

# APPUNTI DI ALGEBRA LINEARE

a cura di Stefano Patrì

indirizzo e-mail: *seriegeo@yahoo.it*

Anno accademico 2008-2009

1 dicembre 2008

Queste note non hanno l'organicità di un testo perché sono da intendersi esclusivamente come integrazione dei veri e propri testi di riferimento bibliografico.

Saranno estremamente graditi commenti, osservazioni, critiche, suggerimenti e segnalazioni di eventuali errori, sviste, imprecisioni che dovessero essere rilevate, sia di carattere matematico, che anche solo semplicemente di carattere linguistico.

## Indice

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Spazi e sottospazi vettoriali</b>                                      | <b>2</b>  |
| 1.1      | Vettori linearmente dipendenti e linearmente indipendenti . . . . .       | 2         |
| 1.2      | Proprietà dei sottospazi vettoriali . . . . .                             | 5         |
| 1.3      | Prodotto interno di vettori . . . . .                                     | 7         |
| 1.4      | Spazio affine . . . . .   | 10        |
| 1.5      | Proiezioni ortogonali su sottospazi . . . . .                             | 13        |
| 1.5.1    | Metodo dei minimi quadrati . . . . .                                      | 16        |
| <b>2</b> | <b>Applicazioni lineari fra spazi vettoriali</b>                          | <b>18</b> |
| 2.1      | Matrici associate ad applicazioni lineari . . . . .                       | 21        |
| 2.2      | Cambiamenti di base e matrici associate ad applicazioni lineari . . . . . | 23        |
| 2.3      | Sottospazi invarianti per applicazioni lineari . . . . .                  | 27        |
| 2.4      | Diagonalizzazione di applicazioni lineari . . . . .                       | 30        |
| 2.5      | Proprietà delle applicazioni simmetriche . . . . .                        | 33        |
| 2.6      | Diagonalizzazione di forme quadratiche . . . . .                          | 38        |
| <b>3</b> | <b>Riduzione a forma canonica di Jordan</b>                               | <b>39</b> |
| 3.1      | Forma canonica di Jordan per matrici di ordine 2 . . . . .                | 39        |
| 3.2      | Forma canonica di Jordan per matrici di ordine 3 . . . . .                | 41        |
| 3.2.1    | Matrice con autovalore di molteplicità algebrica 2 . . . . .              | 41        |
| 3.2.2    | Autovalore di molteplicità 3 con un solo autovettore . . . . .            | 42        |
| 3.2.3    | Autovalore di molteplicità 3 con due autovettori . . . . .                | 44        |
| <b>4</b> | <b>Esercizio risolto</b>  | <b>46</b> |

# 1 Spazi e sottospazi vettoriali

Uno spazio vettoriale, indicato con  $V$ , sul campo reale  $\mathcal{R}$  è un insieme di elementi, chiamati *vettori*, per i quali sono definite le due seguenti operazioni, denominate *addizione di vettori* e *moltiplicazione di uno scalare per un vettore*, indicate con i simboli rispettivamente  $+$ ,  $\cdot$  (anche se poi nel seguito, poiché il significato della scrittura è ugualmente chiaro, indicheremo l'operazione di moltiplicazione omettendo sempre il simbolo  $\cdot$ )

$$\begin{array}{ll} \forall \mathbf{v}_1, \mathbf{v}_2 \in V & \text{si ha } \mathbf{v}_1 + \mathbf{v}_2 \in V \\ \forall k \in \mathcal{R} \text{ e } \forall \mathbf{v} \in V & \text{si ha } k \cdot \mathbf{v} \in V \end{array}$$

in modo che valgano le proprietà riportate nel testo.

Con tali operazioni si possono costruire quelle che vengono denominate *combinazioni lineari di vettori*: la seguente scrittura

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 + \dots + x_k \mathbf{v}_k$$

prende il nome di *combinazione lineare dei vettori*  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$  mediante coefficienti  $x_1, x_2, x_3, \dots, x_k$ .

## 1.1 Vettori linearmente dipendenti e linearmente indipendenti

L'analisi delle combinazioni lineari permette di studiare le proprietà di un insieme di vettori: assegnati i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$ , appartenenti tutti al medesimo spazio vettoriale  $V$  e diversi dal *vettore nullo*, ci domandiamo se ve ne è tra essi qualcuno esprimibile come combinazione lineare degli altri  $k - 1$  vettori.

Supponendo che  $\mathbf{v}_1$  sia un vettore esprimibile come combinazione lineare degli altri  $k - 1$  vettori, si avrà allora

$$\mathbf{v}_1 = \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \dots + \alpha_k \mathbf{v}_k$$

in cui i coefficienti  $\alpha_i$  non possono essere tutti pari a zero altrimenti si avrebbe che il vettore  $\mathbf{v}_1$  è il vettore nullo escluso dall'insieme dei vettori assegnati.

Supponendo quindi che  $\alpha_2$  sia l'unico coefficiente diverso da zero nella combinazione lineare che esprime il vettore  $\mathbf{v}_1$ , avremo allora

$$\mathbf{v}_1 = \alpha_2 \mathbf{v}_2 + 0\mathbf{v}_3 + 0\mathbf{v}_4 + \dots + 0\mathbf{v}_k$$

da cui, esplicitando il vettore  $\mathbf{v}_2$ , deduciamo che vale anche la combinazione lineare

$$\mathbf{v}_2 = \frac{1}{\alpha_2} \mathbf{v}_1 + 0\mathbf{v}_3 + 0\mathbf{v}_4 + \dots + 0\mathbf{v}_k$$

Quando il vettore  $\mathbf{v}_1$  si esprime come combinazione lineare degli altri vettori  $\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \dots, \mathbf{v}_k$ , si dice che  $\mathbf{v}_1$  *dipende* dai vettori  $\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \dots, \mathbf{v}_k$ , perché la conoscenza dei vettori  $\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \dots, \mathbf{v}_k$ , permette, appunto mediante combinazione lineare, di ottenere il vettore  $\mathbf{v}_1$ .

Utilizzando questa terminologia, possiamo concludere allora che quando tra i  $k$  vettori assegnati tutti diversi dal vettore nullo  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$  ce n'è uno che dipende dagli altri  $k - 1$  vettori, allora sono almeno due i vettori dipendenti dagli altri  $k - 1$  vettori.

Assegnati dunque  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$ , appartenenti tutti al medesimo spazio vettoriale  $V$  e diversi dal *vettore nullo*, si possono verificare due differenti situazioni

- alcuni di essi (almeno due, come visto) o anche tutti sono esprimibili come combinazione lineare degli altri  $k - 1$  vettori, ovvero dipendono dagli altri  $k - 1$  vettori;
- nessun vettore tra essi è esprimibile come combinazione lineare degli altri  $k - 1$  vettori, ovvero nessun vettore tra essi dipende dagli altri  $k - 1$  vettori.

Diciamo allora che quando i  $k$  vettori si trovano nella prima situazione, essi formano un *insieme di vettori linearmente dipendente*, quando invece si trovano nella seconda situazione, essi formano un *insieme di vettori linearmente indipendente*.

Si sottolinea dunque che affinché  $k$  vettori formino un insieme di vettori linearmente dipendente, non occorre che tutti i vettori dell'insieme siano necessariamente dipendenti dagli altri, ma è sufficiente che anche solo due lo siano.

Per riconoscere se  $k$  vettori formano un insieme di vettori linearmente dipendente o indipendente, si risolve l'equazione vettoriale

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + \dots + x_k\mathbf{v}_k = \mathbf{0}$$

nelle incognite  $x_1, x_2, x_3, \dots, x_k$ , la quale può avere infinite soluzioni oppure la soluzione unica, detta *soluzione banale*,  $x_1 = x_2 = x_3 = \dots = x_k = 0$ .

Se la soluzione è solo quella banale, allora diremo che i vettori formano un insieme di vettori linearmente indipendente perché dall'uguaglianza

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3 + \dots + 0\mathbf{v}_k = \mathbf{0}$$

non è possibile esprimere nessun vettore come combinazione lineare degli altri  $k - 1$  vettori.

Se invece l'equazione vettoriale possiede infinite soluzioni e supponiamo, per fissare le idee, che una di esse sia

$$0\mathbf{v}_1 + 3\mathbf{v}_2 - 4\mathbf{v}_3 + 0\mathbf{v}_4 + 0\mathbf{v}_5 + \dots + 2\mathbf{v}_{k-1} + 0\mathbf{v}_k = \mathbf{0}$$

allora si ha i vettori con coefficiente diverso da zero dipendono dagli altri vettori, ovvero possono essere espressi come combinazione lineare degli altri  $k - 1$  vettori *portati* al secondo membro, mentre i vettori con coefficiente zero non possono essere espressi come combinazione lineare degli altri  $k - 1$  vettori perché i vettori *portati* al secondo membro non possono poi essere *divisi* per il coefficiente zero del vettore *rimasto* al primo membro.

In ogni caso, quando l'equazione vettoriale dei  $k$  vettori combinati linearmente per dare il vettore nullo possiede infinite soluzioni, allora i  $k$  vettori formano un insieme di vettori linearmente dipendente perché in tali soluzioni ci saranno almeno due coefficienti diversi da zero, i quali appunto permettono di esprimere almeno due vettori come combinazione lineare degli altri  $k - 1$  vettori.

**Definizione:** dato uno spazio vettoriale  $V$ , si dice che esso ha *dimensione*  $n$ , dove  $n$  è un numero intero positivo, se è possibile trovare in  $V$   $n$  vettori linearmente indipendenti e se, comunque si prendano  $n + 1$  vettori, questi formano sempre un insieme linearmente dipendente.  $\triangle$

**Definizione:** dato uno spazio vettoriale  $V$  di dimensione finita  $n$ , si chiama *base* di  $V$  un qualunque insieme di  $n$  vettori di  $V$  linearmente indipendenti.  $\triangle$

**Teorema:** dato uno spazio vettoriale  $V$  di dimensione finita  $n$ , gli  $n$  vettori di base permettono di esprimere in modo unico qualsiasi vettore di  $V$ .

**Dimostrazione:** siano  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$  gli  $n$  vettori di base di  $V$  e consideriamo un qualsiasi altro vettore  $\mathbf{v} \in V$ . Poiché l'insieme degli  $n + 1$  vettori  $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$  è linearmente dipendente, allora si ha che l'equazione vettoriale

$$\alpha \mathbf{v} + x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 + \dots + x_k \mathbf{v}_k = \mathbf{0}$$

avrà infinite soluzioni in cui i coefficienti siano o tutti diversi da zero o alcuni pari a zero e altri (almeno due) diversi da zero.

Se tali coefficienti non fossero tutti diversi da zero, allora il coefficiente  $\alpha$  risulta comunque diverso da zero perché se fosse  $\alpha = 0$  sarebbero poi nulli, in quanto abbinati a vettori linearmente indipendenti, anche tutti gli altri coefficienti  $x_i$ .

Ma se il coefficiente  $\alpha$  è diverso da zero, allora segue, per quanto detto in precedenza, che il vettore  $\mathbf{v}$  è appunto esprimibile come combinazione lineare

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \dots + \alpha_n \mathbf{v}_n$$

Per dimostrare l'unicità di tale combinazione lineare, supponiamo per assurdo che valga anche

$$\mathbf{v} = \alpha'_1 \mathbf{v}_1 + \alpha'_2 \mathbf{v}_2 + \alpha'_3 \mathbf{v}_3 + \dots + \alpha'_n \mathbf{v}_n$$

Sottraendo membro a membro tale combinazione lineare dalla precedente, si ottiene l'equazione

$$\mathbf{0} = (\alpha_1 - \alpha'_1) \mathbf{v}_1 + (\alpha_2 - \alpha'_2) \mathbf{v}_2 + (\alpha_3 - \alpha'_3) \mathbf{v}_3 + \dots + (\alpha_n - \alpha'_n) \mathbf{v}_n$$

da cui, in virtù dell'indipendenza dei vettori di base  $\mathbf{v}_i$ , ricaviamo le uguaglianze  $\alpha_1 = \alpha'_1, \alpha_2 = \alpha'_2, \dots, \alpha_n = \alpha'_n$ , cioè otteniamo che la combinazione lineare avente coefficienti  $\alpha'_i$  coincide con la combinazione lineare avente coefficienti  $\alpha_i$ . *C.d.d.*

**Osservazione:** in virtù della possibilità di esprimere ciascun vettore dello spazio vettoriale  $V$  come combinazione lineare degli  $n$  vettori di base, i vettori di base vengono denominati anche un *sistema di generatori*.  $\triangle$

## 1.2 Proprietà dei sottospazi vettoriali

Dato uno spazio vettoriale  $V$ , il sottoinsieme  $S \subset V$  viene chiamato *sottospazio vettoriale* di  $V$  se per ogni  $\mathbf{u}, \mathbf{v} \in S$  e per ogni  $\alpha, \beta \in \mathcal{R}$  vale la condizione

$$\alpha\mathbf{u} + \beta\mathbf{v} \in S$$

separabile nelle due condizioni

$$\mathbf{u} + \mathbf{v} \in S \quad \text{e} \quad \alpha\mathbf{u} \in S$$

Osserviamo quindi che un sottospazio vettoriale  $S$  dello spazio vettoriale  $V$  è a sua volta uno spazio vettoriale se  $S$  non viene considerato come sottoinsieme dello spazio vettoriale  $V$ .

**Teorema:** sia  $V$  uno spazio vettoriale e siano  $S_1, S_2$  due suoi sottospazi vettoriali, allora il sottoinsieme  $S$  di  $V$  dato dall'intersezione  $S = S_1 \cap S_2$  è anch'esso un sottospazio vettoriale di  $V$ .

**Dimostrazione:** dobbiamo mostrare che se  $\mathbf{u}, \mathbf{v} \in S$  e  $\alpha, \beta \in \mathcal{R}$ , allora segue  $\alpha\mathbf{u} + \beta\mathbf{v} \in S$ . Si ha infatti

$$\begin{aligned} \mathbf{u}, \mathbf{v} \in S &\implies \mathbf{u}, \mathbf{v} \in S_1 \quad \text{e} \quad \mathbf{u}, \mathbf{v} \in S_2 \implies \\ &\implies \alpha\mathbf{u} + \beta\mathbf{v} \in S_1 \quad \text{e} \quad \alpha\mathbf{u} + \beta\mathbf{v} \in S_2 \implies \\ &\implies \alpha\mathbf{u} + \beta\mathbf{v} \in S_1 \cap S_2 \implies \alpha\mathbf{u} + \beta\mathbf{v} \in S \end{aligned}$$

C.d.d.

