrix} 6 & 0 & 3 \\ 0 & -6 & 6\sqrt{2} \end{pmatrix}$$

La moltiplicazione di un numero per una matrice lascia inalterate le dimensioni della matrice: la matrice αA ha le stesse dimensioni della matrice *A*.

Addizione di matrici

Date due matrici *A* e *B*, entrambe di dimensioni $m \times n$, la **matrice somma** è la matrice che si ottiene sommando **posto per posto** i coefficienti di *A* con i coefficienti di *B*:

$$A = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & -1 & \sqrt{2} \end{pmatrix} \quad B = \begin{pmatrix} 0 & 4 & 1 \\ 0 & -1 & -\sqrt{2} \end{pmatrix}$$

$$A + B = \begin{pmatrix} 1 + 0 & 0 + 4 & \frac{1}{2} + 1 \\ 0 + 0 & -1 + (-1) & \sqrt{2} + (-\sqrt{2}) \end{pmatrix} = \begin{pmatrix} 1 & 4 & \frac{3}{2} \\ 0 & -2 & 0 \end{pmatrix}$$

L'addizione fra due matrici può essere fatta solo se le due matrici hanno le stesse dimensioni.

La matrice somma di due matrici $m \times n$ è una nuova matrice di dimensioni $m \times n$.

Combinazioni lineari

Dati due numeri α e β e due matrici *A* e *B*, entrambe di dimensioni $m \times n$, l'espressione

$$\alpha A \; + \; \beta B$$

indica la matrice ottenuta nel seguente modo:

1. si moltiplica la matrice *A* per il numero α
2. si moltiplica la matrice *B* per il numero β
3. si sommano le due matrici così ottenute

La matrice così ottenuta è ancora una matrice $m \times n$ e prende il nome di **combinazione lineare** delle matrici *A* e *B* con coefficienti α e β.

La combinazione lineare di due o più matrici (e altrettanti coefficienti) la si può fare soltanto se tutte le matrici della combinazione lineare hanno le stesse dimensioni.

Combinazione lineare banale

La combinazione lineare banale è una combinazione lineare di due o più matrici (e altrettanti coefficienti) in cui **tutti** i coefficienti della combinazione lineare sono uguali a 0.

$$0 \cdot M_1 + 0 \cdot M_2 + 0 \cdot M_3 + \cdots + 0 \cdot M_n$$

Il risultato di una combinazione lineare banale è sempre una **matrice nulla**.

Moltiplicazione di matrici

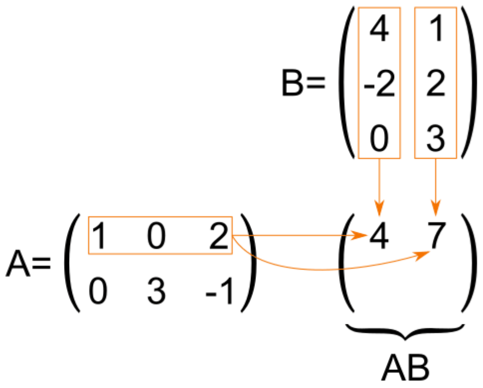
Siano *A* una matrice $m \times n$ e *B* una matrice $n \times p$, entrambe a coefficienti in $\mathbb{K}$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

Data la riga $A^{(s)}$ e la colonna $B_{(t)}$, il coefficiente c_{st} della matrice prodotto è dato da:

$$c_{st} = \sum_{r=1}^{n} a_{ir}\, b_{rj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

i è l'indice che si muove lungo la riga $A^{(s)}$, mentre *j* è l'indice che si muove lungo la colonna $B_{(t)}$.



La matrice prodotto AB è una matrice di dimensioni $m \times p$: ha lo stesso numero di righe di *A* e lo stesso numero di colonne di *B*.

La moltiplicazione fra due matrici *A* e *B* si può eseguire solo se il numero di colonne della matrice *A* coincide con il numero di righe della matrice *B*.

La moltiplicazione fra matrici generalmente non è commutativa: cambiando l'ordine dei fattori, il risultato potrebbe essere diverso.

Dipendenza ed indipendenza lineare

Data una o più matrici, si dice che queste sono **linearmente dipendenti** se esiste una loro combinazione lineare non banale il cui risultato è la matrice nulla. Se invece l'unica combinazione lineare che dà questo risultato è quella banale, allora le matrici si dicono **linearmente indipendenti**.

Per capire se due matrici sono linearmente dipendenti o linearmente indipendenti si risolve un sistema di equazioni lineari.

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

$$\alpha A + \beta B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Se *A* e *B* sono linearmente indipendenti, tutte le equazioni devono essere verificate:

$$\begin{pmatrix} 2\alpha + \beta = 0 & 2\beta = 0 & 2\alpha + 3\beta = 0 \\ 0 = 0 & -2\alpha - 3\beta = 0 & \frac{1}{2}\beta = 0 \end{pmatrix}$$

Le uniche soluzioni di queste equazioni sono $\alpha = 0$ e $\beta = 0$, quindi *A* e *B* sono linearmente indipendenti.

Se all'interno di una combinazione lineare compare la matrice nulla, allora le matrici della combinazione lineare sono linearmente dipendenti.

Dimostrazione

Siano $M(1), M(2), \dots, M(h)$ delle matrici $m \times n$ non nulle. Consideriamo poi la seguente combinazione lineare:

$$1 \cdot O_{m\times n} + 0 \cdot M(1) + M(2) + \cdots + 0 \cdot M(h)$$

Si tratta di una combinazione lineare non banale, tuttavia il suo risultato è comunque la matrice nulla $O_{m\times N}$, perché tutte le matrici non nulle sono state moltiplicate per 0.

Rango di una matrice

Si definisce **rango per righe** il massimo numero di **righe linearmente indipendenti** che è possibile estrarre da una matrice. Analogamente, si definisce **rango per colonne** il massimo numero di **colonne linearmente indipendenti** che è possibile estrarre da una matrice.

Il rango per righe ed il rango per colonne di una matrice sono uguali, ecco perché si parla direttamente di **rango della matrice**.

Il rango di una matrice può essere calcolato in diversi modi:

- metodo di Gauss-Jordan: si trasforma la matrice in una matrice a gradini: il suo rango è dato dal numero di righe non nulle e coincide con il rango della matrice di partenza;
- metodo del determinante (solo per matrici quadrate): se il determinante della matrice è diverso da 0, allora il rango è pari all'ordine della matrice. Se invece il determinante è uguale a 0, allora il rango è strettamente minore all'ordine della matrice;
- metodo dei minori: se nella matrice esiste un minore diverso da 0, allora il rango è maggiore o uguale all'ordine di quel minore;

Il rango di una matrice è pari alla dimensione del sottospazio vettoriale generato dalle sue righe.

Proprietà del rango

Se *A* è una matrice $m \times n$ a coefficienti in un campo numerico $\mathbb{K}$, allora

$$rk(A^t) = rk(A)$$

dove $rk(A^t)$ è il rango della matrice trasposta.

Dimostrazione

Segue dalla definizione di matrice trasposta:

- il rango per righe della matrice *A* è uguale al rango per colonne della matrice A^t
- il rango per colonne della matrice *A* è uguale al rango per righe della matrice A^t

Siccome il rango per righe ed il rango per colonne di una matrice sono uguali, allora il rango di *A* è uguale al rango di A^t.

Se *A* è una matrice $m \times n$ a coefficienti in un campo numerico $\mathbb{K}$ e *B* è una matrice ottenuta da *A* applicando a quest'ultima una qualunque **operazione elementare sulle righe**, allora

$$rk(A) = rk(B)$$

Dimostrazione

Il sottospazio vettoriale di $\mathbb{K}^n$ generato dalle righe di *A* coincide con il sottospazio vettoriale di $\mathbb{K}^n$ generato dalle righe di *B*, quindi anche le dimensioni di questi due sottospazi coincidono, dunque anche i ranghi delle matrici sono uguali.

Siano *A* una matrice $m \times n$ e *B* una matrice $n \times p$, entrambe a coefficienti in un campo $\mathbb{K}$. Allora

$$rk(AB) \leq \min\{rk(A),\, rk(B)\}$$

Dimostrazione

Poniamo

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ & & \vdots & \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix}$$

Per $i = 1,2, \dots, m$, l'*i*-esima riga della matrice prodotto AB è data da:

$$(AB)^{(i)} = \left(A^{(i)}B_{(1)}, A^{(i)}B_{(2)}, \dots, A^{(i)}B_{(p)}\right) =$$

$$= \left(a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1}, \dots, a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np}\right) =$$

$$= \left(a_{i1}b_{11}, \dots, a_{i1}b_{1p}\right) + \left(a_{i2}b_{21}, \dots, a_{i2}b_{2p}\right) + \cdots + \left(a_{in}b_{n1}, \dots, a_{in}b_{np}\right) =$$

$$= a_{i1}\left(b_{11}, \dots, b_{1p}\right) + a_{12}\left(b_{21}, \dots, b_{2p}\right) + \cdots + a_{in}\left(b_{n1}, \dots, b_{np}\right) =$$

$$= a_{i1}B^{(1)} + a_{i2}B^{(2)} + \cdots + a_{in}B^{(n)}$$

La riga $(AB)^{(i)}$ può essere espressa come combinazione lineare delle righe di *B*, dunque le righe di *B* formano un sistema di generatori per la riga $(AB)^{(i)}$.

Il sottospazio vettoriale generato dalle righe di AB è contenuto all'interno del sottospazio vettoriale generato dalle righe di *B*, quindi la dimensione del sottospazio vettoriale generato dalle righe di AB non può superare la dimensione del sottospazio vettoriale generato dalle righe della matrice *B*.

$$rk(AB) \leq rk(B)$$

Questa proprietà la si può enunciare anche in un altro modo equivalente:

Il rango della matrice prodotto non supera il rango del secondo fattore.

Dimostrazione

$$rk(AB) = rk((AB)^t) = rk(B^tA^t) \leq rk(A^t) = rk(A)$$

La trasposta della matrice prodotto è uguale al prodotto delle trasposte, prese in ordine inverso. Tale matrice prodotto inoltre ha un rango che non supera quello del secondo fattore (A^t in questo caso). Ma il rango di A^t è uguale al rango di *A* e dunque la proprietà è dimostrata.

Sia M una matrice $m \times n$ a coefficienti in $\mathbb{K}$. Sia H una matrice quadrata invertibile d'ordine m a coefficienti in $\mathbb{K}$ e sia J una matrice quadrata invertibile d'ordine n a coefficienti in $\mathbb{K}$. Allora:

$$rk(HM) = rk(M) \qquad rk(MJ) = rk(M)$$

Quindi moltiplicando a sinistra o a destra una matrice M per una matrice invertibile, il rango della matrice prodotto è uguale al rango di M.

Dimostrazione

Il rango della matrice prodotto non supera mai quello del secondo fattore, quindi

$$rk(HM) \leq rk(M)$$

La matrice M però può anche essere espressa come prodotto tra la matrice M e la **matrice identità**:

$$rk(M) = rk(I_m M)$$

Dato che H è invertibile, la matrice identità può essere espressa come prodotto di $H^{-1} \cdot H$:

$$rk(I_m M) = rk\big((H^{-1}H)M\big)$$

Si applica poi la proprietà associativa della moltiplicazione tra matrici:

$$rk\big((H^{-1}H)M\big) = rk\big(H^{-1}(HM)\big)$$

e di nuovo si ha che il rango della matrice prodotto non supera quello del secondo fattore, quindi:

$$rk\big(H^{-1}(HM)\big) \leq rk(HM)$$

Tutti questi passaggi si possono riassumere nella seguente catena di uguaglianze:

$$rk(HM) \leq rk(M) = rk(I_m M) = rk\big((H^{-1}H)M\big) = rk\big(H^{-1}(HM)\big) \leq rk(HM)$$

Il fatto di avere la doppia relazione:

$$rk(HM) \leq rk(M) \leq rk(HM)$$

porta a concludere che $rk(HM) = rk(M)$, quindi la proprietà è dimostrata.

Utilizzando passaggi analoghi si può dimostrare anche che vale la proprietà $rk(MJ) = rk(M)$.

Una matrice quadrata di ordine *n* è invertibile se e solo se ha rango *n*.

Dimostrazione

Parte 1: *A* matrice quadrata d'ordine *n* invertibile ⇒ *A* ha rango *n*

Dalla definizione di matrice invertibile, se *A* è invertibile significa che esiste una matrice quadrata d'ordine *n* tale che:

$$I_n = A^{-1} \cdot A = A \cdot A^{-1}$$

Quindi anche A^{-1} è invertibile a sua volta ed ha inversa *A*, quindi risulta che:

$$n = rk(I_n) = rk(A^{-1} \cdot A) = rk(A)$$

Parte 2: *A* è una matrice quadrata d'ordine *n* con $rk(A) = n$ ⇒ *A* è invertibile

Se *A* ha rango *n*, allora in *A* esistono *n* righe linearmente indipendenti, che costituiscono una base per lo spazio vettoriale $\mathbb{K}^n$.

In particolare, se si considera una matrice le cui righe sono date dalla base canonica di $\mathbb{K}^n$:

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Per ogni indice $i = 1,2, \dots, n$, esistono coefficienti $b_{i1}, b_{i2}, \dots, b_{in} \in \mathbb{K}$ tali che l'*i*-esimo vettore canonico risulti:

$$(0,0, \dots ,0,1,0, \dots ,0) = b_{i1}A^{(1)} + b_{i2}A^{(2)} + \cdots + b_{in}A^{(n)}$$

Scrivendo quindi i coefficienti $b_{i1}, b_{i2}, \dots, b_{in}$ in una matrice:

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

si ha che

$$I_n = B \cdot A$$

Analogamente, se i vettori canonici di $\mathbb{K}^n$ si pensano come colonne, allora per ogni indice $j = 1,2, \dots, n$ esistono coefficienti $c_{1j}, c_{2j}, \dots, c_{nj} \in \mathbb{K}$ tali che il *j*-esimo vettore canonico risulti:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A_{(1)}c_{1j} + A_{(2)}c_{2j} + \cdots + A_{(n)}c_{nj}$$

Scrivendo quindi i coefficienti $c_{1j}, c_{2j}, \dots, c_{nj}$ in una matrice:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

si ha che

$$I_n = A \cdot C$$

Quindi vale la seguente catena di uguaglianze:

$$B = B \cdot I_n = B \cdot (A \cdot C) = (B \cdot A) \cdot C = I_n \cdot C = C$$

da cui segue che $B = C$.

Risulta quindi che la matrice *A* è invertibile e la sua inversa è proprio la matrice $B = C$.

Pivot di una riga

Il **pivot** di una riga è il primo elemento non nullo all'interno di una data riga, considerandola da sinistra verso destra. L'indice di colonna del pivot prende il nome di **posizione pivotale**.

Esempio:

$$\begin{pmatrix} 5 & 9 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

1. il pivot della 1° riga è 5 ed ha posizione pivotale 1;
2. il pivot della 2° riga è 3 ed ha posizione pivotale 2;
3. la 3° riga non ha nessun pivot, siccome è una riga nulla

Teorema: metodo di Gauss-Jordan

> Ogni matrice può essere trasformata in una matrice a gradini mediante un'opportuna sequenza di **operazioni elementari sulle righe**.

Dimostrazione

Quando una riga nulla precede una riga nulla oppure quando una riga ha il pivot più a destra rispetto ad una delle righe che la seguono, si può sempre utilizzare l'operazione elementare di scambio fra due righe.

Se in una matrice la riga $A^{(s)}$ e la riga $A^{(t)}$ sono entrambe non nulle, entrambe hanno pivot diverso da 0 ed entrambe hanno il pivot nella stessa posizione:

$$\begin{array}{c} \\ A^{(s)} \\ \\ A^{(t)} \\ \\ \end{array} \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \alpha & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \beta & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \qquad \alpha, \beta \neq 0$$

allora si può sommare ad $A^{(t)}$ il prodotto $-\frac{\beta}{\alpha}A^{(s)}$:

$$R^{(t)} \to R^{(t)} - \frac{\beta}{\alpha}R^{(s)}$$

$$\begin{array}{c} \\ A^{(s)} \\ \\ A^{(t)} \\ \\ \end{array} \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \alpha & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & 0 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

In questo modo nella riga $A^{(t)}$ è comparso uno 0, quindi ora il pivot di $A^{(t)}$, se esiste, è più a destra rispetto al pivot della riga $A^{(s)}$.

Infine l'operazione di moltiplicazione di una riga per uno scalare diverso da 0 la si può utilizzare quando il pivot di una riga è diverso da 1: per ottenere 1 è sufficiente moltiplicare la riga per il reciproco quel pivot.

Sottomatrici

Sia *A* una matrice $m \times n$ a coefficienti nel campo numerico $\mathbb{K}$. Una **sottomatrice** $p \times q$ di *A* è una matrice costituita dagli elementi di *A* che si trovano all'incrocio di *p* righe e *q* colonne fissate di *A*.

Una sottomatrice di *A* si indica con la seguente notazione:

$$A\left(i_1 i_2 \dots i_p \middle| j_1 j_2 \dots j_q\right)$$

dove $i_1, i_2, \dots, i_p$ e $j_1, j_2, \dots, j_q$ sono rispettivamente le righe e le colonne di *A* che sono state scelte per comporre la sottomatrice.

Una matrice può essere vista come sottomatrice di sé stessa.

Esempio: se si considera la seguente matrice 4×5 a coefficienti reali

$$A = \begin{pmatrix} 5 & -2 & 1 & 0 & \sqrt{3} \\ 8 & 0 & -1 & 3 & \frac{1}{10} \\ 2 & -5 & \pi & \sqrt{2} & 6 \\ \frac{1}{2} & 1 & 1 & 0 & -5 \end{pmatrix}$$

allora la sottomatrice 2×3 di *A* corrispondente alle righe 2, 4 e alle colonne 3, 4, 5 è data da:

$$A(2\ 4|3\ 4\ 5) = \begin{pmatrix} -1 & 3 & \frac{1}{10} \\ 1 & 0 & -5 \end{pmatrix}$$

Se *B* è una sottomatrice di *A*, allora

$$rk(B) \leq rk(A)$$

Dimostrazione

Essendo *B* sottomatrice di *A*, per definizione si ha che la dimensione di *B* non supera la dimensione di *A*, quindi il rango di *B* non può superare il rango di *A*.

Il calcolo del rango di una matrice qualsiasi può sempre essere ricondotto al calcolo del rango di sottomatrici quadrate, andando a cercare, fra tutte le sottomatrici quadrate di rango massimo, quelle il cui ordine sia il più grande possibile.

Il rango di una matrice *A* è uguale al massimo tra gli ordini delle sue sottomatrici quadrate invertibili.

Minori di una matrice

SI chiama **minore di ordine *k*** della matrice *A* il determinante di una sottomatrice quadrata d'ordine *k* di *A*.

Il rango di una matrice *A* è uguale al massimo tra gli ordini dei suoi minori non nulli.

Si chiama **minore orlato** di un dato minore di ordine *k* della matrice *A* il minore di ordine $k+1$ ottenuto dal precedente minore aggiungendo una riga e una colonna.

Esempio: sia *M* la seguente matrice composta da coefficienti in campo reale:

$$M = \begin{pmatrix} -1 & 0 & \frac{1}{2} & 0 & -3 \\ \sqrt{2} & 3 & -7 & 0 & -1 \\ 1 & 0 & 0 & 0 & -2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Un possibile minore di ordine 2 è il seguente:

$$\det\big(M(1\,4|2\,5)\big) = \begin{vmatrix} 0 & -3 \\ 1 & 1 \end{vmatrix}$$

un minore orlato del precedente, ottenuto aggiungendo la riga 3 e la colonna 1, è il seguente:

$$\det\big(M(1\,3\,4|1\,2\,5)\big) = \begin{vmatrix} -1 & 0 & -3 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{vmatrix}$$

Il rango di una matrice *A* è uguale a *k* se e solo se esiste in *A* un minore d'ordine *k* diverso da 0 i cui minori orlati (se esistono) siano tutti uguali a 0.

Matrice nulla

La **matrice nulla** è una matrice di dimensioni $m \times n$ i cui coefficienti sono **tutti** uguali a 0:

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

La matrice nulla è l'**elemento neutro** dell'addizione fra matrici.

La matrice nulla si indica con le notazioni $O_{m\times n}$ e $O_{m,n}$.

Matrice a gradini

Si dice che una matrice non nulla *A* è **a gradini** se soddisfa le seguenti proprietà:

- le righe non nulle di *A* **precedono** le eventuali righe nulle;
- le **posizioni dei pivot** delle righe non nulle si dispongono in ordine **crescente** da sinistra verso destra

Esempio:

$$A = \begin{pmatrix} 5 & 9 \\ 0 & 3 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 9 \\ 9 & 0 \\ 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 0 \\ 5 & 3 \\ 0 & 9 \end{pmatrix}$$

- *A* è una matrice a gradini;
- *B* non è una matrice a gradini, perché il pivot della 2° riga ha la stessa posizione (1) del pivot nella 1° riga;
- *C* non è una matrice a gradini, perché una riga nulla precede le righe non nulle

Le righe non nulle di una matrice a gradini sono linearmente indipendenti, dunque il rango di una matrice a gradini è dato dal numero di righe non nulle.

Matrice a gradini ridotta

Una matrice a gradini si dice **ridotta** se:

- i pivot di tutte le righe non nulle sono uguali ad 1;
- ciascun pivot è l'unico elemento non nullo all'interno della propria colonna (cioè se né sopra né sotto al pivot compaiono altri coefficienti diversi da 0)

Una qualunque matrice si può trasformare in una matrice a gradini ridotta mediante una opportuna sequenza di operazioni elementari sulle righe.

Le righe non nulle di una matrice a gradini *G* sono linearmente indipendenti.

Dimostrazione

Supponiamo che la matrice *G* abbia *r* righe non nulle e quindi *r* pivot. Gli indici di colonna $j_1, j_2, \dots, j_r$ rappresentano le posizioni pivotali.

$$\alpha_1 G^{(1)} + \alpha_2 G^{(2)} + \cdots + \alpha_r G^{(r)} = (0,0,\dots,0)$$

Se le *r* righe di *G* sono linearmente indipendenti, allora i coefficienti $\alpha_1, \alpha_2, \dots, \alpha_r$ devono essere tutti uguali a 0.

Per ogni riga $G^{(1)}, G^{(2)}, \dots, G^{(r)}$ consideriamo un generico coefficiente con posizione di colonna *j*:

$$\alpha_1 g_{1j} + \alpha_2 g_{2j} + \cdots + \alpha_r g_{rj} = 0 \quad (\star)$$

Ora specializziamo questa relazione in ciascuna delle posizioni pivotali $j_1, j_2, \dots, j_r$:

$$\alpha_1 g_{1j_1} + \alpha_2 g_{2j_1} + \cdots + \alpha_r g_{rj_1} = 0$$

Siccome *G* è una matrice a gradini, sotto al pivot g_{1j_1} compaiono solo degli zeri, per cui $g_{2j_1}, g_{3j_1}, \dots, g_{rj_1}$ sono tutti uguali a 0. Dunque la precedente relazione diventa:

$$\alpha_1 g_{1j_1} = 0$$

g_{1j_1} è un pivot, quindi è diverso da 0, per cui segue che $\alpha_1 = 0$.

Ora specializziamo la relazione $(\star)$ con la posizione pivotale j_2:

$$\alpha_1 g_{1j_2} + \alpha_2 g_{2j_2} + \cdots + \alpha_r g_{rj_2} = 0$$

g_{2j_2} è un pivot, quindi i coefficienti sotto di lui sono tutti uguali a 0. Anche $\alpha_1 = 0$, quindi la relazione diventa:

$$\alpha_2 g_{2j_2} = 0$$

g_{2j_2} è un pivot, quindi è diverso da 0, per cui segue che $\alpha_2 = 0$.

Procedendo in maniera induttiva fino alla posizione pivotale j_r si dimostra che tutti i coefficienti $\alpha_1, \alpha_2, \dots, \alpha_r$ sono uguali a 0, dunque le righe $G^{(1)}, G^{(2)}, \dots, G^{(r)}$ sono linearmente indipendenti.

Matrice trasposta

La trasposta di una matrice assegnata è la matrice che si ottiene andando a scambiare le righe con le colonne:

$$A = \begin{pmatrix} 1 & 5 \\ 6 & 9 \end{pmatrix} \rightarrow A^t = \begin{pmatrix} 1 & 6 \\ 5 & 9 \end{pmatrix}$$

Matrici invertibili

Una matrice quadrata *A* d'ordine *n* si dice **invertibile** se esiste una matrice quadrata *L* d'ordine *n* per la quale risulti

$$AL = LA = I_n$$

Se la matrice *L* esiste, allora è anche unica.

La matrice *L* prende il nome di **matrice inversa** di *A* e si indica con A^{-1}.

La matrice identica d'ordine *n* è invertibile e la sua inversa coincide con la matrice stessa:

$$I_n^{-1} = I_n$$

Se *A* è una matrice invertibile di ordine *n* ed A^{-1} è la sua inversa, allora a sua volta anche A^{-1} è invertibile e si ha che

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

Se *A* e *B* sono due matrici quadrate invertibili d'ordine *n*, allora anche il prodotto AB è una matrice quadrata invertibile d'ordine *n* e risulta

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

L'insieme di tutte le matrici quadrate d'ordine *n* invertibili a coefficienti nel campo numerico $\mathbb{K}$ viene indicato con $GL_n(\mathbb{K})$.

Come conseguenza del teorema di Binet si ha che se una matrice è invertibile, allora il suo determinante è diverso da 0.

Matrice diagonale

Una **matrice diagonale** è una **matrice quadrata** in cui gli elementi che non appartengono alla diagonale principale sono tutti nulli.

$$A \in M_{n,n}(\mathbb{K}) \text{ matrice diagonale } \Leftrightarrow a_{ij} = 0 \; \forall i \neq j$$

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Non c'è nessun vincolo invece sui coefficienti che appartengono alla diagonale principale: possono essere tutti nulli, qualcuno nullo e qualcuno no, oppure tutti non nulli.

In una matrice diagonale:

- il rango è dato dal numero di righe non nulle (infatti una matrice diagonale è una matrice a gradini);
- il determinante è dato dal prodotto dei coefficienti che stanno sulla diagonale principale

Se i coefficienti che compaiono sulla diagonale principale sono tutti uguali ad 1, allora si tratta della **matrice identità**.

Matrici simili

Due matrici quadrate *A* e *B* si dicono **simili** se e solo se esiste una **matrice invertibile** *P* tale che:

$$A = P^{-1} \cdot B \cdot P$$

La relazione di similitudine è:

- **riflessiva**: ogni matrice è simile a sé stessa – la matrice che va a soddisfare questa similitudine è la **matrice identità**;
- **simmetrica**: se *A* è simile a *B*, allora *B* è simile ad *A*;
- **transitiva**: se *A* è simile a *B* e *B* è simile a *C*, allora *A* è simile a *C*

Due o più matrici simili:

- hanno lo stesso determinante e lo stesso rango;
- hanno lo stesso polinomio caratteristico

Matrice triangolare

Una matrice triangolare è una matrice quadrata in cui tutti gli elementi sopra (matrice triangolare **superiore**) o al di sotto (matrice triangolare **inferiore**) della diagonale principale sono nulli.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 7 & 0 \\ 1 & 4 & -2 \end{pmatrix} \quad B = \begin{pmatrix} 6 & 1 & 4 \\ 0 & -15 & 9 \\ 0 & 0 & 3 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

In questo caso *A* è una matrice triangolare superiore, *B* è una matrice triangolare inferiore e *C* è una matrice triangolare sia superiore che inferiore, ovvero una matrice diagonale.