Abbiamo dimostrato dunque che l'intersezione di due sottospazi vettoriali dello spazio vettoriale  $V$  è anch'essa sottospazio vettoriale di  $V$ . Vediamo ora quale altro sottospazio vettoriale di  $V$  si può costruire con due sottospazi vettoriali di  $V$ .

**Definizione:** dati i sottospazi vettoriali  $U, W$  di  $V$ , il sottoinsieme  $S$  di  $V$  contenente tutti i vettori esprimibili come somma di un qualsiasi vettore  $\mathbf{u} \in U$  con un qualsiasi vettore  $\mathbf{w} \in W$ , ovvero il sottoinsieme

$$S = \{\mathbf{u} + \mathbf{w}, \text{ con } \mathbf{u} \in U, \mathbf{w} \in W\}$$

viene denominato *sottoinsieme somma* dei sottospazi  $U$  e  $W$  e viene indicato con il simbolo  $S = U + W$ .  $\triangle$

**Teorema:** dati i sottospazi vettoriali  $U$  e  $W$  di  $V$ , il sottoinsieme  $S$  di  $V$ , somma dei sottospazi vettoriali  $U$  e  $W$ , è anch'esso un sottospazio vettoriale di  $V$ .

**Dimostrazione:** per dimostrare che  $S$  è un sottospazio vettoriale di  $V$ , dobbiamo dimostrare che se  $\mathbf{v}_1, \mathbf{v}_2 \in S$ , allora segue

$$\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 \in S, \quad \forall \alpha, \beta \in \mathcal{R}$$

Abbiamo dunque

$$\begin{aligned} \mathbf{v}_1, \mathbf{v}_2 \in S &\implies \mathbf{v}_1 = \mathbf{u}_1 + \mathbf{w}_1 \quad \text{e} \quad \mathbf{v}_2 = \mathbf{u}_2 + \mathbf{w}_2 \implies \\ &\implies \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = (\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) + (\alpha\mathbf{w}_1 + \beta\mathbf{w}_2) \in S \end{aligned}$$

perché  $\mathbf{u}_1, \mathbf{u}_2 \in U$  e  $\mathbf{w}_1, \mathbf{w}_2 \in W$  e inoltre  $U, W$  sono sottospazi di  $V$ .  $C.d.d.$

Vogliamo analizzare ora che relazione c'è fra l'intersezione di sottospazi, la somma di sottospazi e l'indipendenza lineare di vettori.

**Teorema:** dati i sottospazi vettoriali  $U$  e  $W$  di  $V$ , si ha l'equivalenza fra le seguenti tre proposizioni

1. per ogni  $\mathbf{u} \in U$  e ogni  $\mathbf{w} \in W$  si ha  $\alpha\mathbf{u} + \beta\mathbf{w} = \mathbf{0}$  solo se  $\alpha = \beta = 0$ ;
2.  $U \cap W = \{\mathbf{0}\}$ ;
3. se un qualsiasi vettore  $\mathbf{v} \in U + W$  si esprime come combinazione lineare dei vettori  $\mathbf{u} \in U$  e  $\mathbf{w} \in W$  attraverso la relazione  $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{w}$ , allora i coefficienti  $\alpha, \beta$  sono unici.

**Dimostrazione:** per dimostrare l'equivalenza delle tre proposizioni, dimostreremo che dalla prima discende la seconda, dalla seconda discende la terza e dalla terza discende la prima.

Dimostriamo dunque che dalla prima proposizione si ottiene la seconda. A tale scopo ammettiamo che nell'intersezione vi sia un vettore diverso dal vettore nullo e vediamo quale conseguenza se ne trae

$$\begin{aligned} \mathbf{u} \in U \cap W, \text{ con } \mathbf{u} \neq \mathbf{0} &\implies \mathbf{u} \in U \text{ e } \mathbf{u} \in W \implies \\ &\implies \mathbf{u} \in U \text{ e } -\mathbf{u} \in W \implies 1\mathbf{u} + 1(-\mathbf{u}) = \mathbf{0} \end{aligned}$$

dove quest'ultimo risultato è assurdo perché per ipotesi la combinazione lineare dei vettori  $\mathbf{u} \in U$  e  $-\mathbf{u} \in W$  non può dare come risultato il vettore nullo con i coefficienti pari a 1. Quindi concludiamo che l'intersezione  $U \cap W$  non può contenere nessun altro vettore oltre al vettore nullo,

Dimostriamo ora che dalla seconda proposizione si ottiene la terza. A tale scopo e senza perdita di generalità, supponiamo che con  $\alpha \neq \alpha'$  e  $\beta \neq \beta'$  si abbia

$$\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{w} \quad \text{e} \quad \mathbf{v} = \alpha'\mathbf{u} + \beta'\mathbf{w}$$

Allora, sottraendo membro a membro la prima e la seconda di queste due equazioni, si ottiene

$$(\alpha - \alpha')\mathbf{u} + (\beta - \beta')\mathbf{w} = \mathbf{0}$$

ovvero  $\mathbf{u} = h\mathbf{w}$ , da cui segue la conseguenza assurda

$$\mathbf{u} \in U \quad \text{e} \quad \mathbf{u} = h\mathbf{w} \in W$$

cioè che esiste il vettore diverso dal nullo  $\mathbf{u} \in U \cap W$ .

Per dimostrare che dalla terza proposizione si ottiene la prima, osserviamo immediatamente che, essendo  $0\mathbf{u} + 0\mathbf{w} = \mathbf{0}$ , se si avesse  $\alpha\mathbf{u} + \beta\mathbf{w} = \mathbf{0}$  con  $\alpha, \beta$  non entrambi nulli, allora seguirebbe che il vettore  $\mathbf{v} = \mathbf{0}$  può essere ottenuto con due combinazioni lineari distinte.

C.d.d.

**Definizione:** quando due sottospazi vettoriali  $U$  e  $W$  di  $V$  sono tali che valga  $U \cap W = \{\mathbf{0}\}$  (ovvero una delle proposizioni del teorema precedente e dunque tutte e tre), allora  $U$  e  $W$  si dicono *sottospazi indipendenti* di  $V$  e il sottospazio somma  $S = U + W$  viene chiamato *somma diretta dei sottospazi  $U$  e  $W$* , indicata con il simbolo  $S = U \oplus W$ .  $\triangle$

Aggiungiamo, omettendo la dimostrazione, che la dimensione del sottospazio vettoriale  $S = U \oplus W$  è pari alla somma della dimensione del sottospazio  $V$  e della dimensione del sottospazio  $W$ .

### 1.3 Prodotto interno di vettori

Finora abbiamo studiato le proprietà degli spazi e sottospazi vettoriali dovute alla linearità. Passiamo ora dunque a studiare le cosiddette proprietà metriche degli spazi e sottospazi vettoriali. A tale scopo introduciamo il concetto di *prodotto interno* di due vettori.

**Definizione:** dato uno spazio vettoriale  $V$  sul campo reale  $\mathcal{R}$  (più avanti vedremo come si modifica la definizione quando il campo sarà quello dei numeri complessi), definiamo *prodotto interno* l'applicazione, indicata con il simbolo  $\langle \cdot, \cdot \rangle$

$$\langle \cdot, \cdot \rangle : V \longrightarrow V$$

tale che valgano le seguenti tre proprietà

1.  $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ ;
2.  $\langle \alpha \mathbf{u} + \beta \mathbf{w}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{w}, \mathbf{v} \in V, \quad \forall \alpha, \beta \in \mathcal{R}$ ;
3.  $\langle \mathbf{0}, \mathbf{0} \rangle = 0$  e  $\langle \mathbf{u}, \mathbf{u} \rangle > 0, \quad \forall \mathbf{u} \neq \mathbf{0}$ .  $\triangle$

**Osservazione:** se lo spazio vettoriale  $V$  è lo spazio vettoriale  $S_n$  delle  $n$ -ple di numeri reali, allora il prodotto interno coincide con il prodotto scalare che possiamo esprimere con la relazione (tra le tante che si potrebbero scegliere!)

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^n u_k v_k = u_1 v_1 + u_2 v_2 + u_3 v_3 + \dots + u_n v_n$$

dove le  $u_k$  e le  $v_k$  sono le componenti rispettivamente dei vettori  $\mathbf{u}$  e  $\mathbf{v}$ .  $\triangle$

**Definizione:** uno spazio vettoriale  $V$  in cui sia definito un prodotto interno, viene denominato *spazio euclideo*. Due vettori  $\mathbf{u}, \mathbf{v}$  di uno spazio euclideo si dicono *ortogonali* se vale  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$   $\triangle$

**Definizione:** in virtù della terza proprietà del prodotto interno, dato un vettore  $\mathbf{v} \in V$ , con  $V$  spazio euclideo, il numero, indicato con il simbolo  $\| \cdot \|$ , ottenuto eseguendo

$$\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

esiste sempre reale e non negativo e viene denominato *norma* del vettore  $\mathbf{v}$ .  $\triangle$

**Definizione:** dato uno spazio euclideo  $V$  di dimensione  $n$ , una base di  $V$  data da  $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$  si dice *ortonormale* se i suoi vettori possiedono tutti norma pari a 1 e inoltre sono a due a due ortogonali.

La proprietà di *ortonormalità* della base  $\mathcal{B}$  viene espressa formalmente dicendo che fra i suoi vettori sussiste la relazione  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$ , dove il simbolo  $\delta_{ij}$ , denominato *delta di Kronecker*, vale 1 se  $i = j$ , mentre vale 0 se  $i \neq j$ .  $\triangle$

Vediamo ora che relazione intercorre fra i prodotti interni e le norme di vettori: dimostriamo quella che prende il nome di *disuguaglianza di Cauchy-Schwarz*.

**Teorema:** dati due vettori  $\mathbf{u}, \mathbf{v}$  di uno spazio euclideo  $V$ , vale la disuguaglianza, detta disuguaglianza di Cauchy-Schwarz

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1 \quad (1)$$

**Dimostrazione:** dati i due vettori  $\mathbf{u}, \mathbf{v} \in V$ , costruiamo il vettore

$$\mathbf{w} = \lambda \mathbf{u} + \mathbf{v}$$

la cui norma, per la terza proprietà del prodotto interno, non è mai negativa per qualunque valore del coefficiente  $\lambda$ . Sviluppando allora la norma di  $\mathbf{w}$ , abbiamo dunque, in virtù della proprietà distributiva del prodotto interno

$$0 \leq \|\mathbf{w}\|^2 = \langle \mathbf{w}, \mathbf{w} \rangle = \langle \lambda \mathbf{u} + \mathbf{v}, \lambda \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle \lambda^2 + 2 \langle \mathbf{u}, \mathbf{v} \rangle \lambda + \langle \mathbf{v}, \mathbf{v} \rangle$$

che è un polinomio di secondo grado nella variabile  $\lambda$ .

Dalla ben nota proprietà dei polinomi di secondo grado per cui essi, quando hanno il coefficiente di  $\lambda^2$ , forniscono risultati mai negativi se e solo se per il loro discriminante vale la disuguaglianza  $b^2 - 4ac \leq 0$ , ricaviamo allora

$$4 \langle \mathbf{u}, \mathbf{v} \rangle^2 - 4 \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \leq 0$$

da cui, separando i due addendi, si ottiene

$$\frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2} \leq 1$$

e quindi, estraendo le radici quadrate di ambo i membri, la disuguaglianza (1).

*C.d.d.*

Utilizzando il prodotto interno di due vettori, vediamo una condizione per stabilire se due vettori (facilmente estensibile al caso di più di due vettori) sono linearmente indipendenti.

**Teorema:** due vettori  $\mathbf{u}, \mathbf{v}$  di uno spazio euclideo  $V$  sono linearmente indipendenti se e solo se vale

$$\det \begin{pmatrix} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{u}, \mathbf{v} \rangle \\ \langle \mathbf{v}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix} \neq 0 \quad (2)$$



**Dimostrazione:** dimostriamo che se  $\mathbf{u}, \mathbf{v}$  sono linearmente indipendenti, allora il determinante nella (2) è diverso da zero.

A tale scopo consideriamo l'equazione vettoriale  $\alpha\mathbf{u} + \beta\mathbf{v} = \mathbf{0}$  che avrà l'unica soluzione  $\alpha = \beta = 0$ . Eseguendo il prodotto interno di ambo i membri di tale equazione vettoriale prima per il vettore  $\mathbf{u}$  e poi per il vettore  $\mathbf{v}$ , si ottengono le due equazioni che formano il sistema (scritto in simbolismo matriciale)

$$\begin{pmatrix} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{u}, \mathbf{v} \rangle \\ \langle \mathbf{v}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Poiché tale sistema ha, come detto, soluzione unica  $\alpha = \beta = 0$ , allora segue che, per il teorema di Rouché-Capelli, il determinante della matrice dei coefficienti è diverso da zero, cioè la (2).

Se, viceversa, vale la relazione (2), allora, per il teorema di Rouché-Capelli, il sistema

$$\begin{pmatrix} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{u}, \mathbf{v} \rangle \\ \langle \mathbf{v}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

possiede l'unica soluzione  $\alpha = \beta = 0$ , da cui segue dunque che anche l'equazione vettoriale  $\alpha\mathbf{u} + \beta\mathbf{v} = \mathbf{0}$  possiede l'unica soluzione  $\alpha = \beta = 0$ , ovvero i vettori  $\mathbf{u}, \mathbf{v}$  sono linearmente indipendenti. E' poi immediata l'estensione di tale teorema al caso di più di due vettori. *C.d.d.*

**Definizione:** dati due vettori  $\mathbf{u}, \mathbf{v}$  di uno spazio euclideo  $V$ , il determinante

$$\det \begin{pmatrix} \langle \mathbf{u}, \mathbf{u} \rangle & \langle \mathbf{u}, \mathbf{v} \rangle \\ \langle \mathbf{v}, \mathbf{u} \rangle & \langle \mathbf{v}, \mathbf{v} \rangle \end{pmatrix}$$

viene denominato *determinante di Gram* dei vettori  $\mathbf{u}, \mathbf{v}$ . E' immediata l'estensione del determinante di Gram al caso di più di due vettori.  $\triangle$

**Teorema:** se  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$  appartenenti ad uno spazio euclideo  $V$  sono a due a due ortogonali, allora essi formano un insieme linearmente indipendente.