Le matrici triangolari superiori/inferiori formano uno spazio vettoriale, infatti:

- somma e prodotto di matrici triangolari dello stesso tipo formano una matrice triangolare;
- l'insieme delle matrici triangolari contiene la matrice nulla $n \times n$

In una matrice triangolare:

- il rango è pari al numero di coefficienti non nulli della diagonale principale;
- il determinante è dato dal prodotto degli elementi della diagonale principale;
- gli autovalori di una matrice triangolare sono gli elementi della diagonale principale

Sistemi di equazioni lineari

Un sistema di equazioni lineari è una lista finita di equazioni lineari nelle stesse incognite.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Una soluzione del sistema è una soluzione simultanea di tutte le sue equazioni.

Si dice che un sistema è **compatibile** se ammette **almeno una soluzione**, altrimenti si dice **incompatibile**. Se in un sistema di equazioni lineari compare un'equazione incompatibile, allora tutto il sistema è incompatibile.

Due sistemi di equazioni lineari si dicono **equivalenti** se ammettono le stesse soluzioni oppure se sono entrambi incompatibili.

Un sistema di equazioni lineari può essere scritto anche in **forma matriciale**

$$AX = b$$

dove *A* è la matrice dei coefficienti delle incognite (detta anche **matrice incompleta** del sistema):

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

X è la **matrice colonna** di dimensioni $n \times 1$ delle incognite:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

e *b* è la matrice colonna di dimensioni $m \times 1$ dei termini noti:

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

La matrice B, di dimensioni $m \times (n+1)$, definita come:

$$B = (A|b) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ & & \vdots & & | & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_n \end{pmatrix}$$

che si ottiene aggiungendo alla matrice A la colonna dei termini noti, prende il nome di **matrice completa del sistema**. Ciascuna riga di questa matrice corrisponde ad una equazione del sistema.

Teorema di Rouché-Capelli

Un sistema di equazioni lineari è compatibile se e solo se la matrice incompleta e la matrice completa del sistema hanno lo stesso rango.

Grazie al teorema di Rouché-Capelli si può dire che se il pivot di una riga cade nella posizione dei termini noti, allora il sistema è incompatibile.

Numero di soluzioni di un sistema di equazioni lineari

Se un sistema è compatibile, il numero di soluzioni è dato da:

$$\infty^{n-rk(M)}$$

dove n è il numero di variabili del sistema, mentre $rk(M)$ è il rango della **matrice completa** associata al sistema. Equivalentemente, $rk(M)$ può essere visto anche come il **numero di variabili pivotali**.

Il valore $n - rk(M)$ è il **numero di variabili libere** del sistema, quindi si può anche dire che il numero di soluzioni del sistema è

$$\infty^{\text{nr. variabili libere}}$$

Se $n - rk(M) = 0$, allora il sistema ammette una sola soluzione.

Risoluzione di un sistema di equazioni lineari

Uno dei metodi con cui si può risolvere un sistema di equazioni lineari è quello di Gauss-Jordan. Partendo dal sistema, questo metodo consiste nei seguenti passi:

1. si scrive la **matrice completa** del sistema;
2. si trasforma la matrice completa del sistema in una **matrice a gradini ridotta**;
3. all'interno di questa matrice si identificano le **variabili pivotali** e le **variabili libere**;
4. si scrive il sistema associato alla matrice ottenuta al punto precedente in cui le variabili pivotali vengono espresse in funzione delle variabili libere (se presenti) e dei termini noti

Sistemi di equazioni lineari parametrici

Un sistema di equazioni lineari si dice **parametrico** se i suoi coefficienti sono delle espressioni che dipendono da uno o più parametri.

Un sistema parametrico non esprime in realtà un singolo sistema, ma ne esprime uno diverso per ogni valore dei suoi parametri. Risolvere un sistema parametrico, quindi, significa trovare per quali valori dei parametri il sistema è compatibile e per ciascuno di tali valori trovare le soluzioni del relativo sistema.

Tipicamente avviene che per tutti i valori del parametro il sistema si comporta allo stesso modo, ad eccezione di un numero finito di casi. Questi valori eccezionali vengono trattati a parte, assegnando al parametro uno di questi valori eccezionali e analizzando il sistema corrispondente.

Sistema omogeneo di equazioni lineari

Un sistema di equazioni lineari si dice **omogeneo** se è composto da equazioni lineari omogenee, ovvero se il termine noto di ogni equazione è uguale a 0.

Un sistema omogeneo ammette sempre almeno una soluzione, che viene chiamata **soluzione banale**, ed è quella soluzione in cui tutti i coefficienti delle incognite sono uguali a 0.

Le soluzioni di un sistema omogeneo di equazioni lineari a coefficienti in un campo $\mathbb{K}$ formano un **sottospazio vettoriale** di $\mathbb{K}^n$. La dimensione di tale sottospazio vettoriale è pari al numero di soluzioni del sistema.

Sistema di Cramer

Un sistema di Cramer è un sistema di equazioni lineari dove la **matrice incompleta** è una **matrice quadrata** ed è **invertibile**.

Un sistema di Cramer è sempre compatibile ed ha sempre una ed una sola soluzione.

Per risolvere un sistema di Cramer si possono utilizzare 2 metodi:

- metodo dell'inversa;
- regola di Cramer

Metodo dell'inversa

La soluzione $(x_1, x_2, \dots, x_m)$ si trova moltiplicando la matrice inversa della matrice incompleta per la colonna dei termini noti:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

Regola di Cramer

Per ogni colonna $C_{(1)}, C_{(2)}, \dots, C_{(n)}$ della matrice incompleta, il corrispondente valore x_j appartenente alla soluzione si ottiene utilizzando la formula

$$x_j = \frac{det(A_j)}{det(A)}$$

dove A_j è la matrice che si ottiene andando a sostituire la *j*-esima colonna della matrice *A* (la matrice incompleta) con la colonna dei termini noti.

Spazi vettoriali

Sia $\mathbb{K}$ un campo numerico. Gli elementi di $\mathbb{K}$ vengono chiamati numeri oppure **scalari**.

Uno **spazio vettoriale** sul campo $\mathbb{K}$ è un insieme non vuoto **V**, i cui elementi vengono chiamati **vettori**, dotato di due operazioni:

- addizione vettoriale
- moltiplicazione di uno scalare per un vettore

Tali operazioni devono soddisfare un certo insieme di proprietà per poter essere definite tali.

Addizione vettoriale

Per ogni scelta di vettori $u, v \in V$ è definito un vettore $u + v \in V$ (univocamente determinato) detto la **somma vettoriale** di **u** con **v**.

La somma di due vettori è a sua volta un vettore.

L'addizione vettoriale deve soddisfare le seguenti proprietà:

1. proprietà associativa:
$$\forall u, v, w \in V,\ (u + v) + w = u + (v + w)$$
2. proprietà commutativa:
$$\forall u, v \in V,\ \ u + v = v + u$$
3. esistenza del vettore nullo (indicato con **0**):
$$\exists 0 \in V : \forall v \in V \quad 0 + v = v + 0 = v$$
4. esistenza del vettore opposto (indicato con $-v$):
$$\forall v \in V,\ \exists(-v) \in V : v + (-v) = (-v) + v = 0$$

Moltiplicazione di uno scalare per un vettore

Per ogni scelta di $\alpha \in \mathbb{K}$ e per ogni scelta di un vettore $v \in V$, è definito un vettore $\alpha v \in V$ (univocamente determinato) detto il **prodotto** di α con **v**.

La moltiplicazione di uno scalare per un vettore dà luogo ad un altro vettore.

La moltiplicazione di uno scalare per un vettore deve soddisfare le seguenti proprietà:

1. proprietà distributiva rispetto alla somma di vettori:
$$\forall \alpha \in \mathbb{K}, \forall u, v \in V \quad \alpha(u + v) = (\alpha \cdot u) + (\alpha \cdot v)$$
2. proprietà distributiva rispetto alla somma di scalari:
$$\forall \alpha, \beta \in \mathbb{K}, \forall v \in V \quad (\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$$
3. proprietà associativa:
$$\forall \alpha,\ \beta \in \mathbb{K},\ \forall v \in V \quad (\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$$
4. moltiplicazione unitaria:
$$\exists 1 \in \mathbb{K} : \forall v \in V \quad 1 \cdot v = v$$

Proprietà degli spazi vettoriali

- se $v_1, v_2, \dots, v_s$ è una sequenza di vettori dello spazio vettoriale **V**, allora l'espressione
$$v_1 + v_2 + \cdots + v_s$$
risulta univocamente determinata, come conseguenza della proprietà associativa dell'addizione vettoriale;
- nello spazio vettoriale **V** vi è un solo vettore nullo: se per assurdo esistesse un altro vettore nullo **0'**, si avrebbe $0' = 0' + 0 = 0$.
- per ogni vettore $v \in V$, esiste un unico vettore opposto $-v$. Se **w** fosse un altro vettore opposto di **v**, allora:
$$w = w + 0 = w + [v + (-v)] = [w + v] + (-v) = 0 + (-v) = -v$$
- per ciascun vettore $v \in V$ si ha che $0 \cdot v = 0$.
$$0 \cdot v = (0 + 0) \cdot v = (0 \cdot v) + (0 \cdot v) \Leftrightarrow 0 = 0 \cdot v$$
- per ciascuno scalare α si ha che $\alpha \cdot 0 = 0$.
$$\alpha \cdot 0 = \alpha \cdot (0 + 0) = (\alpha \cdot 0) + (\alpha \cdot 0) \Leftrightarrow 0 = \alpha \cdot 0$$
- per ciascuno scalare α e vettore **v** si ha che:
$$(-\alpha) \cdot v = -(\alpha \cdot v) = \alpha \cdot (-v)$$

Combinazioni lineari di vettori

Data la sequenza di vettori $v_1, v_2, \dots, v_s \in V$ e la sequenza di scalari $\alpha_1, \alpha_2, \dots, \alpha_s \in \mathbb{K}$, risulta univocamente determinato il vettore

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_s v_s$$

e tale operazione prende il nome di **combinazione lineare** dei vettori $v_1, v_2, \dots, v_s$ con i coefficienti $\alpha_1, \alpha_2, \dots, \alpha_s$.

Spazio vettoriale delle matrici $m \times n$

L'insieme di tutte le matrici $m \times n$ su un campo numerico $\mathbb{K}$, indicato con $M_{m,n}(\mathbb{K})$, forma uno spazio vettoriale:

- l'addizione vettoriale è rappresentata dalla somma di due matrici;
- la moltiplicazione di uno scalare per un vettore è rappresentata dalla moltiplicazione di uno scalare per una matrice;
- $M_{m,n}(\mathbb{K})$ contiene il vettore nullo, ovvero la matrice nulla di dimensioni $m \times n$

Nello spazio vettoriale $M_{m,n}(\mathbb{K})$ vi sono due casi particolari:

- $M_{1,n}(\mathbb{K}) = \{(\alpha_1, \alpha_2, \dots, \alpha_n): \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}\}$
- $M_{n,1}(\mathbb{K}) = \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} : \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K} \right\}$

Nel primo caso i vettori sono righe, mentre nel secondo sono colonne. In entrambi i casi si tratta di sequenze finite di elementi $\alpha_1, \alpha_2, \dots, \alpha_n$ del campo $\mathbb{K}$. Entrambi gli spazi vettoriali si indicano con la scrittura $\mathbb{K}^n$: questo spazio vettoriale prende il nome di **spazio vettoriale numerico standard**.

Sottospazi vettoriali

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$ e sia **U** un sottoinsieme non vuoto di **V**. Si dice che **U** è un **sottospazio vettoriale** di **V** se per **U** valgono le proprietà degli spazi vettoriali e se risulta chiuso rispetto alle combinazioni lineari dei propri elementi, ovvero:

1. per ogni $u, u' \in U,\ u + u' \in U$;
2. per ogni $\alpha \in \mathbb{K}$ e per ogni $u \in U$, $\alpha u \in U$

Qualsiasi spazio vettoriale **V** ammette sempre almeno 2 sottospazi vettoriali:

1. il sottospazio vettoriale costituito dal solo vettore nullo;
2. il sottospazio vettoriale costituito da **V** stesso

Se **U** è un sottospazio vettoriale di **V**, allora **U** contiene il vettore nullo.

Dimostrazione

Per definizione di sottospazio vettoriale, **U** non è vuoto, quindi contiene almeno un vettore **u**. Per la proprietà 2) **U** contiene anche tutti i multipli di ogni suo vettore, quindi contiene anche il vettore $0u$, ovvero il vettore nullo.

Se **U** è un sottospazio vettoriale di **V**, allora **U** risulta chiuso rispetto alle combinazioni lineari dei propri elementi.

Dimostrazione

Se $u_1 \in U$ e $\alpha_1 \in \mathbb{K}$, allora per la proprietà 2) dei sottospazi vettoriali si ha che anche $\alpha_1 u_1 \in U$.

Se $u_1, u_2 \in U$ e $\alpha_1, \alpha_2 \in \mathbb{K}$, allora per la proprietà 2) dei sottospazi vettoriali $\alpha_1 u_1 \in U$ e $\alpha_2 u_2 \in U$, quindi per la proprietà 1) dei sottospazi vettoriali, anche $\alpha_1 u_1 + \alpha_2 u_2 \in U$, quindi la tesi è dimostrata.

Se **V** è uno spazio vettoriale sul campo numerico $\mathbb{K}$ e **U** è un sottoinsieme di **V** contenente il vettore nullo e che risulta chiuso rispetto alle combinazioni lineari dei suoi elementi, allora **U** è un sottospazio vettoriale di **V**.

Dimostrazione

Si tratta di dimostrare che per **U** valgono le proprietà 1) e 2) dei sottospazi vettoriali.

Si possono considerare quindi le seguenti combinazioni lineari:

$$(1) \quad 1 \cdot u_1 + 1 \cdot u_2 \quad u_1, u_2 \in U$$

$$(2) \quad \alpha \cdot u + 1 \cdot 0 \quad \alpha \in \mathbb{K}, u \in U$$

Per ipotesi **U** è chiuso rispetto alle combinazioni lineari dei propri elementi, dunque entrambi questi vettori devono appartenere ad **U**, dunque la tesi è dimostrata, perché:

- la combinazione lineare (1) dimostra la proprietà 1) dei sottospazi vettoriali;
- la combinazione lineare (2) dimostra la proprietà 2) dei sottospazi vettoriali

Intersezione di sottospazi vettoriali

Dati due sottospazi vettoriali **U** e **W** di uno spazio vettoriale **V**, la loro intersezione $U \cap W$ è un sottospazio vettoriale.

Dimostrazione

Il vettore nullo è certamente contenuto in $U \cap W$, in quanto appartiene sia ad **U** che a **W**.

Se **v** e **v'** sono due vettori di $U \cap W$ ed α, α' sono due numeri, allora la combinazione lineare

$$\alpha v + \alpha' v'$$

appartiene sia ad **U** che a **W**, in quanto risultano chiusi rispetto alle combinazioni lineari dei propri elementi.

In generale, l'intersezione di *n* sottospazi vettoriali di **V** (con *n* che può anche essere infinito) è a sua volta un sottospazio vettoriale di **V**. Questo è il motivo per cui l'insieme delle soluzioni di un sistema omogeneo di equazioni lineari è uno spazio vettoriale: ogni soluzione di ogni equazione del sistema è un sottospazio vettoriale di $\mathbb{K}^n$, le soluzioni del sistema sono date dall'intersezione delle soluzioni delle singole equazioni, quindi le soluzioni del sistema sono un'intersezione di sottospazi vettoriali di $\mathbb{K}^n$, dunque sono a loro volta un sottospazio vettoriale di $\mathbb{K}^n$.

Unione di sottospazi vettoriali

A differenza dell'intersezione, non sempre l'unione di due sottospazi vettoriali di *V* dà luogo ad un nuovo sottospazio vettoriale di *V*.

Esempio:

$$U = \{(x, y) \in \mathbb{R}^2 : x = 0\}$$

$$W = \{(x, y) \in \mathbb{R}^2 : y = 0\}$$

Sia *U* che *W* risultano sottospazi vettoriali di $\mathbb{R}^3$.

Si considerino ora due vettori $u \in W$ e $w \in W$ e la loro somma:

$$u = (0,1) \quad w = (1,0) \quad u + w = (1,1)$$

Si ha che $u \in U$ e $w \in W$, quindi $u, w \in U \cup W$. Tuttavia la loro somma $u + w$ non appartiene ad $U \cup W$, perché il vettore $(1,1)$ non appartiene né ad **U** né a **W**. Dunque $U \cup W$ è un insieme di vettori che non risulta chiuso rispetto alle combinazioni lineari dei propri elementi, pertanto $U \cup W$ NON è un sottospazio vettoriale di **V**.

Sottospazio somma

Siano **U** e **W** sottospazi vettoriali di **V**. Il **sottospazio somma** è definito come:

$$U + W = \{v \in V \colon \exists u \in U, w \in W : \ v = u + w\}$$

Il sottospazio somma è generato dall'unione di **U** con **W**:

$$U + W = L(U \cup W)$$

Sistema di generatori

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$ e sia *S* un sottoinsieme non vuoto di **V**.

Sia ora *F* la famiglia di sottospazi vettoriali di **V** definita come:

$$F = \{U : \text{U è un sottospazio vettoriale di V}, \ S \subseteq U\}$$

Dunque *F* è un insieme formato da tutti i sottospazi vettoriali di **V** che contengono l'insieme *S*.

L'insieme *F* non è vuoto, perché contiene almeno lo spazio vettoriale **V** stesso, dato che $S \subseteq V$.

Sia ora $L(S)$ l'intersezione di tutti i sottospazi vettoriali **U** che appartengono alla famiglia *F*:

$$L(S) = \bigcap_{U \in F} U$$

L'insieme $L(S)$ prende il nome di **sistema di generatori** per **V** ed è l'intersezione di tutti i sottospazi vettoriali $U \in V$ che contengono l'insieme *S*.

> Per ogni insieme non vuoto *S* di vettori di **V**, valgono le seguenti proprietà:
> 1. $L(S)$ è un sottospazio vettoriale di **V** e risulta $S \subseteq L(S)$;
> 2. se **W** è un sottospazio vettoriale di **V** tale che $S \subseteq W$, allora risulta anche $L(S) \subseteq W$

Queste proprietà caratterizzano l'insieme $L(S)$ come il più piccolo sottospazio vettoriale di **V** che contiene l'insieme *S*.

Dimostrazione

Parte 1: $L(S)$ è un sottospazio vettoriale

Per ipotesi si ha che $S \subseteq L(S)$, quindi siccome *S* non è vuoto, allora nemmeno $L(S)$ è vuoto.

Siano ora u, u' due vettori di un qualsiasi sottospazio vettoriale $U \in L(S)$ e siano α, α' due scalari. Siccome **U** è un sottospazio vettoriale, esso è chiuso rispetto alle combinazioni lineari dei propri elementi, dunque $\alpha u + \alpha' u' \in U$. Dato che questo vale per ogni sottospazio vettoriale **U**, allora il vettore $\alpha u + \alpha' u'$ è anche nell'intersezione di tutti i sottospazi vettoriali **U** che contengono *S*, quindi $\alpha u + \alpha' u \in L(S)$.

Parte 2: $S \subseteq W \Rightarrow L(S) \subseteq W$, con **W** sottospazio vettoriale di **V**

Per ipotesi $S \subseteq W \subseteq V$, quindi **W** appartiene alla famiglia *F* dei sottospazi vettoriali di **V** che contengono l'insieme *S*, la cui intersezione dà luogo esattamente ad $L(S)$, quindi $L(S) \subseteq W$.

Base di un sottospazio vettoriale

Sia **V** uno spazio vettoriale sul campo $\mathbb{K}$. Un sottoinsieme finito di vettori di **V**:

$$S = \{v_1, v_2, \dots, v_h\} \subseteq V$$

è una **base** di **V** se S è un sistema di generatori per **V** e se i vettori $v_1, v_2, \dots, v_h$ sono **linearmente indipendenti**.

Componenti di un vettore rispetto ad una base

Sia **V** uno spazio vettoriale *n*-dimensionale sul campo numerico $\mathbb{K}$ e sia $B = \{v_1, v_2, \dots, v_n\}$ una base di **V**. Poiché *B* è un sistema di generatori per **V**, ogni vettore $v \in V$ può essere espresso come combinazione lineare dei vettori di *B*:

$$v = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$$

Siccome i vettori di *B* sono linearmente indipendenti, coefficienti $x_1, x_2, \dots, x_n$ sono univocamente determinati e prendono il nome di **componenti** del vettore **v** rispetto alla base *B*.

Per determinare le componenti di un vettore rispetto ad una base occorre risolvere un sistema di equazioni lineari.

Estrazione di una base da un sistema di generatori

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$. Supponiamo che **V** non si riduca al solo vettore nullo e che sia finitamente generato, dunque sia $S = \{v_1, v_2, \dots, v_s\}$ un sistema finito di generatori per **V**.
Allora esiste una base *B* di **V** tale che risulti $B \subseteq S$.

Per estrarre una base da un sistema di generatori ci sono 2 algoritmi:

- algoritmo del sistema
- algoritmo della matrice

Algoritmo del sistema

1. si inizializza l'insieme *B* a vuoto;
2. per ogni vettore $v_i \in S$:
    - se $v_i \neq 0$ e $B = \emptyset$, allora si aggiunge il vettore all'insieme *B*
    - se $v_i \neq 0$ e $B \neq \emptyset$, allora v_i viene aggiunto all'insieme *B* solo se questo è linearmente indipendente rispetto ai vettori già presenti in *B*. Per determinare se v_i è linearmente indipendente dagli altri, cioè se non è esprimibile come combinazione lineare dei vettori presenti nell'insieme *B*, si risolve un sistema.

Algoritmo della matrice

Questo algoritmo si basa sul fatto che modificare la lista dei vettori del sistema di generatori tramite una sequenza di una o più operazioni analoghe alle operazioni elementari sulle righe non comporta modifiche dello spazio vettoriale generato. Quindi è possibile:

- scrivere i vettori del sistema di generatori come righe di una matrice;
- trasformare la matrice in una matrice a gradini mediante una sequenza di operazioni elementari sulle righe;
- le righe non nulle della matrice formano una base per lo spazio vettoriale

Teorema del completamento ad una base

Dato uno spazio vettoriale **V**, una base $E = \{v_1, v_2, \dots, v_n\}$ di **V** ed un insieme $S = \{w_1, w_2, \dots, w_k\}$ di vettori linearmente indipendenti di **V**, è possibile completare la base *E* aggiungendo ad essa opportuni vettori dell'insieme *S*, scelti in modo tale che tutti i vettori della base siano linearmente indipendenti.

Gli algoritmi per il completamento ad una base sono gli stessi utilizzabili per l'estrazione di una base.

Vettori linearmente dipendenti ed indipendenti

Se tra i vettori $v_1, v_2, \dots, v_h$ vi è il vettore nullo, allora i vettori della lista sono linearmente dipendenti.

Dimostrazione

Analoga a quanto visto per le matrici: supponiamo che $v_1 = 0$ e consideriamo la combinazione lineare

$$1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_h$$

Non è una combinazione lineare banale, tuttavia il suo risultato è comunque il vettore nullo, quindi la proprietà è dimostrata.

Un singolo vettore **v** è linearmente indipendente se e solo se $v \neq 0$.

Dimostrazione

Se $v = 0$, allora:

$$\alpha v = 0 \quad \forall \alpha \in \mathbb{K}$$

dunque se $\alpha \neq 0$, la combinazione lineare non è banale ma dà comunque luogo al vettore nullo.

Se invece $v \neq 0$, allora l'unico coefficiente $\alpha \in \mathbb{K}$ tale per cui

$$\alpha v = 0$$

è proprio 0, dunque **v** è linearmente indipendente.

Due vettori v_1, v_2 sono linearmente dipendenti se e solo se uno dei due risulta multiplo dell'altro.

Dimostrazione

Supponiamo che

$$\alpha_1 v_1 + \alpha_2 v_2 = 0$$

sia una combinazione lineare non banale. In particolare, supponiamo che $\alpha_1 \neq 0$. Allora si possono dividere entrambi i membri dell'equazione per α_1:

$$v_1 + \frac{\alpha_2}{\alpha_1} v_2 = 0$$

ovvero

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2$$

I vettori $v_1, v_2, \dots, v_h$ (con $h \geq 2$) dello spazio vettoriale **V** sono linearmente dipendenti se e solo se almeno uno di essi può essere espresso come combinazione lineare dei rimanenti.

Dimostrazione

Supponiamo che

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_h v_h = 0$$

sia una combinazione lineare non banale. Allora esiste un coefficiente $\alpha_j \neq 0$. Se si portano gli altri termini a secondo membro, l'uguaglianza diventa:

$$\alpha_j v_j = -\alpha_1 v_1 - \alpha_2 v_2 - \cdots - \alpha_h v_h$$

e dividendo entrambi i membri per α_j si ottiene:

$$v_j = -\frac{\alpha_1}{\alpha_j} v_1 - \frac{\alpha_2}{\alpha_j} v_2 - \cdots - \frac{\alpha_h}{\alpha_j} v_h$$

dunque v_j è stato espresso come combinazione lineare dei rimanenti, quindi il teorema è dimostrato: se in una combinazione lineare non banale, che dà come risultato il vettore nullo, uno dei vettori si può scrivere come combinazione lineare dei rimanenti, allora i vettori sono linearmente dipendenti.

L'implicazione vale anche nell'altro verso: se tra i vettori $v_1, v_2, \dots, v_h$ ce n'è almeno uno che si può esprimere come combinazione lineare dei rimanenti, allora i vettori sono linearmente dipendenti. Per dimostrarlo è sufficiente partire dall'uguaglianza precedente e portare a destra il termine v_j.

Se i vettori $v_1, v_2, \dots, v_k$ sono linearmente dipendenti, allora lo sono anche i vettori $v_1, v_2, \dots, v_h$ (con $1 \leq k < h$).

In altre parole, se i vettori di una lista sono linearmente dipendenti, aggiungendo altri vettori a questa lista si avrà comunque una lista di vettori linearmente dipendenti.

Dimostrazione

Per ipotesi, i vettori $v_1, v_2, \dots, v_k$ sono linearmente dipendenti, quindi deve esistere una combinazione lineare non banale di questo tipo:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

Se a questa combinazione lineare si aggiunge $h - k$ volte il vettore nullo:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k + 0 v_{k+1} + \cdots + 0 v_h = 0$$

si tratta ancora di una combinazione lineare non banale, ma che dà comunque luogo al vettore nullo. Quindi la proprietà è dimostrata.

Se i vettori $v_1, v_2, \dots, v_h$ sono linearmente indipendenti, allora lo sono anche i vettori $v_1, v_2, \dots, v_k$ (con $1 \leq k < h$).

In altre parole, rimuovendo dei vettori da una lista di vettori linearmente indipendenti, si avrà comunque una nuova lista di vettori linearmente indipendenti.

Dimostrazione

Supponiamo per assurdo che i vettori della lista corta $v_1, v_2, \dots, v_k$ siano linearmente dipendenti. Allora anche i vettori della lista lunga $v_1, v_2, \dots, v_h$ dovrebbero essere linearmente dipendenti. Ma questa è una contraddizione dell'ipotesi, quindi la proprietà è dimostrata.