**Dimostrazione:** se i  $k$  vettori assegnati sono a due a due ortogonali, allora per ogni coppia di indici  $i, j$  distinti si ha  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ .

Applicando dunque il criterio del determinante di Gram a tale insieme di vettori, si ottiene

$$\det \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \langle \mathbf{v}_1, \mathbf{v}_3 \rangle & \dots & \langle \mathbf{v}_1, \mathbf{v}_k \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \langle \mathbf{v}_2, \mathbf{v}_3 \rangle & \dots & \langle \mathbf{v}_2, \mathbf{v}_k \rangle \\ \langle \mathbf{v}_3, \mathbf{v}_1 \rangle & \langle \mathbf{v}_3, \mathbf{v}_2 \rangle & \langle \mathbf{v}_3, \mathbf{v}_3 \rangle & \dots & \langle \mathbf{v}_3, \mathbf{v}_k \rangle \\ \dots & \dots & \dots & \dots & \dots \\ \langle \mathbf{v}_k, \mathbf{v}_1 \rangle & \langle \mathbf{v}_k, \mathbf{v}_2 \rangle & \langle \mathbf{v}_k, \mathbf{v}_3 \rangle & \dots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle \end{pmatrix} =$$

$$= \det \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & 0 & 0 & \dots & 0 \\ 0 & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & 0 & \dots & 0 \\ 0 & 0 & \langle \mathbf{v}_3, \mathbf{v}_3 \rangle & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \langle \mathbf{v}_k, \mathbf{v}_k \rangle \end{pmatrix} \neq 0$$

da cui deduciamo che i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$  a due a due ortogonali formano un insieme linearmente indipendente. *C.d.d.*

**Definizione:** dato lo spazio euclideo  $V$  di dimensione  $n$  e un suo sottospazio vettoriale  $S$  di dimensione  $k$ , con  $k < n$ , nel quale siano fissati i  $k$  vettori di base  $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_k$ , il sottoinsieme di  $V$ , indicato con  $S^\perp$ , definito come

$$S^\perp = \{ \mathbf{v} \in V, \text{ tale che } \langle \mathbf{v}, \mathbf{u}_1 \rangle = \langle \mathbf{v}, \mathbf{u}_2 \rangle = \dots = \langle \mathbf{v}, \mathbf{u}_k \rangle = 0 \}$$

viene denominato *complemento ortogonale del sottospazio  $S$* .  $\triangle$

**Teorema:** dato uno spazio vettoriale  $V$  di dimensione  $n$  e un suo sottospazio vettoriale  $S$  di dimensione  $k < n$ , il sottoinsieme  $S^\perp$  di  $V$  è anch'esso un sottospazio vettoriale di  $V$  di dimensione  $n - k$ .

**Dimostrazione:** indicati con  $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_k$  i vettori di base di  $S$ , dobbiamo dimostrare che se i vettori  $\mathbf{v}, \mathbf{w}$  appartengono a  $S^\perp$ , ovvero se, in base alla definizione, essi sono ortogonali a tutti i vettori  $\mathbf{u}_i$  di base per  $S$ , allora anche la loro combinazione lineare appartiene a  $S^\perp$ . A tal proposito calcoliamo

$$\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u}_1 \rangle = \alpha \langle \mathbf{v}, \mathbf{u}_1 \rangle + \beta \langle \mathbf{w}, \mathbf{u}_1 \rangle = \alpha 0 + \beta 0 = 0$$

Poiché vale analogamente  $\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u}_i \rangle = 0$ , con  $i = 2, 3, \dots, k$ , allora concludiamo che anche il vettore  $\alpha \mathbf{v} + \beta \mathbf{w}$  è ortogonale a tutti i vettori  $\mathbf{u}_i$  di base di  $S$  e dunque appartiene a  $S^\perp$ .

Omettiamo la dimostrazione, peraltro molto semplice, del fatto che la dimensione di  $S^\perp$  è pari a  $n - k$ . *C.d.d.*

## 1.4 Spazio affine

Poiché una  $n - pla$  di numeri reali rappresenta tanto le coordinate di un punto quanto le componenti di un vettore, vogliamo collegare i due concetti di punto e di vettore per capire dove sono applicati i vettori.

**Definizione:** dato uno spazio vettoriale  $V$  e un insieme  $E$  di punti, tale insieme  $E$  viene denominato *spazio affine associato allo spazio vettoriale  $V$* , se fra i punti di  $E$  e i vettori di  $V$  vi è una relazione che verifica le tre seguenti proprietà

1. per ogni coppia di punti  $P, Q \in E$ , esiste un solo vettore  $\mathbf{v} \in V$  tale che si abbia  $\mathbf{v} = \vec{PQ}$ ;
2. per ogni punto  $P \in E$  e per ogni vettore  $\mathbf{v} \in V$  esiste unico il punto  $Q \in E$  tale che si abbia  $\mathbf{v} = \vec{PQ}$ ;

3. per ogni  $P, Q, R \in E$  si ha  $\vec{PQ} + \vec{QR} = \vec{PR}$ . △

**Definizione:** se lo spazio vettoriale  $V$  ha dimensione  $n$ , allora diciamo che anche la dimensione dello spazio affine  $E$ , associato a  $V$ , è pari a  $n$ : talvolta per indicare esplicitamente che lo spazio  $E$  ha dimensione  $n$ , scriveremo  $E^n$ . △

Se ora interpretiamo le  $n$ -ple  $P(x_1, x_2, x_3, \dots, x_n)$  e  $Q(y_1, y_2, y_3, \dots, y_n)$  come punti di  $E$ , allora come regola per associare a tali punti il vettore  $\mathbf{v} \in V$  adottiamo quella per cui il vettore  $\mathbf{v} \in V$  ha componenti

$$\mathbf{v} = \vec{PQ} = (x_1 - y_1, x_2 - y_2, x_3 - y_3, \dots, x_n - y_n)$$

perché il vettore  $\mathbf{v} \in V$  così ottenuto è associato ai punti  $P, Q \in E$  attraverso una relazione che, come si vede facilmente, verifica le tre proprietà richieste.

Come è facile vedere, tale associazione del vettore  $\mathbf{v}$  ai due punti  $P$  e  $Q$  è coerente con la rappresentazione di un vettore mediante segmento orientato che congiunge l'origine con il punto avente per coordinate le componenti del vettore stesso: infatti, se l'origine  $O$  ha coordinate tutte nulle e il punto  $P$  ha coordinate  $P(x_1, x_2, x_3, \dots, x_n)$ , allora si ha

$$\mathbf{v} = \vec{OP} = (x_1 - 0, x_2 - 0, x_3 - 0, \dots, x_n - 0) = (x_1, x_2, x_3, \dots, x_n)$$

A questo punto allora possiamo scrivere le equazioni della retta  $r$  passante per i punti  $P(x_1, y_1, z_1)$  e  $Q(x_2, y_2, z_2)$  di  $E^3$ .

Tali punti individuano il vettore  $\mathbf{v} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) = (l, m, n)$  che rappresenta la direzione della retta e allora un qualsiasi punto  $R(x, y, z)$  della retta  $r$  (che cioè *describe* tutta la retta  $r$ ) è tale che il vettore  $\vec{RP}$  sia multiplo del vettore  $\mathbf{v}$ , ovvero valga la relazione dello spazio affine

$$\vec{RP} = t\mathbf{v}$$

esprimibile nelle forma, detta *equazione parametrica* della retta

$$\begin{cases} x = x_1 + tl \\ y = y_1 + tm \\ z = z_1 + tn \end{cases} \quad \text{oppure} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + t \begin{pmatrix} l \\ m \\ n \end{pmatrix}$$

Esprimendo il parametro  $t$  nelle tre equazioni parametriche

$$t = \frac{x - x_1}{l} = \frac{y - y_1}{m} = \frac{z - z_1}{n}$$

e uguagliando due dei tre membri che esprimono  $t$ , si ottiene l'equazione cartesiana della retta nella forma

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

Se qualcuna delle componenti  $l, m, n$  del vettore  $\mathbf{v}$  dovesse essere nulla, diciamo senza perdita di generalità  $l = 0$ , allora una delle due equazioni della forma cartesiana è  $x = x_1$  e l'altra è

$$\frac{y - y_1}{m} = \frac{z - z_1}{n}$$

Se fossero nulle due componenti del vettore  $\mathbf{v}$ , diciamo senza perdita di generalità  $l = m = 0$ , allora le due equazioni della forma cartesiana sono

$$\begin{cases} x = x_1 \\ y = y_1 \end{cases}$$

Scriviamo ora l'equazione di un piano immerso nello spazio affine  $E^3$ . La *giacitura* di un piano è data da due vettori linearmente indipendenti e quindi per individuare un piano occorre aggiungere alla giacitura il passaggio per un punto.

Dati dunque il punto  $P(x_0, y_0, z_0)$  e i due vettori linearmente indipendenti  $\mathbf{u} = (u_1, u_2, u_3)$  e  $\mathbf{v} = (v_1, v_2, v_3)$ , il generico punto  $Q(x, y, z)$  che *describe* tutto il piano è tale che il vettore  $\vec{PQ}$  sia combinazione lineare dei vettori  $\mathbf{u}, \mathbf{v}$ , ovvero valga la relazione dello spazio affine

$$\vec{PQ} = \alpha \mathbf{u} + \beta \mathbf{v}$$

esprimibile nelle forma, detta *equazione parametrica* del piano

$$\begin{cases} x = x_0 + \alpha u_1 + \beta v_1 \\ y = y_0 + \alpha u_2 + \beta v_2 \\ z = z_0 + \alpha u_3 + \beta v_3 \end{cases} \quad \text{oppure} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \alpha \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \beta \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Eliminando attraverso facili passaggi algebrici i parametri  $\alpha$  e  $\beta$  nelle equazioni parametriche del piano, si ottiene l'equazione cartesiana dello stesso piano avente la forma

$$ax + by + cz = d, \quad \text{con } a, b, c, d \in \mathcal{R}$$

Dunque una sola equazione cartesiana nelle indeterminate  $x, y, z$  rappresenta un piano immerso in  $E^3$  passante per un punto e avente giacitura data da due vettori linearmente indipendenti di  $S_3$ . In generale, una sola equazione cartesiana nelle indeterminate  $x_1, x_2, \dots, x_n$  rappresenta un *iperpiano* immerso in  $E^n$  passante per un punto e avente giacitura data da  $n - 1$  vettori linearmente indipendenti dello spazio vettoriale  $S_n$ .

Per concludere la trattazione delle rette e dei piani di  $E^3$ , dimostriamo che la direzione ortogonale al piano  $\pi$  di equazione cartesiana  $ax + by + cz = d$  è quella del vettore  $\mathbf{v} = (a, b, c)$ .

A tale scopo utilizziamo il piano  $\pi'$  di equazione cartesiana  $ax + by + cz = 0$ , parallelo al piano  $\pi$ , perché se il vettore  $\mathbf{v} = (a, b, c)$  è ortogonale al piano  $\pi'$ , allora  $\mathbf{v}$  sarà ortogonale anche al piano  $\pi$ .

Il vantaggio di utilizzare il piano  $\pi'$  deriva dal fatto che il piano  $\pi'$  passa per l'origine e quindi un vettore che congiunge l'origine con un punto di  $\pi'$  appartiene tutto a tale piano: poiché, come detto, un vettore che congiunge l'origine ad un punto ha componenti coincidenti con le coordinate del punto secondo estremo, allora, se il piano  $\pi'$  passa per l'origine, un punto appartenente a  $\pi'$  ha coordinate coincidenti con le componenti di un vettore appartenente a  $\pi'$ .

Pertanto, scrivendo l'equazione cartesiana del piano  $\pi'$  data da  $ax+by+cz=0$  nella forma di prodotto interno

$$\langle (a, b, c), (x, y, z) \rangle = 0$$

si ricava che, in virtù della coincidenza fra le coordinate di un punto appartenente al piano  $\pi'$  e le componenti di un vettore giacente su  $\pi'$ , la condizione di appartenenza di un punto a  $\pi'$ , data da  $ax+by+cz=0$ , coincide con la condizione di ortogonalità di un vettore di  $\pi'$  al vettore  $(a, b, c)$ , data da  $\langle (a, b, c), (x, y, z) \rangle = 0$ : quindi se il prodotto interno di ogni vettore di  $\pi'$  con il vettore  $(a, b, c)$  è pari a zero, allora concludiamo che il vettore  $(a, b, c)$  è ortogonale all'intero piano  $\pi'$  e di conseguenza a tutto il fascio di piani paralleli a  $\pi'$  fra i quali vi è il piano  $\pi$  di equazione cartesiana  $ax+by+cz=d$ .

## 1.5 Proiezioni ortogonali su sottospazi

Dato un sottospazio vettoriale  $U$  di uno spazio euclideo  $V$ , vogliamo definire la *proiezione ortogonale* di un qualsiasi vettore di  $V$  sul sottospazio  $U$ .

**Definizione:** dato un vettore  $\mathbf{v} \in V$ , con  $V$  spazio euclideo, il vettore  $\mathbf{u} \in U$ , con  $U$  sottospazio di  $V$ , viene chiamato *proiezione ortogonale* del vettore  $\mathbf{v}$  su  $U$  se il vettore  $\mathbf{v} - \mathbf{u}$  è ortogonale a tutti i vettori del sottospazio  $U$ .  $\triangle$

Vediamo ora come si può determinare la proiezione ortogonale di un certo vettore  $\mathbf{v} \in V$  sul sottospazio  $U$ .

**Teorema:** dato un vettore  $\mathbf{v} \in V$ , con  $V$  spazio euclideo, e un sottospazio  $U$  di  $V$ , esiste unico il vettore  $\mathbf{u} \in U$ , *proiezione ortogonale* del vettore  $\mathbf{v}$  su  $U$ .