Sia $v_1, v_2, \dots, v_h$ una lista di vettori linearmente indipendenti dello spazio vettoriale **V** e siano $\alpha_1, \alpha_2, \dots, \alpha_h, \beta_1, \beta_2, \dots, \beta_h$ scalari del campo $\mathbb{K}$. Se risulta:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_h v_h = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_h v_h$$

allora si ha che:

$$\alpha_1 = \beta_1 \quad \alpha_2 = \beta_2 \quad \dots \quad \alpha_h = \beta_h$$

In altre parole, se un vettore può essere rappresentato come combinazione lineare di vettori linearmente indipendenti, allora tale combinazione lineare è unica.

Dimostrazione

Si porta tutto a primo membro e si raccolgono i coefficienti dei vettori:

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_h - \beta_h)v_h = 0$$

Per ipotesi i vettori $v_1, v_2, \dots, v_h$ sono linearmente indipendenti, quindi l'unica loro combinazione lineare che dà luogo al vettore nullo è quella banale, quindi si deve avere che:

$$\alpha_1 - \beta_1 = 0 \quad \alpha_2 - \beta_2 = 0 \quad \dots \quad \alpha_h - \beta_h = 0$$

ovvero:

$$\alpha_1 = \beta_1 \quad \alpha_2 = \beta_2 \quad \dots \quad \alpha_h = \beta_h$$

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$.
Supponiamo che **V** sia finitamente generato e che $S = \{v_1, v_2, \dots, v_n\}$ sia un sistema di generatori per **V**.

Allora, se $w_1, w_2, \dots, w_m$ sono vettori di **V** con $m > n$, i vettori $w_1, w_2, \dots, w_m$ sono linearmente dipendenti.

Dimostrazione

Siccome $m > n$, conviene trattare la lista $w_1, w_2, \dots, w_m$ come l'unione di due liste:

$$w_1, w_2, \dots, w_n \qquad w_{n+1}, w_{n+2}, \dots, w_m$$

Se i vettori della sotto-lista $w_1, w_2, \dots, w_n$ sono linearmente dipendenti, allora anche i vettori dell'intera lista $w_1, w_2, \dots, w_m$ sono linearmente dipendenti.

Se invece i vettori della lista $w_1, w_2, \dots, w_n$ sono linearmente indipendenti, allora per dimostrare che i vettori $w_1, w_2, \dots, w_m$ sono linearmente dipendenti occorre dimostrare che i vettori $w_1, w_2, \dots, w_n$ formano un sistema di generatori per **V**: se così fosse i vettori $w_{n+1}, w_{n+2}, \dots, w_m$ si potrebbero esprimere come combinazione lineare dei vettori $w_1, w_2, \dots, w_n$ e dunque i vettori $w_1, w_2, \dots, w_m$ sarebbero linearmente dipendenti.

Sottospazio vettoriale generato da un insieme finito di vettori

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$. Sia $w_1, w_2, \dots, w_s$ una lista finita di vettori di **V** e sia **W** il sottospazio vettoriale di **V** generato dai vettori di questa lista:

$$W = L(w_1, w_2, \dots, w_s)$$

Se la lista $L(w_1, w_2, \dots, w_s)$ viene modificata in uno dei seguenti modi:

- scambio di due vettori:
$$W' = L(w_1, w_2, \dots, w_k, \dots, w_h, \dots, w_s) \quad 1 \leq h < k \leq s$$
- moltiplicazione di un vettore della lista per un numero diverso da 0:
$$W' = L\left(w_1, w_2, \dots, \alpha w_j, \dots, w_s\right) \quad 1 \leq j \leq s,\ \alpha \in \mathbb{K},\ \alpha \neq 0$$
- addizione ad un vettore della lista di un multiplo di un altro vettore della lista:
$$W' = L(w_1, w_2, \dots, w_h + \beta w_k, \dots, w_s) \quad 1 \leq h,\ k \leq s,\ h \neq k,\ \beta \in \mathbb{K}$$

Allora il sottospazio vettoriale **W'** generato dalla lista modificata coincide con lo spazio vettoriale **W** generato dalla lista originale.

Questa proprietà è il motivo per cui il rango di una matrice (le cui righe formano un sottospazio vettoriale) non viene modificato quando la matrice viene trasformata in una matrice a gradini.

Dimostrazione

L'obiettivo è quello di dimostrare che il vettore di **W**' che ha subito la modifica è contenuto anche in **W**, in modo da dimostrare la doppia inclusione $W \subseteq W'$ e $W' \subseteq W$, ovvero $W' = W$.

Parte 1: scambio di due vettori

Se il vettore w_k viene scambiato di posto con il vettore w_h, è evidente che le due liste **W** e **W'** rimangono identiche, a meno dell'ordine, pertanto tutti i vettori di **W** sono contenuti in **W'** e viceversa, dunque $W = W'$.

Parte 2: moltiplicazione di un vettore per uno scalare diverso da 0

Il vettore $\alpha w_j \in W'$ è presente anche in **W**: essendo uno spazio vettoriale, contenendo w_j contiene anche ogni suo multiplo. Dunque tutti i vettori di **W** sono contenuti in **W'** e viceversa, pertanto $W = W'$.

Parte 3: addizione ad un vettore di un multiplo di un altro vettore

Il vettore $w_h + \beta w_k \in W'$ è contenuto anche in **W**: essendo uno spazio vettoriale, siccome $w_h \in W$ e $w_k \in W$, esso contiene anche tutte le loro combinazioni lineari. Dunque tutti i vettori di **W** sono contenuti in **W'** e viceversa, pertanto $W = W'$.

Dimensione di uno spazio vettoriale

Sia $V \neq \{0\}$ uno spazio vettoriale sul campo numerico $\mathbb{K}$ e sia *n* un intero positivo. Si dice che **V** ha **dimensione** *n* se esiste in **V** una base costituita da esattamente *n* vettori.

La dimensione di uno spazio vettoriale **V** si indica come

$$dim(V) = n$$

Per definizione si ha che

$$dim(\{0\}) = 0$$

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$ e sia **W** un sottospazio vettoriale di **V**.
Se **V** è finitamente generato con $dim(V) = n$, allora anche **W** è finitamente generato e risulta che

$$dim(W) \leq n$$

L'uguaglianza $dim(W) = dim(V) = n$ risulta vera se e solo se $W = V$.

Dimostrazione

Supponiamo per assurdo che il sottospazio **W** non sia finitamente generato, quindi in **W** esistono infiniti vettori linearmente indipendenti. Essendo **W** un sottospazio vettoriale di **V**, questa lista infinita di vettori linearmente indipendenti è presente anche in **V**, quindi se **W** non è finitamente generato, nemmeno **V** è finitamente generato. Ma questa è una contraddizione dell'ipotesi, quindi **W** deve essere finitamente generato.

Posto quindi $dim(W) = k$, si ha che una base di **W** è costituita da *k* vettori linearmente indipendenti. Questi vettori linearmente indipendenti esistono anche in **V**, quindi se $dim(V) = n$, allora $k \leq n$.

Nel caso particolare in cui $k = n$ significa che una base di **W** è anche una base di **V**, quindi si ha che $W = V$ perché i due spazi vettoriali hanno lo stesso sistema di generatori.

Esempi di dimensione

$dim(\mathbb{K}^n) = n$

La base canonica di $\mathbb{K}^n$ è costituita dagli *n* vettori canonici.

$$e_1 = (1,0,0,\dots,0)$$
$$e_2 = (0,1,0,\dots,0)$$
$$\vdots$$
$$e_n = (0,0,0,\dots,1)$$

$dim\left(M_{m,n}(\mathbb{K})\right) = m \times n$

Per $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n,$ si definisce la matrice E_{ij} come la matrice $m \times n$ che ha il coefficiente 1 nell'incrocio tra la riga *i* e la colonna *j*, e 0 in tutte le altre posizioni. L'insieme di tutte queste matrici E_{ij} forma la base canonica dello spazio vettoriale $M_{m,n}(\mathbb{K})$.

Per esempio, le 4 matrici

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

formano una base dello spazio vettoriale $M_{2,2}(\mathbb{K})$ che ha dimensione 4.

Formula dimensionale di Grassmann

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$ e siano **U** e **W** sottospazi vettoriali di **V**, ciascuno di essi finitamente generato.
Allora anche il sottospazio vettoriale $U + W$ è finitamente generato e risulta

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W)$$

Se $dim(U \cap W) = 0$, allora la formula di Grassmann assume una forma particolare:

$$dim(U + W) = dim(U) + dim(W)$$

In questo caso si dice che il sottospazio somma $U + W$ è la **somma diretta** dei due sottospazi vettoriali **U** e **W** e si scrive

$$U \bigoplus W$$

Dimostrazione

Sia $dim(U) = m$ e $dim(W) = n$. L'insieme $U \cap W$ è un sottospazio vettoriale di **V** che risulta contenuto sia in **U** che in **W**. Posto quindi $dim(U \cap W) = k$, si ha che:

$$k \leq m \quad k \leq n$$

Sia ora $\{z_1, z_2, \dots, z_k\}$ una base di $U \cap W$. Per il teorema del completamento ad una base, esistono:

$$\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m\} \quad \text{base di U}$$

$$\{z_1, z_2, \dots, z_n, w_{k+1}, \dots, w_n\} \quad \text{base di W}$$

Per dimostrare la validità della formula di Grassmann, occorre dimostrare che l'insieme

$$\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m, w_{k+1}, \dots, w_n\}$$

forma una base di $U + W$. In questo modo si dimostra che $U + W$ è finitamente generato e che la sua dimensione è data da:

$$dim(U + W) = k + (m - k) + (n - k) = m + n - k$$

Poiché $\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m\}$ è una base di **U**, allora un generico vettore $u \in U$ può essere espresso come combinazione lineare dei vettori di tale base:

$$u = \alpha_1 z_1 + \alpha_2 z_2 + \dots + \alpha_{k+1} u_{k+1} + \dots + \alpha_m u_m$$

Lo stesso deve valere per un qualsiasi vettore $w \in W$ e la sua base $\{z_1, z_2, \dots, z_k, w_{k+1}, \dots, w_n\}$:

$$w = \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k w_k + \beta_{k+1} w_{k+1} + \dots + \beta_n w_n$$

Quindi il generico vettore $(u + w) \in (U + W)$ è dato da:

$$u + w = \alpha_1 z_1 + \alpha_2 z_2 + \dots + \alpha_m u_m + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_n w_n$$

E si tratta di una combinazione lineare dei vettori $\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m, w_{k+1}, \dots, w_n\}$.

Ora occorre dimostrare che i vettori dell'insieme $\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m, w_{k+1}, \dots, w_n\}$ sono linearmente indipendenti. Si considera quindi la seguente combinazione lineare:

$$\lambda_1 z_1 + \dots + \lambda_k z_k + \mu_{k+1} u_{k+1} + \dots + \mu_m u_m + \nu_{k+1} w_{k+1} + \dots + \nu_n w_n = 0$$

I vettori $w_{k+1}, \dots, w_n$ si possono portare a secondo membro:

$$\lambda_1 z_1 + \dots + \lambda_k z_k + \mu_{k+1} u_{k+1} + \dots + \mu_m u_m = -\nu_{k+1} w_{k+1} - \dots - \nu_n w_n$$

Il vettore a primo membro $\lambda_1 z_1 + \dots + \mu_m u_m$ appartiene ad **U**, in quanto combinazione lineare dei vettori della base di **U**.
Il vettore a secondo membro $-\nu_{k+1} w_{k+1} - \dots - \mu_n w_n$ appartiene a **W**, in quanto combinazione lineare dei vettori della base di **W**.
Ne segue quindi che il vettore espresso dalla combinazione lineare sopra appartiene ad $U \cap W$, quindi il vettore $-\nu_{k+1} w_1 - \dots - \nu_n w_n$ si deve poter esprimere come combinazione lineare dei vettori della base di $U \cap W$, quindi devono esistere scalari $\eta_1, \dots, \eta_k$ tali che:

$$\eta_1 z_1 + \eta_2 z_2 + \dots + \eta_k z_k = -\nu_{k+1} w_{k+1} - \dots - \nu_n w_n$$

Portando tutti a primo membro:

$$\eta_1 z_1 + \dots + \eta_k z_k + \nu_{k+1} w_{k+1} + \dots + \nu_n w_n = 0$$

I vettori $z_1, z_2, \dots, z_k, w_{k+1}, \dots, w_n$ sono linearmente indipendenti, in quanto formano una base di **W**. Pertanto si deve avere che:

$$\eta_1 = 0 \quad \eta_2 = 0 \quad \dots \quad \eta_\mathrm{k} = 0 \quad \nu_{k+1} = 0 \quad \dots \quad \nu_n = 0$$

Inserendo queste informazioni nella relazione iniziale, si ha che

$$\lambda_1 z_1 + \lambda_2 z_2 + \dots + \lambda_k z_k + \mu_{k+1} u_{k+1} + \dots + \mu_m u_m = 0$$

Ma anche i vettori $z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m$ sono linearmente indipendenti, perché formano una base di **U**. Allora si deve avere che

$$\lambda_1 = 0 \quad \dots \quad \lambda_k = 0 \quad \mu_{k+1} = 0 \quad \dots \quad \mu_m = 0$$

Dunque è stato dimostrato che l'insieme

$$\{z_1, z_2, \dots, z_k, u_{k+1}, \dots, u_m, w_{k+1}, \dots, w_n\}$$

è composto da vettori linearmente indipendenti, quindi forma una base di $U + W$. Pertanto la formula dimensionale di Grassmann è dimostrata.

Sia **V** uno spazio vettoriale sum campo numerico $\mathbb{K}$ e siano **U** e **W** sottospazi vettoriali di **V**.
Il sottospazio somma $U + W$ è una somma diretta se e solo se ogni vettore $y \in U + W$ può essere rappresentato in un unico modo come

$$y = u + w \quad u \in U,\, w \in W$$

Dimostrazione

Supponiamo che il sottospazio somma $U + W$ sia una somma diretta, quindi $U \cap W = \{0\}$, e che il vettore $y \in U + W$ possa essere rappresentato in due modi:

$$(1) \quad y = u + w \quad u \in U, w \in W$$

$$(2) \quad y = u' + w' \quad u' \in U, w' \in W$$

Uguagliando le due espressioni si ha che:

$$u + w = u' + w' \to u - u' = w - w'$$

Il vettore $u + u'$ appartiene ad **U**, in quanto combinazione lineare di vettori di **U**, mentre il vettore $w + w'$ appartiene a **W**, in quanto combinazione lineare di vettori di **W**. Quindi entrambi i vettori $u + u'$, $w + w'$ appartengono ad $U \cap W$. Ma per ipotesi $U \cap W = \{0\}$, quindi si deve avere che:

$$u - u' = 0 = w - w' \to u = u' \quad w = w'$$

Quindi le due rappresentazioni sono in realtà una sola.