**Dimostrazione:** sia  $U$  un sottospazio vettoriale di dimensione  $k$  dello spazio euclideo  $V$  e i  $k$  vettori  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$  siano una base di  $U$ .

Per dimostrare che esiste unico il vettore  $\mathbf{u} \in U$ , proiezione ortogonale del vettore  $\mathbf{v}$  su  $U$ , basta dimostrare che tale vettore  $\mathbf{u} \in U$  si può sempre ricavare e può essere sempre espresso come combinazione lineare dei vettori di base di  $U$  mediante coefficienti unici  $\alpha_1, \alpha_2, \dots, \alpha_k$ .

Sviluppiamo allora il vettore incognito  $\mathbf{u} \in U$  come combinazione lineare dei  $k$  vettori di base di  $U$  mediante coefficienti incogniti

$$\mathbf{u} = \sum_{i=1}^k \alpha_i \mathbf{u}_i = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3 + \dots + \alpha_k \mathbf{u}_k \quad (3)$$

Affinché il vettore  $\mathbf{u}$  sia la proiezione ortogonale su  $U$  del vettore  $\mathbf{v}$ , bisogna imporre, in base alla definizione di proiezione ortogonale, che il vettore  $\mathbf{v} - \mathbf{u}$

sia ortogonale a tutti i vettori di  $U$ . A tale scopo però è sufficiente imporre l'ortogonalità del vettore  $\mathbf{v} - \mathbf{u}$  ai soli  $k$  vettori di base di  $U$ , data dalle relazioni

$$\langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle = 0, \quad \forall i = 1, 2, 3, \dots, k \quad (4)$$

Consideriamo infatti un qualsiasi vettore  $\mathbf{w} \in U$  che esprimiamo come combinazione lineare dei  $k$  vettori di base di  $U$

$$\mathbf{w} = \sum_{i=1}^k \beta_i \mathbf{u}_i = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \beta_3 \mathbf{u}_3 + \dots + \beta_k \mathbf{u}_k$$

Se il vettore  $\mathbf{v} - \mathbf{u}$  è ortogonale ai  $k$  vettori di base di  $U$ , allora tale vettore è automaticamente ortogonale al generico vettore  $\mathbf{w} \in U$ , come si ricava immediatamente eseguendo il prodotto scalare

$$\langle \mathbf{v} - \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v} - \mathbf{u}, \sum_{i=1}^k \beta_i \mathbf{u}_i \rangle = \sum_{i=1}^k \beta_i \langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle = 0$$

A questo punto allora possiamo ricavare dalle equazioni (4) i coefficienti incogniti  $\alpha_1, \alpha_2, \dots, \alpha_k$  della combinazione lineare che fornisce il vettore  $\mathbf{u}$ .

L'esistenza e l'unicità della proiezione ortogonale  $\mathbf{u} \in U$  sarà allora ottenuta dall'esistenza e unicità dei coefficienti  $\alpha_j$  mediante i quali si esprime il vettore  $\mathbf{u}$  nella base del sottospazio vettoriale  $U$ .

Poiché si ha

$$\begin{aligned} \langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle &= \langle \mathbf{v} - \sum_{j=1}^k \alpha_j \mathbf{u}_j, \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle \sum_{j=1}^k \alpha_j \mathbf{u}_j, \mathbf{u}_i \rangle = \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \sum_{j=1}^k \alpha_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle, \quad \forall i = 1, 2, 3, \dots, k \end{aligned}$$

allora le equazioni (4) diventano

$$\langle \mathbf{v}, \mathbf{u}_i \rangle - \sum_{j=1}^k \alpha_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0, \quad \forall i = 1, 2, 3, \dots, k$$

ovvero, scambiando l'ordine dei vettori nei prodotti interni

$$\sum_{j=1}^k \langle \mathbf{u}_i, \mathbf{u}_j \rangle \alpha_j = \langle \mathbf{u}_i, \mathbf{v} \rangle, \quad \forall i = 1, 2, 3, \dots, k \quad (5)$$

Le equazioni (5) rappresentano il sistema scritto in forma matriciale

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & \langle \mathbf{u}_1, \mathbf{u}_3 \rangle & \dots & \langle \mathbf{u}_1, \mathbf{u}_k \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle & \langle \mathbf{u}_2, \mathbf{u}_3 \rangle & \dots & \langle \mathbf{u}_2, \mathbf{u}_k \rangle \\ \langle \mathbf{u}_3, \mathbf{u}_1 \rangle & \langle \mathbf{u}_3, \mathbf{u}_2 \rangle & \langle \mathbf{u}_3, \mathbf{u}_3 \rangle & \dots & \langle \mathbf{u}_3, \mathbf{u}_k \rangle \\ \dots & \dots & \dots & \dots & \dots \\ \langle \mathbf{u}_k, \mathbf{u}_1 \rangle & \langle \mathbf{u}_k, \mathbf{u}_2 \rangle & \langle \mathbf{u}_k, \mathbf{u}_3 \rangle & \dots & \langle \mathbf{u}_k, \mathbf{u}_k \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \langle \mathbf{u}_2, \mathbf{v} \rangle \\ \langle \mathbf{u}_3, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{v} \rangle \end{pmatrix}$$

Poiché i vettori  $\mathbf{u}_i$  costituiscono una base di  $U$ , allora essi sono linearmente indipendenti e dunque la matrice dei coefficienti di tale sistema ha determinante di Gram diverso da zero, da cui ricaviamo in definitiva che il sistema possiede sempre un'unica soluzione  $\alpha_1, \alpha_2, \dots, \alpha_k$ , la quale, sostituita nella (3), fornirà dunque il vettore  $\mathbf{u}$ , proiezione ortogonale unica del vettore  $\mathbf{v}$  sul sottospazio  $U$ . *C.d.d.*

Se la base nel sottospazio  $U$  viene scelta ortonormale, allora il primo membro del sistema (5) assume la forma

$$\sum_{j=1}^k \langle \mathbf{u}_i, \mathbf{u}_j \rangle \alpha_j = \sum_{j=1}^k \delta_{ij} \alpha_j = \delta_{i1} \alpha_1 + \delta_{i2} \alpha_2 + \delta_{i3} \alpha_3 + \dots + \delta_{ik} \alpha_k = \alpha_i$$

perché nella somma con i fattori  $\delta$  soltanto l'addendo con  $\delta_{ii}$  vale  $1\alpha_i$ , mentre tutti gli altri  $k-1$  addendi sono pari a zero in quanto contenenti la  $\delta$  con indici diversi.

Pertanto nel caso in cui nel sottospazio  $U$  si scelga una base ortonormale, il sistema (5) diventa allora molto semplicemente

$$\alpha_i = \langle \mathbf{u}_i, \mathbf{v} \rangle, \quad \forall i = 1, 2, 3, \dots, k \quad (6)$$

ovvero, nel caso particolare di base ortonormale in  $U$ , i coefficienti  $\alpha_i$  della combinazione lineare che fornisce  $\mathbf{u}$  sono uguali al prodotto interno fra il corrispondente vettore  $i$ -esimo di base  $\mathbf{u}_i$  per il vettore  $\mathbf{v}$  da proiettare su  $U$ .

Vediamo ora da quale altro punto di vista può essere considerato il problema del ricerca della proiezione ortogonale di un vettore  $\mathbf{v}$  su un sottospazio  $U$ .

**Teorema:** dato un vettore  $\mathbf{v} \in V$ , con  $V$  spazio euclideo, e un sottospazio  $U$  di  $V$ , il vettore  $\mathbf{u} \in U$ , *proiezione ortogonale* del vettore  $\mathbf{v}$  su  $U$ , è la soluzione del problema di minimizzazione

$$\min_{\mathbf{u}'} \|\mathbf{v} - \mathbf{u}'\|^2 \quad (7)$$

**Dimostrazione:** dimostriamo questo teorema solo nel caso in cui la base in  $U$  sia ortonormale, ovvero dimostriamo che risolvendo il problema di minimizzazione posto nella (7) si ottengono i coefficienti per il vettore  $\mathbf{u}$  dati dalle equazioni (6).

Considerando  $U$  un sottospazio di dimensione  $k$  in cui si abbia la base ortonormale contenente i vettori  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$  ed esprimendo il vettore incognito  $\mathbf{u}' \in U$  in questa base mediante coefficienti incogniti  $\alpha_1, \alpha_2, \dots, \alpha_k$ , si ottiene

$$\begin{aligned} \|\mathbf{v} - \mathbf{u}'\|^2 &= \langle \mathbf{v} - \mathbf{u}', \mathbf{v} - \mathbf{u}' \rangle = \left\langle \mathbf{v} - \sum_{i=1}^k \alpha_i \mathbf{u}_i, \mathbf{v} - \sum_{j=1}^k \alpha_j \mathbf{u}_j \right\rangle = \\ &= \|\mathbf{v}\|^2 + \left[ \sum_{j=1}^k \sum_{i=1}^k \alpha_i \alpha_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle - 2 \sum_{i=1}^k \alpha_i \langle \mathbf{u}_i, \mathbf{v} \rangle \right] = \\ &= \|\mathbf{v}\|^2 + \left[ \sum_{i=1}^k \alpha_i^2 - 2 \sum_{i=1}^k \alpha_i \langle \mathbf{u}_i, \mathbf{v} \rangle + \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{v} \rangle^2 \right] - \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{v} \rangle^2 = \end{aligned}$$

$$= \| \mathbf{v} \|^2 + \sum_{i=1}^k [\alpha_i - \langle \mathbf{u}_i, \mathbf{v} \rangle]^2 - \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{v} \rangle^2$$

dove si è utilizzato la proprietà dei vettori di una base ortonormale per cui

$$\sum_{j=1}^k \sum_{i=1}^k \alpha_i \alpha_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{j=1}^k \sum_{i=1}^k \alpha_i \alpha_j \delta_{ij} = \sum_{j=1}^k \alpha_j^2$$

Il problema (7) può essere allora riformulato scrivendo

$$\min_{\mathbf{u}'} \| \mathbf{v} - \mathbf{u}' \|^2 = \min_{\alpha_i} \left\{ \| \mathbf{v} \|^2 + \sum_{i=1}^k [\alpha_i - \langle \mathbf{u}_i, \mathbf{v} \rangle]^2 - \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{v} \rangle^2 \right\}$$

da cui si ricava che il minimo di  $\| \mathbf{v} - \mathbf{u}' \|^2$  è realizzato se i coefficienti  $\alpha_i$  annullano la base del quadrato nel secondo addendo, ovvero quando, per ogni  $i = 1, 2, \dots, k$ , si hanno le uguaglianze  $\alpha_i - \langle \mathbf{u}_i, \mathbf{v} \rangle = 0$ , che sono proprio le equazioni (6). *C.d.d.*

**Osservazione:** ogni volta dunque che ci si trova di fronte ad un problema posto nella forma della (7), tale problema può essere allora equivalentemente interpretato come un problema di ricerca della proiezione ortogonale del vettore  $\mathbf{v}$  sul sottospazio a cui deve appartenere il vettore  $\mathbf{u}'$  e può essere risolto quindi utilizzando le equazioni (5) oppure le equazioni (6) a seconda che la base in  $U$  sia, rispettivamente, non ortonormale oppure ortonormale.

### 1.5.1 Metodo dei minimi quadrati

Immaginiamo di voler determinare la relazione funzionale esistente fra le grandezze  $x, y, z$  che caratterizzano un certo fenomeno che stiamo studiando e supponiamo che la struttura di tale relazione funzionale sia

$$z = ax + by + c$$

in modo che attraverso una sequenza statistica di misurazioni si debba ricavare semplicemente il valore dei parametri  $a, b, c$ .

A prima vista si potrebbe pensare che effettuando direttamente  $n$  misurazioni empiriche  $x_i, y_i, z_i$  delle tre grandezze  $x, y, z$ , con  $i = 1, 2, \dots, n$ , si possa risolvere poi il sistema

$$\begin{cases} ax_1 + by_1 + c = z_1 \\ ax_2 + by_2 + c = z_2 \\ \dots\dots\dots \\ ax_n + by_n + c = z_n \end{cases}$$

nelle incognite  $a, b, c$  e determinare così il valore di tali parametri.

Non è difficile però rendersi conto che tale procedura è inapplicabile perché nella pratica, a causa delle indeterminazioni insite nel processo di misura, il sistema che si ottiene, in cui i valori misurati  $x_i, y_i, z_i$  sono i coefficienti delle incognite



$a, b, c$ , risulterà in generale incompatibile per il teorema di Rouché-Capelli, dal momento che se la caratteristica della matrice ad  $n$  righe e 3 colonne dei coefficienti è pari a 3, poi la matrice completa ad  $n$  righe e 4 colonne avrà in generale caratteristica 4.

Poiché dunque, date le  $n$  misurazioni delle grandezze  $x, y, z$ , non è possibile ricavare il valore dei parametri  $a, b, c$  risolvendo il sistema precedente, allora la procedura che si segue per ricavare il valore di tali parametri  $a, b, c$  è quella chiamata *metodo dei minimi quadrati*.

Questa procedura consiste nel determinare il valore dei parametri  $a, b, c$  tali che sia minimo lo scarto fra il valore  $z_i$  misurato e il valore di  $z_i$  ricavato dalla relazione funzionale  $ax_i + by_i + c$ . Poiché si è interessati allo scarto assoluto e non al suo segno, allora si cerca il valore dei parametri  $a, b, c$  tali che sia minima l'espressione

$$\begin{aligned} & \sum_{i=1}^n (z_i - ax_i - by_i - c)^2 = \\ & = (z_1 - ax_1 - by_1 - c)^2 + (z_2 - ax_2 - by_2 - c)^2 + \dots + (z_n - ax_n - by_n - c)^2 \end{aligned}$$

scrivibile quindi nella forma

$$\| \mathbf{z} - a\mathbf{x} - b\mathbf{y} - c\mathbf{1} \|^2$$

dove è stata utilizzata la notazione dei vettori  $n$ -dimensionali  $\mathbf{z} = (z_1, z_2, \dots, z_n)$ ,  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  e  $\mathbf{1} = (1, 1, \dots, 1)$ .