Supponiamo viceversa che nel sottospazio somma $U + W$ vi sia l'unicità della rappresentazione di ciascuno dei suoi vettori come somma di un vettore di **U** e di un vettore di **W**.

Sia quindi $z \in U \cap W$, quindi:

$$(1) \quad z = z + 0$$

$$(2) \quad z = 0 + z$$

Nella rappresentazione (1), i due vettori **z** e **0** si possono pensare come $z \in U$ e $0 \in W$. Viceversa, nella rappresentazione (2) i due vettori **z** e **0** si possono pensare come $0 \in U$ e $z \in W$.

Entrambe le rappresentazioni però danno luogo allo stesso vettore dello spazio vettoriale $U + W$, quindi devono coincidere, pertanto si deve avere che $z = 0$, dove **0** è l'unico vettore del sottospazio vettoriale $U \cap W$, quindi la proprietà è dimostrata.

Permutazioni

Sia Ω un insieme finito. Una **permutazione** di Ω (o su Ω) è una funzione **biiettiva** $f : \Omega \to \Omega$.

L'insieme di tutte le permutazioni su Ω si indica con $Sym(\Omega)$.

Se *n* è un intero positivo, indicheremo con S_n l'insieme $Sym(\Omega)$, intendendo che Ω è un insieme con *n* elementi.

Una permutazione si può rappresentare in forma tabellare verticale:

$$\begin{array}{cccc} f: & \Omega & \mapsto & \Omega \\ & 1 & \mapsto & f(1) \\ & 2 & \mapsto & f(2) \\ & & \vdots & \\ & n & \mapsto & f(n) \end{array}$$

oppure in forma tabellare orizzontale, ovvero con una matrice:

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ f(1) & f(2) & f(3) & \dots & f(n) \end{pmatrix}$$

Se gli elementi del dominio sono scritti in ordine crescente (come in questo caso) la tabella la si può comprimere ad una sola riga:

$$\begin{pmatrix} f(1) & f(2) & f(3) & \dots & f(n) \end{pmatrix}$$

O equivalentemente

$$\begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_n \end{pmatrix}$$

La stringa $(f_1\, f_2\, \dots\, f_n)$ rappresenta una permutazione di Ω se e solo se ciascuno degli elementi $1, 2, \dots, n$ compare nella stringa una e una sola volta (come da definizione di funzione biiettiva).

Il numero di permutazioni di *n* oggetti è $n!$.

Prodotto di permutazioni

Quando *f* e *g* sono entrambe permutazioni di Ω (cioè $f, g \in Sym(\Omega)$), la loro **funzione composta** $f \circ g$ esiste sempre:

$$(g \circ f)(x) = g\big(f(x)\big) \quad \text{per ogni } x \in \Omega$$

Inoltre essendo *f* e *g* funzioni biiettive, anche la funzione composta $g \circ f$ è una funzione biiettiva, quindi anche $g \circ f \in Sym(\Omega)$.

Un modo rapido per calcolare il prodotto di permutazioni, una volta che *f* e *g* sono scritte in forma tabellare, è quello di scrivere una nuova tabella di 2 righe, in cui la 1° riga corrisponde alla 1° riga di *f*, mentre la 2° riga corrisponde alla 2° riga di *g*.

Esempio:

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$$

$$g = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

$$g \circ f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

Segno di una permutazione

Sia *n* un intero positivo e sia *N* il valore definito come:

$$N := \prod_{1 \leq i < j \leq n} (j - i) = (2 - 1)(3 - 1) \dots \big(n - (n - 1)\big)$$

N è quindi il prodotto di tutti i termini $(j - i)$ dove $i < j$.

Se $f \in S_n$ è una permutazione di *n* oggetti, allora si può esprimere N_f come:

$$N_f := \prod_{1 \leq i < j \leq n} \big(f(j) - f(i)\big)$$

I fattori di *N* sono gli stessi di N_f, a meno del segno. Pertanto, detto t_f il numero delle coppie $\{i, j\}$ di elementi di $\{1, 2, \dots, n\}$ con $i < j$ e $f(i) > f(j)$, si ha che:

$$N_f = (-1)^{t_f} N$$

Il numero $(-1)^{t_f}$ si chiama **segno di *f*** e si indica con $\epsilon(f)$ (oppure $sgn(f)$). In base alla relazione definita sopra, il segno di *f* lo si può calcolare come:

$$\epsilon(f) = \frac{N_f}{N}$$

La **permutazione identica** $id : \Omega \to \Omega$, definita ponendo $id(x) = x$ per ogni $x \in \Omega$ ha evidentemente segno uguale ad 1.

Determinanti

Sia *A* una matrice $n \times n$ a coefficienti nel campo numerico $\mathbb{K}$. Si chiama **determinante** di *A*, e si indica con $det(A)$, il numero

$$det(A) = \sum_{p \in S_n} \epsilon(p) a_{1p(1)}\, a_{2p(2)} \dots a_{np(n)}$$

dove S_n denota l'insieme di tutte le permutazioni di *n* oggetti e dove $\epsilon(p)$ è il **segno** della permutazione $p \in S_n$.

Il determinante di *A*, quindi, è la somma dei prodotti degli elementi delle diverse permutazioni moltiplicati per il segno della permutazione. Gli addendi di questa somma sono $n!$, tanti quante sono le permutazioni $p \in S_n$.

Un modo alternativo per indicare il determinante di *A* è quello delimitare la matrice *A* con delle barre verticali:

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Proprietà del determinante

Sia *A* una matrice quadrata d'ordine *n*. Valgono le seguenti proprietà:

$$\boxed{det(A^t) = det(A)}$$

Dimostrazione

Si ha che:

$$det(A) = \sum_{p \in S_n} \epsilon(p) a_{1p(1)}\, a_{2p(2)} \dots a_{np(n)}$$

per ottenere il determinante della trasposta occorre invertire gli indici dei coefficienti:

$$det(A^t) = \sum_{p \in S_n} \epsilon(p) a_{p(1)1}\, a_{p(2)2} \dots a_{p(n)n}$$

A meno del segno, gli addendi di $det(A^t)$ sono gli stessi di $det(A)$, perché il termine $a_{p(1)1} a_{p(2)2} \dots a_{p(n)n}$ può essere riscritto come:

$$a_{1p^{-1}(1)} a_{2p^{-1}(2)} \dots a_{np^{-1}(n)}$$

Dato che $\epsilon(p^{-1}) = \epsilon(p)$, si conclude che $det(A) = det(A^t)$.

Se l'i-esima riga di A può essere espressa come combinazione lineare di due righe R ed R' di A, per esempio:

$$A^{(i)} = \alpha R + \alpha' R'$$

allora

$$det(A) = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ \alpha R + \alpha' R' \\ \vdots \\ A^{(n)} \end{pmatrix} + \alpha det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ R \\ \vdots \\ A^{(n)} \end{pmatrix} + \alpha' det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ R' \\ \vdots \\ A^{(n)} \end{pmatrix}$$

Questa proprietà viene espressa dicendo che il determinante è lineare su ciascuna riga (o colonna) della matrice A.

Caso particolare: quando $\alpha' = 0$, si ha che:

$$det(A) = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ \alpha R \\ \vdots \\ A^{(n)} \end{pmatrix} = \alpha det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ R \\ \vdots \\ A^{(n)} \end{pmatrix}$$

Questo caso si può esprimere dicendo che se in una matrice quadrata si moltiplica una riga per uno scalare $\alpha \neq 0$, allora:

$$det(A) = \alpha det(A)$$

Se A presenta due righe uguali oppure due colonne uguali, allora

$$det(A) = 0$$

Dimostrazione

Se la riga A ha due righe uguali, scambiandole tra loro la matrice A non cambia d'aspetto, ma il suo determinante cambia di segno, quindi:

$$det(A) = -det(A) \rightarrow 2det(A) = 0 \rightarrow det(A) = 0$$

$$det(Id_n) = 1$$

Più in generale, il determinante di una **matrice diagonale** è uguale al prodotto dei coefficienti che stanno sulla diagonale principale. Nel caso della matrice identica questi coefficienti sono tutti uguali ad 1, dunque $det(Id_n) = 1$.

Se *A* ha una riga o una colonna nulla, allora

$$det(A) = 0$$

Dimostrazione

Supponiamo che la riga $A^{(i)}$ abbia tutti i coefficienti uguali a 0. Dalla definizione di determinante:

$$det(A) = \sum_{p \in S_n} \epsilon(p) a_{1p(1)}\, a_{2p(2)} \dots a_{np(n)}$$

All'interno del prodotto $a_{1p(1)}a_{2p(2)} \dots a_{np(n)}$ compaiono anche i coefficienti della riga $A^{(i)}$, che vanno quindi ad annullare l'intero prodotto. Il determinante quindi si riduce ad una sommatoria in cui tutti gli addendi sono uguali a 0 e dunque $det(A) = 0$.

Sia *L* la matrice ottenuta scambiando due righe oppure due colonne della matrice *A*. Allora

$$det(L) = -det(A)$$

Se le righe o le colonne di *A* sono linearmente dipendenti, allora

$$det(A) = 0$$

Come conseguenza di questa proprietà, si ha che se *A* ha rango *n*, allora il suo determinante è diverso da 0.

Dimostrazione

Supponiamo che la riga $A^{(1)}$ possa essere espressa come combinazione lineare delle altre righe di *A*. Allora, dalla linearità del determinante:

$$det(A) = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} = det\begin{pmatrix} \alpha_1 A^{(2)} + \alpha_2 A^{(3)} + \cdots + \alpha_n A^{(n)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} =$$

$$= \alpha_1 det\begin{pmatrix} A^{(2)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \alpha_2 det\begin{pmatrix} A^{(3)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \cdots + \alpha_n det\begin{pmatrix} A^{(n)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix}$$

Tutte queste matrici hanno due righe uguali, quindi il loro determinante è 0, dunque anche il determinante di *A* è 0 e la proprietà è dimostrata.

Sia A una matrice quadrata d'ordine n e siano i e j indici di riga diversi. Sia B la matrice ottenuta da A sostituendo la riga $A^{(i)}$ con $A^{(i)} + \beta A^{(j)}$, con $\beta \in \mathbb{K}$. Allora

$$det(B) = det(A)$$

Dimostrazione

Segue dalla linearità del determinante sulla riga *i*:

$$det(B) = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(i)} + \beta A^{(j)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(n)} \end{pmatrix} = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(i)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \beta det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(n)} \end{pmatrix}$$

L'ultima matrice contiene due righe uguali, cioè contiene due volte $A^{(j)}$, quindi il suo determinante è 0. Allora la precedente relazione si può scrivere come:

$$det(B) = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(i)} + \beta A^{(j)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(n)} \end{pmatrix} = det\begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(i)} \\ \vdots \\ A^{(j)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \beta \cdot 0$$

da cui segue che $det(B) = det(A)$.

Teorema di Binet: moltiplicatività del determinante

Se A e B sono due matrici quadrate d'ordine n, allora

$$det(AB) = det(A)det(B)$$

Come diretta conseguenza del teorema di Binet si ha un altro teorema:

A è invertibile se e solo se il suo determinante è diverso da 0.

Dimostrazione

Parte 1: A invertibile $\Rightarrow det(A) \neq 0$

Dalla relazione

$$AA^{-1} = Id_n$$

prendendo il determinante di entrambi i membri si ha che:

$$det(AA^{-1}) = det(Id_n)$$

applicando il teorema di Binet e ricordando che $det(Id_n) = 1$, si ha che:

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

dunque si deve necessariamente avere che sia $det(A)$ che $det(A^{-1})$ sono diversi da 0, altrimenti la relazione non sarebbe soddisfatta. Nello specifico si deve avere che:

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

Parte 2: $det(A) \neq 0 \Rightarrow A$ è invertibile

Se $det(A) \neq 0$, allora $rk(A) = n$.

Supponiamo per assurdo che invece $rk(A) < n$, quindi che le righe di A siano linearmente dipendenti: allora, dalle proprietà del determinante, si ha che $det(A) = 0$. Ma questa è una contraddizione dell'ipotesi, dunque se $det(A) \neq 0$ allora A è invertibile.

Questo teorema permette di dire che le seguenti affermazioni sono tutte equivalenti tra loro:

1. $rk(A) = n$;
2. A è una matrice invertibile;
3. $det(A) \neq 0$

Calcolo del determinante per matrici quadrate di ordine 1

$$A = (a_{11})$$

Il determinante in questo caso è dato dall'unico elemento della matrice:

$$det(A) = a_{11}$$

Calcolo del determinante per matrici quadrate di ordine 2

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

Il determinante è dato dal prodotto degli elementi della **diagonale principale** meno il prodotto degli elementi della **diagonale secondaria**:

$$det(A) = (a_{11} \cdot a_{22}) - (a_{12} \cdot a_{21})$$

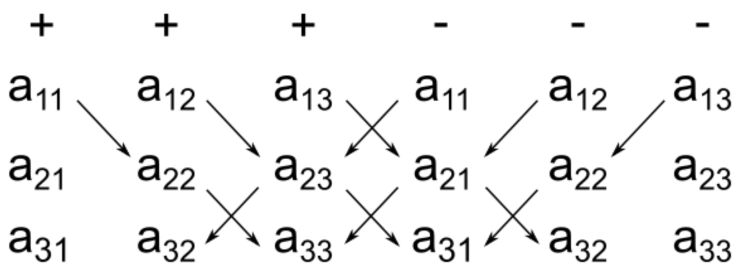
Calcolo del determinante per matrici quadrate di ordine 3 – regola di Sarrus

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

La regola di Sarrus consiste nello scrivere 2 volte di seguito la matrice *A*:

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & a_{33} \end{matrix}$$

Il determinante è dato dalla somma dei prodotti delle linee parallele alla diagonale principale meno la somma dei prodotti delle linee parallele alla diagonale secondaria:



Determinante di matrici 3x3 con la regola di Sarrus

$$det(A) = a_{11}a_{22}a_{23} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31})$$

Teorema di Laplace per il calcolo del determinante

Un altro metodo per il calcolo del determinante è quello degli **sviluppi di Laplace**. Si tratta di formule ricorsive che calcolano il determinante della matrice di partenza calcolando il determinante delle sue sottomatrici.