Poiché il problema di minimizzazione

$$\min_{a,b,c} \| \mathbf{z} - a\mathbf{x} - b\mathbf{y} - c\mathbf{1} \|^2 = \min_{a,b,c} \| \mathbf{z} - (a\mathbf{x} + b\mathbf{y} + c\mathbf{1}) \|^2$$

equivale, come visto, al problema della ricerca della proiezione ortogonale del vettore  $\mathbf{z} \in S_n$  sul sottospazio  $U$  di  $S_n$  individuato dai vettori  $\mathbf{x}$ ,  $\mathbf{y}$  e  $\mathbf{1}$ , allora per determinare il valore dei parametri  $a, b, c$  applichiamo le equazioni (5) perché utilizziamo come vettori di base in  $U$  i vettori  $\mathbf{x}$ ,  $\mathbf{y}$  e  $\mathbf{1}$  che non formano, in generale, una base ortonormale

$$\begin{cases} \langle \mathbf{x}, \mathbf{x} \rangle a + \langle \mathbf{x}, \mathbf{y} \rangle b + \langle \mathbf{x}, \mathbf{1} \rangle c = \langle \mathbf{x}, \mathbf{z} \rangle \\ \langle \mathbf{x}, \mathbf{y} \rangle a + \langle \mathbf{y}, \mathbf{y} \rangle b + \langle \mathbf{y}, \mathbf{1} \rangle c = \langle \mathbf{y}, \mathbf{z} \rangle \\ \langle \mathbf{x}, \mathbf{z} \rangle a + \langle \mathbf{y}, \mathbf{z} \rangle b + \langle \mathbf{z}, \mathbf{1} \rangle c = \langle \mathbf{z}, \mathbf{z} \rangle \end{cases}$$

Si verifica immediatamente che tali equazioni, scritte con la notazione vettoriale, corrispondono alle equazioni che si ricavano risolvendo il problema di minimizzazione mediante derivate parziali.

E' immediato poi estendere tale discussione al caso generale in cui le grandezze misurabili per via statistica siano  $x_1, x_2, \dots, x_n$  e la struttura della relazione funzionale che lega tali grandezze sia

$$z = a_1x_1 + a_2x_2 + \dots + a_nx_n + a_{n+1}$$

dove si debba ricavare da misurazioni dirette delle grandezze  $x_1, x_2, \dots, x_n$  il valore degli  $n + 1$  parametri  $a_1, a_2, \dots, a_n, a_{n+1}$ .

## 2 Applicazioni lineari fra spazi vettoriali

Dati due spazi vettoriali  $V, W$  di dimensione finita, un'applicazione lineare fra gli spazi vettoriali  $V$  e  $W$ , indicata con il simbolo  $L : V \longrightarrow W$ , è una legge che ad un vettore  $\mathbf{v} \in V$  associa, come risultato, un vettore  $\mathbf{w} \in W$ , verificando le seguenti due proprietà

$$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) \quad \text{e} \quad L(k\mathbf{v}) = kL(\mathbf{v})$$

per ogni  $\mathbf{v}_1, \mathbf{v}_2 \in V$  e per ogni  $k \in \mathcal{R}$ , esprimibili nell'unica relazione

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$$

per ogni  $\mathbf{v}_1, \mathbf{v}_2 \in V$  e per ogni  $\alpha, \beta \in \mathcal{R}$ . Quando si scrive  $L(\mathbf{v}) = \mathbf{w}$ , si dice che il vettore  $\mathbf{w} \in W$  è *immagine* del vettore  $\mathbf{v} \in V$ .

**Teorema:** un'applicazione lineare  $L : V \longrightarrow W$  trasforma il vettore nullo dello spazio vettoriale  $V$  nel vettore nullo dello spazio vettoriale  $W$ .

**Dimostrazione:** sia  $\mathbf{v}$  un generico vettore di  $V$ , segue allora

$$L(\mathbf{0}) = L(\mathbf{v} - \mathbf{v}) = L(\mathbf{v}) - L(\mathbf{v}) = \mathbf{0}$$

dove, con ovvio significato dei simboli, il vettore nullo che appare in  $L(\mathbf{0})$  appartiene allo spazio  $V$ , mentre il vettore nullo finale appartiene allo spazio  $W$ .

*C.d.d.*

**Definizione:** siano dati due spazi vettoriali  $V$  e  $W$  e un'applicazione lineare  $L : V \longrightarrow W$ , il sottoinsieme di  $V$ , indicato con  $N(L)$ , dato da

$$N(L) = \{\mathbf{v} \in V \text{ tale che } L(\mathbf{v}) = \mathbf{0}\}$$

viene denominato *nucleo* dell'applicazione lineare  $L$ . △

Osserviamo innanzitutto che il nucleo di un'applicazione lineare non è mai l'insieme vuoto perché almeno il vettore nullo di  $V$  viene trasformato nel vettore nullo di  $W$ : resta dunque da stabilire se il vettore nullo è l'unico vettore di  $V$  ad essere trasformato nel vettore nullo di  $W$ , o se altri (infiniti) vettori di  $V$  vengono pure trasformati nel vettore nullo di  $W$ .

**Teorema:** siano dati due spazi vettoriali  $V$  e  $W$  e un'applicazione lineare  $L : V \longrightarrow W$ , allora il nucleo di  $L$  è un sottospazio vettoriale di  $V$ .

**Dimostrazione:** se il nucleo di  $L$  contiene solo il vettore nullo di  $V$ , allora tale nucleo è un sottospazio vettoriale e la dimostrazione è conclusa.

Se il nucleo non contiene solo il vettore nullo, allora dobbiamo dimostrare che se due generici vettori  $\mathbf{v}_1, \mathbf{v}_2$  appartengono al nucleo di  $L$ , allora anche una loro qualsiasi combinazione lineare appartiene al nucleo di  $L$ .

Se dunque  $\mathbf{v}_1, \mathbf{v}_2 \in N(L)$ , allora si ha  $L(\mathbf{v}_1) = L(\mathbf{v}_2) = \mathbf{0}$  da cui, per la linearità dell'applicazione lineare  $L$ , segue

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) = \alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0}$$

ovvero il vettore  $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2$  appartiene al nucleo di  $L$ .

*C.d.d.*

**Teorema:** un'applicazione lineare  $L : V \longrightarrow W$  è iniettiva (cioè trasforma ogni coppia di vettori distinti in  $V$  in una coppia di vettori distinti in  $W$ ) se e solo se il suo nucleo  $N(L)$  contiene il solo vettore nullo.

**Dimostrazione:** supponiamo che  $N(L)$  contenga soltanto il vettore nullo, allora dobbiamo dimostrare che due generici vettori  $\mathbf{v}_1, \mathbf{v}_2 \in V$  distinti non possono dare due risultati uguali. Infatti per  $\mathbf{v}_1 \neq \mathbf{v}_2$  si ha

$$L(\mathbf{v}_1) - L(\mathbf{v}_2) = L(\mathbf{v}_1 - \mathbf{v}_2) = L(\mathbf{v}) \neq \mathbf{0}$$

dove  $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2 \neq \mathbf{0}$  perché  $\mathbf{v}_1$  e  $\mathbf{v}_2$  sono distinti e  $L(\mathbf{v}) \neq \mathbf{0}$  perché  $\mathbf{v} \neq \mathbf{0}$  non appartiene al nucleo dove c'è solo il vettore nullo.

Quindi, poiché si è ottenuto  $L(\mathbf{v}_1) - L(\mathbf{v}_2) \neq \mathbf{0}$ , allora segue  $L(\mathbf{v}_1) \neq L(\mathbf{v}_2)$ .

Viceversa, se l'applicazione  $L$  è iniettiva, allora dobbiamo dimostrare che nel nucleo non ci può essere nessun vettore diverso dal vettore nullo.

Se infatti per assurdo un certo vettore  $\mathbf{v} \neq \mathbf{0}$  fosse presente nel nucleo, ovvero si avesse  $L(\mathbf{v}) = \mathbf{0}$  con  $\mathbf{v} \neq \mathbf{0}$ , allora si avrebbe  $\mathbf{v} + \mathbf{v}_1 \neq \mathbf{v}_1$  e inoltre

$$L(\mathbf{v} + \mathbf{v}_1) = L(\mathbf{v}) + L(\mathbf{v}_1) = \mathbf{0} + L(\mathbf{v}_1) = L(\mathbf{v}_1)$$

ovvero  $L(\mathbf{v} + \mathbf{v}_1) = L(\mathbf{v}_1)$  che è un risultato appunto assurdo perché, per l'injectività di  $L$ , i due vettori distinti  $\mathbf{v} + \mathbf{v}_1$  e  $\mathbf{v}_1$  non possono dare il medesimo risultato. Quindi concludiamo che il nucleo di  $L$  non può contenere nessun vettore  $\mathbf{v} \neq \mathbf{0}$ .

*C.d.d.*

**Definizione:** siano dati due spazi vettoriali  $V$  e  $W$  e un'applicazione lineare  $L : V \longrightarrow W$ , il sottoinsieme di  $W$ , indicato con  $Im(L)$ , dato da

$$Im(L) = \{\mathbf{w} \in W \text{ tale che esiste } \mathbf{v} \in V \text{ da cui si ottenga } \mathbf{w} = L(\mathbf{v})\}$$

viene denominato *immagine* dell'applicazione lineare  $L$ .

$\triangle$

**Teorema:** siano dati due spazi vettoriali  $V$  e  $W$  e un'applicazione lineare  $L : V \longrightarrow W$ , allora l'immagine di  $L$  è un sottospazio vettoriale di  $W$ .

**Dimostrazione:** dobbiamo dimostrare che se due generici vettori  $\mathbf{w}_1, \mathbf{w}_2$  appartengono all'immagine di  $L$ , allora anche una loro qualsiasi combinazione lineare appartiene all'immagine di  $L$ .

Se dunque  $\mathbf{w}_1, \mathbf{w}_2 \in Im(L)$ , allora si ha che esistono  $\mathbf{v}_1, \mathbf{v}_2 \in V$  tali che  $L(\mathbf{v}_1) = \mathbf{w}_1$  e  $L(\mathbf{v}_2) = \mathbf{w}_2$ . Segue allora per la linearità di  $L$

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2$$

ovvero  $\alpha \mathbf{w}_1 + \beta \mathbf{w}_2$  appartiene all'immagine di  $L$ , perché esiste il vettore  $\alpha \mathbf{v}_1 + \beta \mathbf{v}_2$  appartenente a  $V$  tale che risulti  $L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha \mathbf{w}_1 + \beta \mathbf{w}_2$ . *C.d.d.*

**Definizione:** data un'applicazione lineare  $L : V \longrightarrow W$ , la dimensione del sottospazio vettoriale  $N(L)$  viene denominata *nullità* di  $L$  e la dimensione dell'immagine  $Im(L)$  viene denominata *rango* di  $L$ .  $\triangle$

Vogliamo concludere quest'analisi delle proprietà di un'applicazione lineare  $L : V \longrightarrow W$  con una relazione che lega la nullità e il rango di  $L$  alla dimensione dello spazio vettoriale  $V$ .

**Teorema:** data un'applicazione lineare  $L : V \longrightarrow W$ , si ha che la somma della nullità e del rango di  $L$  è pari alla dimensione di  $V$ .

**Dimostrazione:** supponiamo che  $V$  abbia dimensione pari a  $n$  e che il nucleo e l'immagine di  $L$  abbiano dimensione uguale rispettivamente ad  $h$  e  $k$ : dobbiamo dimostrare che vale la relazione  $n = h + k$ .

Consideriamo nel sottospazio  $Im(L)$  la base formata dai  $k$  vettori linearmente indipendenti  $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k\}$  e siano  $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$  vettori di  $V$  tali che  $L(\mathbf{v}_i) = \mathbf{w}_i$ , per ogni  $i = 1, 2, 3, \dots, k$ .

Verifichiamo subito che i vettori  $\mathbf{v}_i$  sono linearmente indipendenti in  $V$ . Se i vettori  $\mathbf{v}_i$  fossero linearmente dipendenti, allora si avrebbe che uno tra essi, diciamo senza perdita di generalità  $\mathbf{v}_1$ , si esprime come combinazione lineare dei rimanenti  $k - 1$  vettori, cioè

$$\mathbf{v}_1 = \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \dots + \alpha_k \mathbf{v}_k$$

da cui scaturirebbe la conseguenza assurda che anche i vettori  $\mathbf{w}_i$  sarebbero linearmente dipendenti in  $W$ , come si vede immediatamente confrontando primo e ultimo membro dell'uguaglianza

$$\begin{aligned} \mathbf{w}_1 &= L(\mathbf{v}_1) = L(\alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \dots + \alpha_k \mathbf{v}_k) = \\ &= \alpha_2 L(\mathbf{v}_2) + \alpha_3 L(\mathbf{v}_3) + \dots + \alpha_k L(\mathbf{v}_k) = \alpha_2 \mathbf{w}_2 + \alpha_3 \mathbf{w}_3 + \dots + \alpha_k \mathbf{w}_k \end{aligned}$$

Dopo aver mostrato che i vettori  $\mathbf{v}_i$  sono linearmente indipendenti in  $V$ , consideriamo un generico vettore  $\mathbf{v} \in V$  ed esprimiamo la sua immagine come combinazione lineare dei vettori di base di  $Im(L)$

$$L(\mathbf{v}) = \beta_1 \mathbf{w}_1 + \beta_2 \mathbf{w}_2 + \beta_3 \mathbf{w}_3 + \dots + \beta_k \mathbf{w}_k$$

da cui segue

$$L(\mathbf{v}) = \beta_1 L(\mathbf{v}_1) + \beta_2 L(\mathbf{v}_2) + \beta_3 L(\mathbf{v}_3) + \dots + \beta_k L(\mathbf{v}_k)$$

e quindi per la linearità di  $L$ , scrivendo tutti i vettori al primo membro

$$L(\mathbf{v} - \beta_1 \mathbf{v}_1 - \beta_2 \mathbf{v}_2 - \beta_3 \mathbf{v}_3 - \dots - \beta_k \mathbf{v}_k) = \mathbf{0}$$

Da quest'ultima equazione ricaviamo l'informazione che, poiché la sua immagine è il vettore nullo di  $W$ , il vettore di  $V$  dato da  $\mathbf{v} - \beta_1\mathbf{v}_1 - \beta_2\mathbf{v}_2 - \beta_3\mathbf{v}_3 - \dots - \beta_k\mathbf{v}_k$  appartiene al nucleo di  $L$  e allora lo possiamo esprimere come combinazione lineare degli  $h$  vettori linearmente indipendenti  $\mathbf{u}_i$  della base del nucleo

$$\mathbf{v} - \beta_1\mathbf{v}_1 - \beta_2\mathbf{v}_2 - \beta_3\mathbf{v}_3 - \dots - \beta_k\mathbf{v}_k = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \dots + \alpha_h\mathbf{u}_h$$

Da tale uguaglianza segue lo sviluppo del vettore  $\mathbf{v} \in V$  come combinazione lineare dei  $k + h$  vettori di  $V$  che sono  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k; \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h$

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \beta_3\mathbf{v}_3 + \dots + \beta_k\mathbf{v}_k + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \dots + \alpha_h\mathbf{u}_h$$

Poiché i  $k + h$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k; \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h$  permettono di esprimere ogni vettore  $\mathbf{v} \in V$ , allora per completare la dimostrazione, basta a questo punto dimostrare che gli stessi vettori sono linearmente indipendenti in  $V$  perché in tal modo essi costituirebbero una base di  $V$  che deve contenere  $n$  vettori quanto è la dimensione di  $V$  stesso, ovvero  $n = h + k$ .