Sia quindi *A* una matrice definita come segue:

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

e sia A_{ij} la matrice che si ottiene eliminando la riga *i* e la colonna *j* dalla matrice *A* e con $det(A_{ij})$ il determinante di tale matrice.

Fissato un qualsiasi elemento $a_{ij} \in A$, si chiama **complemento algebrico** di a_{ij} il numero

$$(-1)^{i+j} \cdot det(A_{ij})$$

A questo punto lo sviluppo di Laplace può essere fatto **per righe** oppure **per colonne**.

Sviluppo di Laplace per righe

Fissata una riga *i* qualsiasi della matrice *A*, il determinante di *A* è pari alla somma dei prodotti degli elementi della riga *i* per i rispettivi complementi algebrici:

$$det(A) = \sum_{j=1}^{n} \left[a_{ij} \cdot (-1)^{i+j} \cdot det(A_{ij})\right]$$

Sviluppo di Laplace per colonne

Fissata una colonna *j* qualsiasi della matrice *A*, il suo determinante è uguale alla somma dei prodotti degli elementi della colonna *j* per i rispettivi complementi algebrici:

$$det(A) = \sum_{i=1}^{n} \left[a_{ij} \cdot (-1)^{i+j} \cdot det(A_{ij})\right]$$

Funzioni lineari

Siano **V** e **W** due spazi vettoriali sul campo numerico $\mathbb{K}$ e sia $F: V \to W$. Si dice che *F* è una **funzione lineare** se soddisfa le seguenti condizioni:

- **additività**: per ogni $v_1, v_2 \in V$, l'immagine della somma è uguale alla somma delle immagini:
$$F(v_1 + v_2) = F(v_1) + F(v_2)$$
- **omogeneità**: per ogni $v \in V$ e per ogni $\lambda \in \mathbb{K}$, l'immagine del prodotto di **v** per lo scalare λ è uguale al prodotto dello scalare per l'immagine di **v**:
$$F(\lambda v) = \lambda F(v)$$

Le condizioni possono anche essere raggruppate in una sola condizione: *F* è una funzione lineare se conserva le combinazioni lineari dei vettori del dominio:

$$F(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 F(v_1) + \lambda_2 F(v_2)$$

Se si considera uno scalare $\lambda = 0$ e un qualunque vettore $v \in V$, dall'omogeneità dell'applicazione lineare si ha che:

$$F(0 \cdot v) = 0 \cdot F(v) = 0_w \quad 0_w \in W$$

ovvero *F* mappa il vettore nullo di **V** con il vettore nullo di **W**.

Classificazione delle funzioni lineari

- **omomorfismo**: è una funzione lineare dove i due spazi vettoriali del dominio e del codominio condividono lo stesso campo numerico;
- **endomorfismo** (o **operatore lineare**): è una funzione lineare dove lo spazio vettoriale del dominio corrisponde a quello del codominio (es. $F: V \to V$);
- **isomorfismo**: è una funzione lineare **biiettiva** con dominio e codominio distinti;
- **automorfismo**: è una funzione lineare **biiettiva** dove il dominio è uguale al codominio;

Funzione lineare inversa

Se $F: V \to W$ è una funzione lineare **biiettiva** (cioè se è un isomorfismo o un automorfismo), allora esiste la relativa **funzione inversa**, che è a sua volta una funzione lineare biiettiva.

$$F^{-1}: W \to V$$

Composizione di funzioni lineari

Siano $F: V \to W$ e $G: W \to U$ due funzioni lineari. Poiché lo spazio d'arrivo di *F* coincide con lo spazio di partenza di *G*, è possibile definire la **funzione composta** $F \circ G$, che è a sua volta una funzione lineare

$$F \circ G: V \to U$$

Esempi di funzioni lineari

Funzione identità

È un endomorfismo su **V** che ad ogni vettore $v \in V$ associa il vettore **v** stesso:

$$\forall v \in V \quad F(v) = v$$

Funzione lineare nulla

Associa a qualunque elemento $v \in V$ il vettore nullo del codominio **W**:

$$\forall v \in V \quad F(v) = 0_w$$

Forma lineare

Se **V** è uno spazio vettoriale sul campo numerico $\mathbb{K}$ e **W** è uno spazio vettoriale sul campo numerico $\mathbb{K}^1$, la funzione lineare $F: V \to W$ prende il nome di **forma lineare**.
In altre parole, in una forma lineare un vettore viene mappato con uno scalare.

Mappe scalari

Una **mappa scalare** è un automorfismo vettoriale $F: V \to V$ che associa ad ogni vettore $v \in V$ il vettore αv, con $\alpha \in \mathbb{K}$ costante.

$$F(v) = \alpha v \quad \forall v \in V$$

Se $\alpha = 0$ allora si tratta della funzione lineare nulla.
Se invece $\alpha \neq 0$ allora *F* è invertibile e si ha che:

$$F^{-1}: V \to V$$

$$v \mapsto \frac{1}{\alpha} v$$

Isomorfismo canonico definito da una base

Si tratta di una funzione lineare $F: V \to \mathbb{K}^n$ che ad ogni vettore $v \in V$ associa le coordinate di **v** rispetto ad una base di **V**.

È una funzione lineare biiettiva, perché tutti i vettori di **V** possono essere rappresentati come combinazione lineare dei vettori della base (suriettività) e perché tale rappresentazione è univoca per ogni vettore di **V** (iniettività).

Criteri per verificare l'esistenza e l'unicità di un'applicazione lineare

Siano **V** e **W** due spazi vettoriali sul $\mathbb{K}$ e sia $F: V \to W$ una funzione. La **condizione necessaria** affinché *F* esista è che *F* mappi il vettore nullo di **V** con il vettore nullo di **W**:

$$F(0_V) = 0_W$$

Se ciò non accade, allora *F* non è una funzione lineare. Se invece questa condizione è rispettata, per verificare che *F* esiste occorre dimostrare che *F* conserva le combinazioni lineari dei propri vettori:

$$F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 F(v_1) + \alpha_2 F(v_2)$$

Per fare ciò, esiste un importante teorema:

> Siano **V** e **W** due spazi vettoriali sul campo $\mathbb{K}$, sia $B_V = \{v_1, v_2, \ldots, v_n\}$ una base di **V** e siano $\{w_1, w_2, \ldots, w_n\}$ vettori arbitrari di **W**. Allora esiste una ed una sola applicazione lineare $F: V \to W$ tale che:
>
> $$F(v_i) = w_i \quad i = 1,2, \ldots, n$$

Dimostrazione

Se *F* esiste, allora è anche unica, perché le componenti di un qualunque vettore $v \in V$ rispetto alla base B_V sono univocamente determinate. Quindi è sufficiente dimostrare che vale la condizione di linearità. Dati quindi due vettori $v, v' \in V$:

$$v = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

$$v' = y_1 v_1 + y_2 v_2 + \cdots + y_n v_n$$

si deve avere che:

$$F(v + v') = x_1 F(v_1) + \cdots + x_n F(v_n) + y_1 F(v_1) + \cdots + y_n F(v_n)$$

per ipotesi si ha che $F(v_i) = w_i$ per $i = 1,2, \ldots, n$, quindi la precedente relazione si può riscrivere come:

$$F(v + v') = x_1 w_1 + \cdots + x_n w_n + y_1 w_1 + \cdots + y_n w_n$$

Dunque *F* conserva le combinazioni lineari dei propri vettori ed i coefficienti $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ sono univocamente determinati, quindi *F* esiste ed è unica.

Nucleo di una funzione lineare

Sia $F: V \to W$ una funzione lineare tra spazi vettoriali del campo numerico $\mathbb{K}$. Si definisce **nucleo di *F***, e si indica con $N(F)$, l'insieme di tutti gli elementi del dominio *V* che hanno come immagine mediante la *F* lo zero di **W**:

$$N(F) \coloneqq \{v \in V : F(v) = 0_w\}$$

Il nucleo di una funzione lineare è un sottospazio vettoriale del dominio.

Funzione lineare iniettiva

Sia $F: V \to W$ una funzione lineare. *F* è iniettiva se e solo se

$$N(F) = \{0\} \Leftrightarrow dim\big(N(F)\big) = 0$$

Dimostrazione

Parte 1: *F* iniettiva $\Rightarrow N(F) = \{0\}$

Dalla definizione di funzione iniettiva:

$$\forall v_1, v_2 \in V \; v_1 \neq v_2 \Rightarrow F(v_1) \neq F(v_2)$$

Per definizione di funzione lineare, si ha che $F(0_V) = 0_W$. Allora se il nucleo è composto da tutti i vettori di **V** che vengono mappati col vettore nullo di **W** e la funzione è iniettiva, si deve necessariamente avere che l'unico elemento di $N(F)$ è il vettore nullo di **V**. Se il nucleo contenesse altri elementi, significherebbe che ci sono più vettori di **V** che hanno la stessa immagine (ovvero più vettori di **V** che vengono mappati col vettore nullo di **W**), dunque la funzione non sarebbe iniettiva.

Parte 2: $N(F) = \{0\} \Rightarrow$ *F* è iniettiva

Siano $v, u \in V$ due vettori tali che $F(v) = F(u)$. Dato che *F* è lineare, questo significa che:

$$F(v) - F(u) = 0_W \Leftrightarrow F(v-u) = 0_W$$

Quindi $v - u \in N(F)$. Poiché $N(F)$ contiene il solo vettore nullo, si deve necessariamente avere che $v - u = 0_V$, quindi $v = u$, dunque *F* è iniettiva, perché due vettori uguali del dominio hanno la stessa immagine nel codominio.

Immagine di una funzione lineare

Sia $F: V \to W$ una funzione lineare. Si definisce l'**immagine di V tramite la *F***, e si indica con $Im(F)$, l'insieme formato dalle immagini dei vettori di **V** tramite la *F*:

$$Im(F) = \{F(v): v \in V\}$$

L'immagine di una funzione lineare è un sottospazio vettoriale del codominio. La dimensione di $Im(F)$ prende anche il nome di **rango** di *F*.

> Se $dim(V) = n$ e $B = \{v_1, v_2, \dots, v_n\}$ è una base di **V**, allora le immagini dei vettori della base formano un sistema di generatori per $Im(F)$:
>
> $$Im(F) = L\big(F(v_1), F(v_2), \dots, F(v_n)\big)$$

Dimostrazione

Se *B* è una base di **V**, significa che un generico vettore $v \in V$ può essere espresso come combinazione lineare dei vettori della base *B*. Quindi esistono coefficienti $x_1, x_2, \dots, x_n$ tali per cui:

$$v = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

Dato che *F* è una funzione lineare:

$$F(v) = F(x_1 v_1 + x_2 v_2 + \cdots + x_n v_n) = x_1 F(v_1) + x_2 F(v_2) + \cdots + x_n F(v_n)$$

Quindi un generico vettore $F(v) \in Im(F)$ può essere scritto come combinazione lineare dei vettori $x_1 F(v_1), x_2 F(v_2), \dots, x_n F(v_n)$, dunque tali vettori formano un sistema di generatori per $Im(F)$.

Funzione lineare suriettiva

> Sia $F: V \to W$ una funzione lineare. *F* è suriettiva se e solo se
>
> $$dim\big(Im(F)\big) = dim(W)$$

Dimostrazione

F suriettiva $\Rightarrow dim\big(Im(F)\big) = dim(W)$

Dalla definizione di funzione suriettiva:

$$\forall w \in W \; \exists v \in V : F(v) = w$$

Quindi tutti gli elementi $w \in W$ sono raggiunti da un qualche elemento $v \in V$ tramite la *F*. Poiché l'immagine di una funzione lineare è l'insieme di tutti i vettori $F(v) \in W$ tali che $v \in V$, se *F* è suriettiva si deve necessariamente avere che $Im(F)$ e **W** coincidono, dunque $dim\big(Im(F)\big) = dim(W)$.

$dim\big(Im(F)\big) = dim(W) \Rightarrow F$ è suriettiva

Se $dim\big(Im(F)\big) = dim(W)$, significa che tutti i vettori di **W** vengono raggiunti da un qualche vettore $v \in V$. Questa è esattamente la definizione di funzione suriettiva.