Consideriamo allora la solita equazione vettoriale

$$\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \beta_3\mathbf{v}_3 + \dots + \beta_k\mathbf{v}_k + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \dots + \alpha_h\mathbf{u}_h = \mathbf{0}$$

e dimostriamo che essa possiede l'unica soluzione con tutti i coefficienti pari a zero.

Se infatti qualche coefficiente  $\beta_i$  (almeno due coefficienti) fosse diverso da zero, allora facendo agire  $L$  su ambo i membri e ricordando che i vettori  $\mathbf{u}_i$  hanno come immagine il vettore nullo, in quanto appartenenti al nucleo di  $L$ , si otterrebbe l'uguaglianza

$$\beta_1L(\mathbf{v}_1) + \beta_2L(\mathbf{v}_2) + \beta_3L(\mathbf{v}_3) + \dots + \beta_kL(\mathbf{v}_k) = \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \beta_3\mathbf{w}_3 + \dots + \beta_k\mathbf{w}_k = \mathbf{0}$$

con qualche  $\beta_i \neq 0$ , che è assurda perché i vettori  $\mathbf{w}_i$  sono linearmente indipendenti in  $W$ . Se i coefficienti  $\beta_i$  sono dunque tutti nulli, allora segue che anche i coefficienti  $\alpha_j$  dei vettori  $\mathbf{u}_j$  sono tutti nulli perché i vettori  $\mathbf{u}_j$  formano una base del nucleo e pertanto sono anch'essi linearmente indipendenti.

Poiché sia i coefficienti  $\beta_i$  che i coefficienti  $\alpha_j$  sono nulli, allora concludiamo che gli  $h + k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k; \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h$  appartenenti a  $V$  formano una base di  $V$  che è uno spazio vettoriale di dimensione  $n$  e quindi si deve avere l'uguaglianza dimensionale  $n = h + k$ . *C.d.d.*

## 2.1 Matrici associate ad applicazioni lineari

Dato uno spazio vettoriale  $V$  di dimensione  $n$  e un vettore  $\mathbf{v} \in V$ , se in  $V$  viene fissata una base  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$ , allora abbiamo che rispetto a tale base il vettore  $\mathbf{v}$  risulterà rappresentato dall' $n$ -pla  $\mathbf{v}_{\mathcal{B}}$ , le cui componenti sono, come noto, i coefficienti  $x_i$  della combinazione lineare dei vettori di base  $\mathbf{e}_i$  che fornisce appunto il vettore  $\mathbf{v}$ , ovvero

$$\mathbf{v} = (x_1, x_2, x_3, \dots, x_n)_{\mathcal{B}} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + \dots + x_n\mathbf{e}_n$$

dove è stato attribuito l'indice  $\mathcal{B}$  all' $n$ -pla degli  $x_i$  per intendere che tale  $n$ -pla è riferita alla base  $\mathcal{B}$ . Tale indice verrà comunque omesso ogniqualvolta sia inequivocabile la base di riferimento.

Dati ora due spazi vettoriali  $V$  e  $W$ , rispettivamente di dimensione  $n$  e  $m$ , in cui siano state fissate le relative basi, vogliamo mostrare che ad un'applicazione lineare  $L : V \rightarrow W$  si può sempre associare una matrice, indicata con  $\mathcal{A}$ , in modo che l'azione dell'applicazione lineare sui vettori dello spazio vettoriale  $V$  si realizzi mediante il prodotto righe per colonne della matrice associata per la colonna del vettore che si vuole trasformare.

Siano dunque  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$  e  $\tilde{\mathcal{B}} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \dots, \mathbf{f}_m\}$  rispettivamente una base di  $V$  e di  $W$ . Per sapere come l'applicazione lineare  $L$  trasforma tutti i vettori  $\mathbf{v} \in V$  è sufficiente conoscere l'azione di  $L$  solo sui vettori della base  $\mathcal{B}$ .

Poiché l'immagine di un vettore di base  $\mathbf{e}_i$  appartiene allo spazio  $W$ , allora ogni  $L(\mathbf{e}_i) \in W$  è un vettore esprimibile come combinazione lineare dei vettori di base di  $W$ . Per assegnare dunque un'applicazione lineare  $L : V \rightarrow W$ , dobbiamo specificare come essa trasforma gli  $n$  vettori di base di  $V$ .

Date allora le seguenti relazioni, dette *leggi di trasformazione dei vettori di base* di  $V$

$$\begin{cases} L(\mathbf{e}_1) = a_{11}\mathbf{f}_1 + a_{21}\mathbf{f}_2 + a_{31}\mathbf{f}_3 + \dots + a_{m1}\mathbf{f}_m \\ L(\mathbf{e}_2) = a_{12}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + a_{32}\mathbf{f}_3 + \dots + a_{m2}\mathbf{f}_m \\ L(\mathbf{e}_3) = a_{13}\mathbf{f}_1 + a_{23}\mathbf{f}_2 + a_{33}\mathbf{f}_3 + \dots + a_{m3}\mathbf{f}_m \\ \dots \\ L(\mathbf{e}_n) = a_{1n}\mathbf{f}_1 + a_{2n}\mathbf{f}_2 + a_{3n}\mathbf{f}_3 + \dots + a_{mn}\mathbf{f}_m \end{cases}$$

dimostriamo che con esse è possibile trasformare per linearità ogni vettore  $\mathbf{v} \in V$  avente componenti note  $\mathbf{v} = (x_1, x_2, x_3, \dots, x_n)_{\mathcal{B}}$ . Sostituendo le leggi di trasformazione assegnate, si ha infatti

$$\begin{aligned} L(\mathbf{v}) &= L(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + \dots + x_n\mathbf{e}_n) = \\ &= x_1L(\mathbf{e}_1) + x_2L(\mathbf{e}_2) + x_3L(\mathbf{e}_3) + \dots + x_nL(\mathbf{e}_n) = \\ &= x_1(a_{11}\mathbf{f}_1 + a_{21}\mathbf{f}_2 + \dots + a_{m1}\mathbf{f}_m) + x_2(a_{12}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + \dots + a_{m2}\mathbf{f}_m) + \\ &+ x_3(a_{13}\mathbf{f}_1 + a_{23}\mathbf{f}_2 + \dots + a_{m3}\mathbf{f}_m) + \dots + x_n(a_{1n}\mathbf{f}_1 + a_{2n}\mathbf{f}_2 + \dots + a_{mn}\mathbf{f}_m) = \\ &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)\mathbf{f}_1 + (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n)\mathbf{f}_2 + \\ &+ (a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n)\mathbf{f}_3 + \dots + (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n)\mathbf{f}_m \end{aligned}$$

ovvero la sola conoscenza dei coefficienti  $a_{ij}$  nelle leggi di trasformazione  $L(\mathbf{e}_i)$  permette di trasformare ogni vettore  $\mathbf{v} = (x_1, x_2, x_3, \dots, x_n) \in V$ .

Se poniamo  $L(\mathbf{v}) = \mathbf{w}$ , dove  $\mathbf{w} \in W$  è il vettore immagine avente componenti date da  $\mathbf{w} = (y_1, y_2, y_3, \dots, y_m)_{\tilde{\mathcal{B}}}$ , allora dalle due uguaglianze

$$\begin{aligned} \mathbf{w} = L(\mathbf{v}) &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)\mathbf{f}_1 + (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n)\mathbf{f}_2 + \\ &+ (a_{31}x_1 + a_{32}x_2 + \dots + a_{3n}x_n)\mathbf{f}_3 + \dots + (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n)\mathbf{f}_m \end{aligned}$$

e

$$\mathbf{w} = L(\mathbf{v}) = (y_1, y_2, y_3, \dots, y_m)_{\tilde{\mathcal{B}}} = y_1 \mathbf{f}_1 + y_2 \mathbf{f}_2 + y_3 \mathbf{f}_3 + \dots + y_m \mathbf{f}_m$$

seguono, per confronto fra i coefficienti del medesimo vettore  $\mathbf{f}_j$ , le relazioni fra le componenti  $x_1, x_2, x_3, \dots, x_n$  del vettore  $\mathbf{v} \in V$  e le componenti  $y_1, y_2, y_3, \dots, y_m$  del vettore immagine  $\mathbf{w} \in W$ , dette *equazioni della trasformazione L*

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n \\ y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n \\ \dots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{mn}x_n \end{cases}$$

Tali equazioni della trasformazione  $L$ , come si vede immediatamente, si possono ottenere anche mediante prodotto righe per colonne della matrice  $\mathcal{A}$

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

detta *matrice associata all'applicazione L rispetto alle fissate basi*, per il vettore disposto in colonna delle componenti  $x_1, x_2, x_3, \dots, x_n$  del vettore  $\mathbf{v} \in V$ , ovvero il vettore immagine  $\mathbf{w} \in W$  si ottiene dal vettore  $\mathbf{v} \in V$  eseguendo  $\mathbf{w}_{\tilde{\mathcal{B}}} = \mathcal{A}\mathbf{v}_{\mathcal{B}}$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

dove si può osservare che la matrice  $\mathcal{A}$  possiede tante colonne quante sono le componenti del vettore  $\mathbf{v} \in V$ .

Confrontando la matrice  $\mathcal{A}$  (associata all'applicazione  $L$ ) con le leggi di trasformazione dei vettori di base di  $V$ , si osserva, per concludere, che i coefficienti  $a_{11}, a_{21}, a_{31}, \dots, a_{m1}$  dell'immagine  $L(\mathbf{e}_1)$  del vettore di base  $\mathbf{e}_1$  formano la prima colonna della matrice  $\mathcal{A}$ , i coefficienti  $a_{12}, a_{22}, a_{32}, \dots, a_{m2}$  dell'immagine  $L(\mathbf{e}_2)$  del vettore di base  $\mathbf{e}_2$  formano la seconda colonna della matrice  $\mathcal{A}$ , e così via.

## 2.2 Cambiamenti di base e matrici associate ad applicazioni lineari

Quando si effettua un cambiamento di base in uno spazio vettoriale  $V$ , ovvero si passa dalla base  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$  alla nuova base  $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \dots, \mathbf{e}'_n\}$ ,

allora il vettore  $\mathbf{v}$  risulterà rappresentato da una nuova  $n$ -pla  $\mathbf{v}_{\mathcal{B}'}$ , le cui componenti sono, analogamente a prima, i nuovi coefficienti  $x'_j$  della nuova combinazione lineare che fornisce lo stesso vettore  $\mathbf{v}$ , calcolata, in questo secondo caso, con i vettori  $\mathbf{e}'_j$  della base nuova  $\mathcal{B}'$ , ovvero

$$\mathbf{v} = (x'_1, x'_2, x'_3, \dots, x'_n)_{\mathcal{B}'} = x'_1 \mathbf{e}'_1 + x'_2 \mathbf{e}'_2 + x'_3 \mathbf{e}'_3 + \dots + x'_n \mathbf{e}'_n$$

Poiché abbiamo che il vettore  $\mathbf{v}$  deve essere ottenuto sia come combinazione lineare dei vettori della base  $\mathcal{B}$  che come combinazione lineare dei vettori della nuova base  $\mathcal{B}'$ , non è sorprendente allora che il vettore  $\mathbf{v}$  risulti rappresentato da due  $n$ -ple diverse, ciascuna riferita alla relativa base, in modo che comunque valga l'uguaglianza

$$\mathbf{v} = \mathbf{v}_{\mathcal{B}} = \mathbf{v}_{\mathcal{B}'}$$

in virtù della quale possiamo considerare il vettore  $\mathbf{v}$  come un *ente assoluto* di cui ovviamente cambia la rappresentazione a seconda dei vettori di base utilizzati per esprimerlo come combinazione lineare.

Il problema consiste quindi nel determinare le componenti  $\mathbf{v}_{\mathcal{B}'}$  del vettore  $\mathbf{v}$  rispetto alla nuova base  $\mathcal{B}'$  quando siano note le componenti di  $\mathbf{v}$  rispetto alla prima base  $\mathcal{B}$ .

Esprimendo i vettori  $\mathbf{e}'_j$  della nuova base  $\mathcal{B}'$  come combinazione lineare dei vettori  $\mathbf{e}_i$  della prima base  $\mathcal{B}$

$$\begin{cases} \mathbf{e}'_1 = c_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + a_{31}\mathbf{e}_3 + \dots + a_{n1}\mathbf{e}_n \\ \mathbf{e}'_2 = c_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + a_{32}\mathbf{e}_3 + \dots + a_{n2}\mathbf{e}_n \\ \mathbf{e}'_3 = c_{13}\mathbf{e}_1 + a_{23}\mathbf{e}_2 + a_{33}\mathbf{e}_3 + \dots + a_{n3}\mathbf{e}_n \\ \dots \\ \mathbf{e}'_n = c_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + a_{3n}\mathbf{e}_3 + \dots + a_{nn}\mathbf{e}_n \end{cases}$$

si ottiene la matrice  $\mathcal{T}$ , detta *matrice del cambio di base*, data da

$$\mathcal{T} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \dots & c_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & c_{n3} & \dots & c_{nn} \end{pmatrix}$$

in cui, come si verifica immediatamente, la prima colonna contiene i coefficienti della combinazione lineare che fornisce il vettore  $\mathbf{e}'_1$ , la seconda colonna contiene i coefficienti della combinazione lineare che fornisce il vettore  $\mathbf{e}'_2$ , e così via.

A questo punto la relazione che intercorre fra le componenti del vettore  $\mathbf{v}$  rispetto alle due basi è data da

$$\mathbf{v}_{\mathcal{B}'} = \mathcal{T}^{-1} \mathbf{v}_{\mathcal{B}} \quad (8)$$

ovvero le nuove componenti  $\mathbf{v}_{\mathcal{B}'}$  rispetto alla nuova base si ottengono moltiplicando righe per colonne la matrice inversa di quella del cambio di base per la  $n$ -pla delle componenti  $\mathbf{v}_{\mathcal{B}}$  rispetto alla prima base  $\mathcal{B}$ .



Vogliamo ora capire in che modo il cambiamento di base negli spazi vettoriali si ripercuote sulla matrice associata ad una data applicazione lineare  $L : V \longrightarrow W$ .

Abbiamo visto che rispetto alle basi  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$ , fissata in  $V$ , e  $\tilde{\mathcal{B}} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \dots, \mathbf{f}_m\}$ , fissata in  $W$ , l'azione dell'applicazione lineare  $L(\mathbf{v}) = \mathbf{w}$ , con  $\mathbf{v} \in V$  e  $\mathbf{w} \in W$ , si realizza attraverso il prodotto righe per colonne  $\mathcal{A}\mathbf{v}_{\mathcal{B}} = \mathbf{w}_{\tilde{\mathcal{B}}}$ .

Se quindi effettuiamo un cambio di base nei due spazi vettoriali  $V$  e  $W$ , passando in  $V$  dalla base  $\mathcal{B}$  alla base  $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \dots, \mathbf{e}'_n\}$  con la matrice  $\mathcal{T}_V$  e in  $W$  dalla base  $\tilde{\mathcal{B}}$  alla base  $\tilde{\mathcal{B}}' = \{\mathbf{f}'_1, \mathbf{f}'_2, \mathbf{f}'_3, \dots, \mathbf{f}'_m\}$  con la matrice  $\mathcal{T}_W$ , e se imponiamo la condizione che la medesima applicazione lineare  $L$  di prima trasformi il medesimo vettore  $\mathbf{v}$  nel medesimo vettore  $\mathbf{w}$  di prima, essendo i vettori  $\mathbf{v}$  e  $\mathbf{w}$  rappresentati nelle nuove basi, allora dovrà accadere che la trasformazione lineare agisca attraverso il prodotto righe per colonne  $\mathcal{A}'\mathbf{v}_{\mathcal{B}'} = \mathbf{w}_{\tilde{\mathcal{B}}'}$ , dove la matrice  $\mathcal{A}'$  sarà diversa rispetto alla matrice  $\mathcal{A}$  perché le due matrici  $\mathcal{A}$  e  $\mathcal{A}'$  moltiplicano le due  $n$ -ple diverse che sono rispettivamente  $\mathbf{v}_{\mathcal{B}}$  e  $\mathbf{v}_{\mathcal{B}'}$ .

Il problema consiste pertanto nel determinare la relazione che intercorre fra la matrice  $\mathcal{A}'$ , associata all'applicazione lineare  $L : V \longrightarrow W$  rispetto alle basi, rispettivamente,  $\mathcal{B}'$  e  $\tilde{\mathcal{B}}'$ , e la matrice  $\mathcal{A}$ , associata all'applicazione lineare  $L$  rispetto alle basi, rispettivamente,  $\mathcal{B}$  e  $\tilde{\mathcal{B}}$ .

Se nella relazione  $\mathcal{A}'\mathbf{v}_{\mathcal{B}'} = \mathbf{w}_{\tilde{\mathcal{B}}'}$  sostituiamo la relazione (8), otteniamo

$$\mathcal{A}'\mathcal{T}_V^{-1}\mathbf{v}_{\mathcal{B}} = \mathcal{T}_W^{-1}\mathbf{w}_{\tilde{\mathcal{B}}}$$

da cui, moltiplicando per la matrice  $\mathcal{T}_W$  a sinistra in ambo i membri, si ricava la relazione

$$\mathcal{T}_W\mathcal{A}'\mathcal{T}_V^{-1}\mathbf{v}_{\mathcal{B}} = \mathcal{T}_W\mathcal{T}_W^{-1}\mathbf{w}_{\tilde{\mathcal{B}}} = \mathbf{w}_{\tilde{\mathcal{B}}}$$

Confrontando quest'ultima relazione ottenuta  $(\mathcal{T}_W\mathcal{A}'\mathcal{T}_V^{-1})\mathbf{v}_{\mathcal{B}} = \mathbf{w}_{\tilde{\mathcal{B}}}$  con la relazione iniziale relativa alle basi  $\mathcal{B}$  e  $\tilde{\mathcal{B}}$  data da  $\mathcal{A}\mathbf{v}_{\mathcal{B}} = \mathbf{w}_{\tilde{\mathcal{B}}}$ , si ricava l'uguaglianza matriciale  $\mathcal{T}_W\mathcal{A}'\mathcal{T}_V^{-1} = \mathcal{A}$  che con la moltiplicazione per le matrici  $\mathcal{T}_W^{-1}$  da sinistra e  $\mathcal{T}_V$  da destra, può essere invertita e posta nella forma

$$\mathcal{A}' = \mathcal{T}_W^{-1}\mathcal{A}\mathcal{T}_V \quad (9)$$

che è la relazione mediante la quale si ricava appunto la matrice  $\mathcal{A}'$  quando si conosca la matrice  $\mathcal{A}$ . Osserviamo che se in uno dei due spazi vettoriali la base non cambia, allora la matrice  $\mathcal{T}$  del cambio di base in quello spazio vettoriale coincide con la matrice identità del medesimo spazio.

Ricordando che la matrice  $\mathcal{A}$  associata ad  $L$  relativamente alle basi  $\mathcal{B}$  e  $\tilde{\mathcal{B}}$

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

si ottiene disponendo in colonna i coefficienti delle leggi di trasformazione dei vettori di base di  $V$  date da

$$\begin{cases} L(\mathbf{e}_1) = a_{11}\mathbf{f}_1 + a_{21}\mathbf{f}_2 + a_{31}\mathbf{f}_3 + \dots + a_{m1}\mathbf{f}_m \\ L(\mathbf{e}_2) = a_{12}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + a_{32}\mathbf{f}_3 + \dots + a_{m2}\mathbf{f}_m \\ L(\mathbf{e}_3) = a_{13}\mathbf{f}_1 + a_{23}\mathbf{f}_2 + a_{33}\mathbf{f}_3 + \dots + a_{m3}\mathbf{f}_m \\ \dots \\ L(\mathbf{e}_n) = a_{1n}\mathbf{f}_1 + a_{2n}\mathbf{f}_2 + a_{3n}\mathbf{f}_3 + \dots + a_{mn}\mathbf{f}_m \end{cases}$$

è importante precisare che, analogamente, la matrice  $\mathcal{A}'$  ottenuta con la (9) avente elementi

$$\mathcal{A}' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & a'_{23} & \dots & a'_{2n} \\ a'_{31} & a'_{32} & a'_{33} & \dots & a'_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a'_{m1} & a'_{m2} & a'_{m3} & \dots & a'_{mn} \end{pmatrix}$$

ha per colonne i coefficienti delle leggi di trasformazione dei vettori di base di  $V$  relativamente alle basi nuove  $\mathcal{B}'$  e  $\tilde{\mathcal{B}}'$ , ovvero

$$\begin{cases} L(\mathbf{e}'_1) = a'_{11}\mathbf{f}'_1 + a'_{21}\mathbf{f}'_2 + a'_{31}\mathbf{f}'_3 + \dots + a'_{m1}\mathbf{f}'_m \\ L(\mathbf{e}'_2) = a'_{12}\mathbf{f}'_1 + a'_{22}\mathbf{f}'_2 + a'_{32}\mathbf{f}'_3 + \dots + a'_{m2}\mathbf{f}'_m \\ L(\mathbf{e}'_3) = a'_{13}\mathbf{f}'_1 + a'_{23}\mathbf{f}'_2 + a'_{33}\mathbf{f}'_3 + \dots + a'_{m3}\mathbf{f}'_m \\ \dots \\ L(\mathbf{e}'_n) = a'_{1n}\mathbf{f}'_1 + a'_{2n}\mathbf{f}'_2 + a'_{3n}\mathbf{f}'_3 + \dots + a'_{mn}\mathbf{f}'_m \end{cases}$$

Nel caso in cui si abbia  $L : V \longrightarrow V$ , se si effettua un medesimo cambio di base sia nello spazio  $V$  come *spazio di partenza* che nello spazio  $V$  come *spazio di arrivo*, dalla base  $\mathcal{B}$  alla base  $\mathcal{B}'$  attraverso la medesima matrice  $\mathcal{T}$  del cambio di base, allora la relazione (9) assume ovviamente la forma senza indici

$$\mathcal{A}' = \mathcal{T}^{-1}\mathcal{A}\mathcal{T} \quad (10)$$

Per concludere osserviamo che anche quando si abbia  $L : V \longrightarrow V$ , ovvero anche quando i due spazi vettoriali coincidano, si può sempre immaginare che nello spazio vettoriale  $V$  come *spazio di partenza* vi sia una base  $\mathcal{B}_1$  e nel medesimo spazio  $V$  come *spazio di arrivo* vi sia un'altra base  $\mathcal{B}_2$  diversa da  $\mathcal{B}_1$ .

Se ora effettuiamo un cambio di base nello spazio  $V$  come *spazio di partenza* dalla base  $\mathcal{B}_1$  alla base  $\mathcal{B}'_1$  mediante la matrice  $\mathcal{T}_1$  ed effettuiamo un cambio di base nello spazio  $V$  come *spazio di arrivo* dalla base  $\mathcal{B}_2$  alla base  $\mathcal{B}'_2$  mediante la matrice  $\mathcal{T}_2$ , allora la relazione (9) assume la forma

$$\mathcal{A}' = \mathcal{T}_2^{-1}\mathcal{A}\mathcal{T}_1 \quad (11)$$

in cui si riconosce che la matrice del cambio di base nello spazio di partenza si trova sempre a destra e quella del cambio di base nello spazio d'arrivo si trova sempre invertita a sinistra della matrice  $\mathcal{A}$ .

### 2.3 Sottospazi invarianti per applicazioni lineari

Dopo che in uno spazio vettoriale  $V$  di dimensione  $n$  è stata fissata una base  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$ , consideriamo un'applicazione lineare dallo spazio  $V$  in se stesso, cioè  $L : V \rightarrow V$ , rappresentata da una matrice  $\mathcal{A}$ .

**Definizione:** data l'applicazione lineare  $L : V \rightarrow V$ , un vettore  $\mathbf{v} \in V$  diverso dal vettore nullo viene denominato *autovettore* dell'applicazione  $L$  relativamente all'*autovalore*  $\lambda$  se vale la relazione, chiamata *equazione agli autovalori* o *equazione secolare*

$$L(\mathbf{v}) = \lambda \mathbf{v} \quad (12)$$

dove  $\lambda$  è un certo coefficiente appartenente all'insieme dei numeri reali.  $\triangle$

Un autovettore  $\mathbf{v}$  dunque è un vettore diverso dal vettore nullo che dall'applicazione lineare viene trasformato in un vettore appartenente alla medesima retta a cui appartiene  $\mathbf{v}$  stesso; l'autovalore relativo ad un autovettore  $\mathbf{v}$  è il coefficiente che attraverso la moltiplicazione per l'autovettore  $\mathbf{v}$ , fornisce l'immagine di  $\mathbf{v}$ .

Scrivendo  $\mathbf{v} = \mathcal{I}\mathbf{v}$ , dove con  $\mathcal{I}$  si indica la *matrice identità* di  $V$ , la (12) può essere posta nella forma

$$(L - \lambda\mathcal{I})\mathbf{v} = \mathbf{0}$$

dalla quale ricaviamo che l'autovettore  $\mathbf{v} \neq \mathbf{0}$  è un vettore che dalla nuova applicazione  $L - \lambda\mathcal{I}$  viene trasformato nel vettore nullo, ovvero è un vettore appartenente al nucleo dell'applicazione  $L - \lambda\mathcal{I}$ .

Ricordando che quando un'applicazione lineare contiene vettori diversi dal vettore nullo nel proprio nucleo, essa non è iniettiva, ovvero non è invertibile, allora possiamo concludere che un valore  $\lambda \in \mathcal{R}$  è un autovalore dell'applicazione  $L$  se l'applicazione  $L - \lambda\mathcal{I}$  non è invertibile, cioè se è nullo il determinante della matrice  $\mathcal{A} - \lambda\mathcal{I}$ , dove  $\mathcal{A}$  è la matrice associata all'applicazione  $L$  rispetto alla base  $\mathcal{B}$  di  $V$ .

Imponendo dunque

$$\det(\mathcal{A} - \lambda\mathcal{I}) = 0$$

si ottiene un'equazione algebrica di grado  $n$  pari alla dimensione di  $V$ , che possiede le soluzioni  $\lambda_1, \lambda_2, \dots, \lambda_k$  aventi le molteplicità algebriche rispettivamente  $m_a(\lambda_i)$  tali che valga la relazione  $m_a(\lambda_1) + m_a(\lambda_2) + \dots + m_a(\lambda_k) = n$ .

L'equazione algebrica  $\det(\mathcal{A} - \lambda\mathcal{I}) = 0$  viene chiamata *equazione caratteristica* della matrice  $\mathcal{A}$  e il polinomio  $\mathcal{P}_n(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I})$  avente grado  $n$  viene chiamato *polinomio caratteristico* della matrice  $\mathcal{A}$ .

Ci poniamo ora la questione di cosa accade al polinomio caratteristico di una matrice  $\mathcal{A}$  se si effettua un cambiamento di base.