Formula dimensionale per funzioni lineari

Sia $F: V \rightarrow W$ una funzione lineare e sia **V** uno spazio vettoriale finitamente generato, con $dim(V) = n$. Allora anche $Im(F)$ è finitamente generato e risulta che:

$$dim\big(N(F)\big) + dim\big(Im(F)\big) = dim(V)$$

Dimostrazione

Sia $B = \{v_1, v_2, \dots, v_n\}$ una base di **V**. Le immagini dei vettori di *B* formano un sistema di generatori per $Im(F)$, quindi poiché *B* è di dimensione finita, anche $Im(F)$ è di dimensione finita e vale:

$$dim\big(Im(F)\big) \leq n$$

Sia ora $dim\big(N(F)\big) = k$. Per definizione di sottospazio vettoriale si ha che $k \leq n$. Si definisce ora una base per $N(F)$:

$$B_{N(F)} = \{z_1, z_2, \dots, z_k\}$$

Dato che $N(F)$ è un sottospazio vettoriale di **V**, i vettori $z_1, z_2, \dots, z_k$ sono vettori linearmente indipendenti in **V**, quindi per il teorema del completamento ad una base esistono in **V** dei vettori $v_1, v_2, \dots, v_{n-k}$ tali che:

$$B_V = \{z_1, z_2, \dots, z_k, v_1, v_2, \dots, v_{n-k}\} \quad \text{base di V}$$

Le immagini dei vettori di B_V formano un sistema di generatori per $Im(F)$:

$$Im(F) = L\big(F(z_1), F(z_2), \dots, F(z_k), F(v_1), F(v_2), \dots, F(v_{n-k})\big)$$

I vettori $z_1, z_2, \dots, z_k$ appartengono al nucleo di *F*, quindi $F(z_1) = F(z_2) = \cdots = F(z_k) = 0_W$. Perciò il sistema di generatori si può semplificare:

$$Im(F) = L\big(F(v_1), F(v_2), \dots, F(v_{n-k})\big)$$

Quindi un possibile sistema di generatori per $Im(F)$ è formato da $n - k$ vettori. Se questo sistema fosse una base di $Im(F)$, dalla formula dimensionale si avrebbe che:

$$n - k + k = n$$

e dunque tale formula risulterebbe dimostrata. Occorre quindi dimostrare che i vettori $F(v_1), F(v_2), \dots, F(v_{n-k})$ sono linearmente indipendenti (e quindi formano una base di $Im(F)$). Si considera quindi la combinazione lineare:

$$\alpha_1 F(v_1) + \alpha_2 F(v_2) + \cdots + \alpha_{n-k} F(v_{n-k}) = 0 \quad (\star)$$

Dalla linearità di *F*:

$$F(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n-k} v_{n-k}) = 0$$

Quindi il vettore $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n-k} v_{n-k}$ appartiene al nucleo di *F*, siccome viene mappato con il vettore nullo. Dato che $B_{N(F)} = \{z_1, z_2, \dots, z_k\}$, esistono coefficienti $\beta_1, \beta_2, \dots, \beta_k \in \mathbb{K}$ tali che:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n-k} v_{n-k} = \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_k z_k$$

si portano poi a primo membro tutti i termini:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n-k} v_{n-k} - \beta_1 z_1 - \beta_2 z_2 - \cdots - \beta_k z_k = 0$$

Questa è una combinazione lineare dei vettori di B_V. Tutti questi vettori sono linearmente indipendenti, quindi si deve avere che:

$$\alpha_1 = 0 \quad \dots \quad \alpha_{n-k} = 0 \quad \beta_1 = 0 \quad \dots \quad \beta_k = 0$$

Considerando quindi $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_{n-k} = 0$ nella relazione $(\star)$, si ottiene che i vettori $F(v_1), F(v_2), \dots, F(v_{n-k})$ sono linearmente indipendenti e pertanto formano una base di $Im(F)$. Quindi:

$$dim\big(Im(F)\big) = n - k$$

Pertanto il teorema è dimostrato.

Matrice associata ad una funzione lineare

Sia $F: V \to W$ una funzione lineare, con $dim(V) = n$ e $dim(W) = m$. Siano $B_V = \{v_1, v_2, \dots, v_n\}$ e $B_W = \{w_1, w_2, \dots, w_m\}$ rispettivamente una base di **V** e una base di **W**.

Il generico vettore $v_j \in B_V$ può essere espresso come combinazione lineare dei vettori di B_W, quindi per ogni $v_j \in B_V$ esistono coefficienti $\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj} \in \mathbb{K}$ tali che:

$$F(v_j) = \alpha_{1j} w_1 + \alpha_{2j} w_2 + \cdots + \alpha_{mj} w_m$$

I coefficienti $\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj}$, che formano le coordinate di $F(v_j)$ rispetto alla base B_W, possono essere scritti come **colonne** all'interno di una matrice. Questo vale per ogni $F(v_j) \in W$, quindi si arriva ad avere una matrice $m \times n$:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

La matrice *A* prende il nome di **matrice associata** alla funzione lineare *F* rispetto alle basi B_V e B_W. Ogni colonna di *A* è formata dalle coordinate del vettore $F(v_j)$ rispetto alla base di **W**.

Un generico vettore $v \in V$ può essere espresso come combinazione lineare dei vettori di B_V:

$$v = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

mentre un generico vettore $F(v) \in W$ può essere espresso come combinazione lineare dei vettori di B_W:

$$F(v) = y_1 w_1 + y_2 w_2 + \cdots + y_m w_m$$

Le coordinate $x_1, x_2, \dots, x_n$ e $y_1, y_2, \dots, y_m$ sono legate da loro da un'equazione, detta **equazione matriciale** di *F* rispetto alle basi B_V e B_W:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

In base a com'è definita la matrice associata ad una funzione lineare, si ha la relazione:

$$dim(Im(F)) = rk(A)$$

Una funzione ha infinite matrici associate, in quanto uno spazio vettoriale ha infinite basi. Le matrici associate ad una funzione sono simili tra loro.

Autovettori, autovalori e autospazio

Sia **V** uno spazio vettoriale sul campo numerico $\mathbb{K}$ e sia *F* un endomorfismo su **V**. Si dice che **v** è un **autovettore** di *F* se **v** non è nullo e se esiste uno scalare $\lambda \in \mathbb{K}$ tale che

$$F(v) = \lambda v$$

Il coefficiente λ prende il nome di **autovalore** di *F*.

L'autovalore relativo a ciascun autovettore è univocamente determinato.

Dimostrazione

Supponiamo che valgano contemporaneamente $F(v) = \lambda v$ ed $F(v) = \mu v$. Allora:

$$\lambda v = \mu v \to (\lambda - \mu)v = 0$$

Siccome **v** è non nullo per ipotesi, si ha necessariamente che $\lambda = \mu$, dunque i due autovalori in realtà sono uno solo.

Gli autovettori aventi lo stesso autovalore λ, assieme al vettore nullo, formano un sottospazio vettoriale di **V** detto **autospazio** relativo all'autovalore λ, indicato con V_λ:

$$V_\lambda(F) = \{v \in V : F(v) = \lambda v\} \quad \lambda \in \mathbb{K}, \lambda \text{ autovalore}$$

Se λ non è un autovalore, allora $V_\lambda(F)$ si riduce al solo vettore nullo.

La dimensione di $V_\lambda(F)$ prende il nome di **molteplicità geometrica** dell'autovalore λ.

Relazione tra molteplicità algebrica e geometrica

$$1 \le m_g(\lambda_1) \le m_a(\lambda_1)$$

λ_1 è l'autovalore, $m_g(\lambda_1)$ è la molteplicità geometrica ed $m_a(\lambda_1)$ è la molteplicità algebrica.

La molteplicità geometrica di λ_1 si può calcolare risolvendo un sistema lineare omogeneo del tipo $(A - \lambda_1 Id_n)x = 0$:

$$\begin{pmatrix} a_{11} - \lambda_1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda_1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} - \lambda_1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0$$

La dimensione del sottospazio vettoriale generato dalle soluzioni di questo sistema è la molteplicità geometrica dell'autovalore λ_1. Tutto questo si può riassumere con una formula:

$$m_g(\lambda_1) = n - rk(A - \lambda_1 Id_n)$$

dove *n* è il numero totale di variabili del sistema, mentre $rk(A - \lambda_1 Id_n)$ è il numero delle variabili pivotali.

Polinomio caratteristico

Sia $F: V \to W$ una funzione lineare e sia *A* la matrice associata ad *F* rispetto alle basi B_V e B_W. Si chiama **polinomio caratteristico** il polinomio definito da:

$$det(A - \lambda Id_n)$$

dove λ è un'incognita. La matrice $A - \lambda Id_n$ è fatta in questo modo:

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} - \lambda \end{pmatrix}$$

In questa matrice compare il termine $-\lambda$ all'interno di tutti i coefficienti della diagonale principale.

Il polinomio caratteristico è importante perché grazie ad esso si possono trovare gli autovalori di una funzione lineare.

> Gli autovalori di *F* sono le radici del polinomio caratteristico, ovvero le soluzioni dell'equazione:
> $$det(A - \lambda Id_n) = 0$$

Ogni radice del polinomio caratteristico (e quindi ogni autovalore di *F*) presenta una certa molteplicità, che rappresenta il numero di volte in cui quella radice annulla il polinomio caratteristico. Tale molteplicità nel contesto degli autovalori prende il nome di **molteplicità algebrica**.

Diagonalizzazione di un operatore lineare

Con l'espressione "diagonalizzare un operatore lineare $F: V \rightarrow V$" s'intende trovare una base di **V** tale per cui la matrice associata ad *F* rispetto a quella base è una matrice diagonale.

Sia *F* un endomorfismo di **V** e sia $B = \{v_1, v_2, \dots, v_n\}$ una base di **V** tale per cui la matrice associata ad *F* rispetto alla base *B* è una matrice diagonale:

$$A = \begin{pmatrix} \alpha_{11} & 0 & \cdots & 0 \\ 0 & \alpha_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{nn} \end{pmatrix}$$

Questo significa che le immagini dei vettori della base *B* saranno definite come:

$$F(v_1) = \alpha_1 v_1 + 0v_2 + \cdots + 0v_n = \alpha_1 v_1$$

$$F(v_2) = 0v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \alpha_2 v_2$$

e così via.

Quindi i coefficienti $\alpha_1, \alpha_2, \dots, \alpha_n$ sono degli **autovalori** relativi agli **autovettori** della base *B*. Quindi se la matrice associata ad un operatore lineare rispetto ad una certa base è una matrice diagonale, allora ogni coefficiente sulla diagonale principale è un autovalore, e ciascun vettore di tale base è un autovettore.

Si può anche dire che *B* è una base diagonalizzante per *F* se e solo se *B* è formata da autovettori.

Se un operatore lineare è diagonalizzabile, il suo polinomio caratteristico si spezza nel prodotto di polinomi di primo grado:

$$p_A(x) = (-1)^n (x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n)$$

dove *n* è il rango della matrice associata ed $\alpha_1, \alpha_2, \dots, \alpha_n$ sono gli autovalori di *F*. Questa è una condizione necessaria per la diagonalizzabilità, ma non è sufficiente.

Teorema spettrale

Il teorema spettrale dà una condizione necessaria e sufficiente per stabilire se un operatore lineare è diagonalizzabile.

> Sia $F: V \rightarrow V$ un operatore lineare, con **V** spazio vettoriale su $\mathbb{K}$ e $dim(V) = n$. *F* è diagonalizzabile se e solo se:
> 1. il polinomio caratteristico di *F* si spezza nel prodotto di polinomi di primo grado;
> 2. per ogni autovalore λ, molteplicità algebrica e molteplicità geometrica coincidono

Casi particolari:

- se $\mathbb{K} = \mathbb{C}$, la condizione 1 è immediatamente verificata (corollario del teorema fondamentale dell'algebra)
- se $\lambda_1, \lambda_2, \dots, \lambda_n$ sono tutti distinti tra loro, la condizione 2 è immediatamente verificata;

Sia $F: V \to V$ un operatore lineare e siano $\lambda_1, \lambda_2, \dots, \lambda_s$ gli autovalori di *F* e $V_{\lambda_1}(F), V_{\lambda_2}(F), \dots, V_{\lambda_s}(F)$ gli autospazi relativi a ciascun autovalore. Allora i vettori $v_1, v_2, \dots, v_s \in V_{\lambda_j}(F)$ sono linearmente indipendenti.

Lezione del 17/12/2020, parte 2 (minuto 20 circa)

In generale, autovettori relativi ad autovalori distinti sono linearmente indipendenti.

Dimostrazione

Per $s = 1$, si ha un unico autovalore λ_1, quindi un unico autospazio $V_{\lambda_1}(F)$ e un unico autovettore v_1, che è diverso dal vettore nullo per definizione di autovettore. Quindi v_1 è linearmente indipendente.

Per $s \geq 2$ si procede per induzione.

- ipotesi: $v_1, v_2, \dots, v_{s-1}$ sono linearmente indipendenti;
- tesi: $v_1, v_2, \dots, v_s$ sono linearmente indipendenti

Siano $\alpha_1, \alpha_2, \dots, \alpha_s \in \mathbb{K}$ tali che:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_s v_s = 0$$

Si moltiplicano entrambi i membri dell'uguaglianza per λ_1:

$$\lambda_1(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_s v_s) = \lambda_1 \cdot 0$$

$$(\alpha_1 \lambda_1) v_1 + (\alpha_2 \lambda_1) v_2 + \cdots + (\alpha_s \lambda_1) v_s = 0 \quad (\star)$$

poi si applica la funzione lineare *F* ad entrambi i membri dell'uguaglianza:

$$\alpha_1 F(v_1) + \alpha_2 F(v_2) + \cdots + \alpha_s F(v_s) = 0$$

Dato che $v_1, v_2, \dots, v_s$ sono autovettori, $F(v_i)$ può essere scritto come $\lambda_i v_i$:

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_s \lambda_s v_s = 0$$

Si può quindi sottrarre quest'ultima espressione all'espressione $(\star)$:

$$(\alpha_1 \lambda_1) v_1 + \cdots + (\alpha_s \lambda_1) v_s - \alpha_1 \lambda_1 v_1 - \cdots - \alpha_s \lambda_s v_s = 0$$

e si ottiene:

$$\alpha_2(\lambda_2 - \lambda_1) v_2 + \alpha_3(\lambda_3 - \lambda_1) + \cdots + \alpha_s(\lambda_s - \lambda_1) v_s = 0$$

Sono rimasti $s - 1$ vettori, che per ipotesi induttiva sono linearmente indipendenti. Quindi i coefficienti $\alpha_2, \alpha_3, \dots, \alpha_s$ devono essere tutti uguali a 0. La tesi quindi è dimostrata.

Come conseguenza di questo teorema si ha che l'unione delle basi dei vari autospazi è un insieme di vettori linearmente indipendenti.