**Teorema:** il polinomio caratteristico  $\mathcal{P}_n(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I})$  di una matrice  $\mathcal{A}$  è invariante per cambiamento di base.

**Dimostrazione:** dobbiamo dimostrare che effettuando un cambiamento di base in  $V$  rappresentato da una matrice  $\mathcal{T}$ , si ha  $\det(\mathcal{A}' - \lambda\mathcal{I}) = \det(\mathcal{A} - \lambda\mathcal{I})$ .

Si ha infatti

$$\det(\mathcal{A}' - \lambda\mathcal{I}) = \det(\mathcal{T}^{-1}\mathcal{A}\mathcal{T} - \lambda\mathcal{I}) = \det(\mathcal{T}^{-1}\mathcal{A}\mathcal{T} - \lambda\mathcal{T}^{-1}\mathcal{I}\mathcal{T}) =$$

$$= \det[\mathcal{T}^{-1}(\mathcal{A} - \lambda\mathcal{I})\mathcal{T}] = (\det \mathcal{T}^{-1}) [\det(\mathcal{A} - \lambda\mathcal{I})] (\det \mathcal{T}) = \det(\mathcal{A} - \lambda\mathcal{I})$$

perché  $\mathcal{T}^{-1}\mathcal{I}\mathcal{T} = \mathcal{I}$ ,  $[\det(\mathcal{T}^{-1})] = (\det \mathcal{T})^{-1}$  e inoltre vale la regola (stabilita dal teorema di Binet) sul determinante della matrice prodotto di due matrici  $M_1, M_2$  data da  $\det(M_1M_2) = \det M_1 \det M_2$ . *C.d.d.*

Sostituendo un autovalore  $\lambda = \lambda_i$  nel sistema parametrico, si ottiene il sistema numerico

$$(\mathcal{A} - \lambda_i\mathcal{I})\mathbf{v} = \mathbf{0}$$

che, per il teorema di Rouché-Capelli, avrà  $\infty^h$  soluzioni, con  $h \geq 1$ , perché la caratteristica della matrice  $\mathcal{A} - \lambda_i\mathcal{I}$  dei coefficienti è minore di  $n$ .

**Definizione:** il sottoinsieme di  $V$ , indicato con  $E(\lambda_i)$ , dato da

$$E(\lambda_i) = \{\mathbf{v} \in V, \text{ tale che } L(\mathbf{v}) = \lambda_i\mathbf{v}\}$$

viene denominato *autospatio relativo all'autovalore*  $\lambda_i$ .  $\triangle$

**Teorema:** l'autospatio  $E(\lambda_i)$  relativo ad un autovalore  $\lambda_i$  di un'applicazione lineare  $L : V \longrightarrow V$  è un sottospazio vettoriale di  $V$  avente dimensione, indicata con  $h_i$ , tale che risulti  $h_i \leq m_a(\lambda_i)$ .

**Dimostrazione:** per dimostrare che  $E(\lambda_i)$  è un sottospazio vettoriale, consideriamo due vettori  $\mathbf{v}_1, \mathbf{v}_2 \in E(\lambda_i)$ , cioè tali che  $L(\mathbf{v}_1) = \lambda_i\mathbf{v}_1$  e  $L(\mathbf{v}_2) = \lambda_i\mathbf{v}_2$ , e verifichiamo che anche la loro generica combinazione lineare appartiene ad  $E(\lambda_i)$ .

Si ha infatti

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) = \alpha\lambda_i\mathbf{v}_1 + \beta\lambda_i\mathbf{v}_2 = \lambda_i(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2)$$

ovvero  $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 \in E(\lambda_i)$ .

La dimensione  $h_i$  di  $E(\lambda_i)$  risulta essere, per il teorema di Rouché-Capelli, pari alla differenza fra la dimensione  $n$  di  $V$  e la caratteristica della matrice  $\mathcal{A} - \lambda_i\mathcal{I}$ .

Per dimostrare che  $h_i$  non può superare la molteplicità algebrica  $m_a(\lambda_i)$  del corrispondente autovalore  $\lambda_i$ , consideriamo, per fissare le idee, il caso particolare in cui l'autovalore  $\lambda_i$  abbia molteplicità algebrica 1 e lo spazio  $V$  abbia dimensione pari a 3. Per la teoria delle equazioni algebriche, il polinomio caratteristico della matrice  $\mathcal{A}$ , scritto in forma scomposta, è allora in questo caso

$$\mathcal{P}_3(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I}) = (\lambda - \lambda_i) Q_2(\lambda)$$

dove  $Q_2(\lambda)$  è un polinomio di secondo grado che fornirà gli altri due autovalori.

A questo punto supponiamo per assurdo che il sistema  $(\mathcal{A} - \lambda_i\mathcal{I})\mathbf{v} = \mathbf{0}$ , ottenuto sostituendo  $\lambda = \lambda_i$ , abbia  $\infty^2$  soluzioni, ovvero che l'autospatio  $E(\lambda_i)$  sia un sottospazio di dimensione 2 con vettori di base  $\mathbf{v}_1, \mathbf{v}_2$ .

Se quindi effettuiamo un cambiamento di base dalla base iniziale, supposta canonica, alla base  $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}\}$  e ridefiniamo l'applicazione lineare  $L$  mediante

le leggi di trasformazione dei vettori della base  $\mathcal{B}'$ , in cui  $\mathbf{v}_1, \mathbf{v}_2$  sono autovettori, allora otteniamo

$$\begin{cases} L(\mathbf{v}_1) = \lambda_i \mathbf{v}_1 \\ L(\mathbf{v}_2) = \lambda_i \mathbf{v}_2 \\ L(\mathbf{w}) = a'_{13} \mathbf{v}_1 + a'_{23} \mathbf{v}_2 + a'_{33} \mathbf{w} \end{cases}$$

da cui, incolonnando i coefficienti, scaturisce la nuova matrice  $\mathcal{A}'$  associata all'applicazione  $L$  data da

$$\mathcal{A}' = \begin{pmatrix} \lambda_i & 0 & a'_{13} \\ 0 & \lambda_i & a'_{23} \\ 0 & 0 & a'_{33} \end{pmatrix}$$

Tale matrice  $\mathcal{A}'$  è assurda perchè, come si verifica immediatamente, il suo polinomio caratteristico è

$$\mathcal{P}_3(\lambda) = \det(\mathcal{A} - \lambda \mathcal{I}) = (\lambda - \lambda_i)^2 (a'_{33} - \lambda)$$

mentre anche la matrice  $\mathcal{A}'$  deve avere il medesimo polinomio caratteristico invariante  $\mathcal{P}_3(\lambda) = (\lambda - \lambda_i) Q_2(\lambda)$  della matrice  $\mathcal{A}$ .

Quindi l'autovalore  $\lambda_i$  avente molteplicità algebrica 1 non può dar luogo ad un autospazio di dimensione maggiore di 1 e tale dimostrazione si estende facilmente a qualunque dimensione e a qualunque molteplicità algebrica, in modo che si possa concludere che la dimensione dell'autospazio  $E(\lambda_i)$  non può mai essere maggiore della molteplicità algebrica dell'autovalore  $\lambda_i$ . *C.d.d.*

Poiché un vettore  $\mathbf{v}$  dell'autospazio  $E(\lambda_i)$  di un'applicazione lineare  $L$  ha come immagine ancora un vettore dell'autospazio  $E(\lambda_i)$ , ovvero

$$L(\mathbf{v}) \in E(\lambda_i), \quad \forall \mathbf{v} \in E(\lambda_i)$$

allora un autospazio  $E(\lambda_i)$  di un'applicazione  $L$  viene chiamato *sottospazio invariante* per l'applicazione  $L$ .

Vediamo ora che relazione sussiste fra autovettori relativi ad autovalori distinti di un'applicazione lineare  $L$ .

**Teorema:** data un'applicazione lineare  $L$ , si ha che autovettori relativi ad autovalori distinti sono linearmente indipendenti.

**Dimostrazione:** consideriamo due autovettori  $\mathbf{v}_1, \mathbf{v}_2$  dell'applicazione  $L$  che corrispondano ai due autovalori distinti rispettivamente  $\lambda_1, \lambda_2$ .

Per dimostrare il teorema dobbiamo dimostrare che se i due autovettori  $\mathbf{v}_1, \mathbf{v}_2$  fossero linearmente dipendenti, ovvero se valesse la relazione  $\mathbf{v}_1 = k\mathbf{v}_2$ , allora si arriverebbe ad un'uguaglianza assurda. Infatti se valesse  $\mathbf{v}_1 = k\mathbf{v}_2$ , dalle due equazioni secolari con  $\mathbf{v}_1$  e  $\mathbf{v}_2$  si otterrebbe

$$\begin{cases} L(\mathbf{v}_1) = L(k\mathbf{v}_2) = kL(\mathbf{v}_2) = k\lambda_2 \mathbf{v}_2 \\ L(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1 = \lambda_1 k\mathbf{v}_2 = k\lambda_1 \mathbf{v}_2 \end{cases}$$

da cui, uguagliando gli ultimi membri di entrambe le catene di uguaglianze e semplificando  $k$ , si ricava l'uguaglianza  $(\lambda_1 - \lambda_2)\mathbf{v}_2 = \mathbf{0}$  che è assurda perché il coefficiente  $\lambda_1 - \lambda_2$  è diverso da zero in quanto  $\lambda_1 \neq \lambda_2$  e  $\mathbf{v}_2 \neq \mathbf{0}$ .

Quindi se  $\lambda_1 \neq \lambda_2$  non può valere la relazione  $\mathbf{v}_1 = k\mathbf{v}_2$  e allora concludiamo che due autovettori  $\mathbf{v}_1$  e  $\mathbf{v}_2$  corrispondenti a due autovalori distinti sono linearmente indipendenti. *C.d.d.*

## 2.4 Diagonalizzazione di applicazioni lineari

Dato uno spazio vettoriale  $V$  e un'applicazione lineare  $L : V \rightarrow V$ , se per ogni autovalore  $\lambda_i$  il corrispondente autospazio  $E(\lambda_i)$  ha dimensione  $h_i$  uguale alla molteplicità algebrica  $m_a(\lambda_i)$  di  $\lambda_i$ , allora esiste una base di  $V$  costituita da autovettori di  $L$ .

**Definizione:** un'applicazione lineare  $L : V \rightarrow V$  si dice *diagonalizzabile* se esiste una base di  $V$  costituita da autovettori di  $L$ .  $\triangle$

Se esiste una nuova base  $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$  costituita da autovettori dell'applicazione  $L$ , allora si può ridefinire l'applicazione  $L$  sugli autovettori di base, che, senza perdita di generalità, supponiamo associati ad autovalori di molteplicità algebrica 1

$$\begin{cases} L(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1 \\ L(\mathbf{v}_2) = \lambda_2 \mathbf{v}_2 \\ L(\mathbf{v}_3) = \lambda_3 \mathbf{v}_3 \\ \dots \\ L(\mathbf{v}_n) = \lambda_n \mathbf{v}_n \end{cases}$$

Definendo l'applicazione  $L$  sugli autovettori di base si ottiene, incolonnando i coefficienti delle immagini  $L(\mathbf{v}_i)$ , la matrice relativa alla base di autovettori data da

$$\mathcal{A}' = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

che è una matrice detta *diagonale* perché ha gli autovalori di  $L$  sulla diagonale principale e tutti zeri fuori dalla diagonale principale.

Tale matrice diagonale  $\mathcal{A}'$  si ottiene ovviamente anche con la relazione (10) in cui la matrice  $\mathcal{A}$  rappresenta l'applicazione  $L$  rispetto ad una base iniziale  $\mathcal{B}$  costituita da generici vettori iniziali e  $\mathcal{T}$  è la matrice del cambio di base dalle base iniziale  $\mathcal{B}$  alla base  $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$  costituita da autovettori dell'applicazione  $L$ .

Quando dunque gli autovettori di un'applicazione lineare  $L$  costituiscono una base dello spazio vettoriale  $V$ , diagonalizzare l'applicazione lineare  $L$ , rappresentata inizialmente da una certa matrice  $\mathcal{A}$ , significa passare in  $V$  dalla base iniziale

alla base degli autovettori, relativamente alla quale l'applicazione lineare  $L$  risulta quindi rappresentata, secondo la relazione (10), dalla nuova matrice  $\mathcal{A}'$  che ha forma diagonale. Sottolineiamo però, per concludere, che non occorre ricavare tale matrice diagonale mediante la (10) perché essa è in accordo con la definizione dell'applicazione  $L$  mediante le leggi di trasformazione degli autovettori di base.

Consideriamo come esempio l'applicazione lineare  $L : S_3 \rightarrow S_3$  definita, rispetto alla base  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , dalle seguenti leggi di trasformazione dei vettori di base

$$\begin{cases} L(\mathbf{e}_1) = -6\mathbf{e}_1 - 8\mathbf{e}_2 + 7\mathbf{e}_3 \\ L(\mathbf{e}_2) = 3\mathbf{e}_1 + 5\mathbf{e}_2 - 3\mathbf{e}_3 \\ L(\mathbf{e}_3) = -2\mathbf{e}_1 - 2\mathbf{e}_2 + 3\mathbf{e}_3 \end{cases}$$

La matrice  $\mathcal{A}$  associata a tale applicazione lineare si ottiene incolonnando i coefficienti delle immagini dei vettori di base

$$\mathcal{A} = \begin{pmatrix} -6 & 3 & -2 \\ -8 & 5 & -2 \\ 7 & -3 & 3 \end{pmatrix}$$

Gli autovalori e autovettori di  $L$  si ottengono dall'equazione secolare  $L(\mathbf{v}) = \lambda\mathbf{v}$  che si traduce nel sistema parametrico

$$\begin{pmatrix} -6 - \lambda & 3 & -2 \\ -8 & 5 - \lambda & -2 \\ 7 & -3 & 3 - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

la cui matrice dei coefficienti ha come determinante il polinomio caratteristico della matrice  $\mathcal{A}$

$$\mathcal{P}_3(\lambda) = -\lambda^3 + 2\lambda^2 + \lambda - 2$$

Gli autovalori di  $L$  sono gli zeri del polinomio caratteristico  $\mathcal{P}_3(\lambda)$  e valgono  $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ,  $\lambda_3 = -1$ . Risolvendo quindi i sistemi numerici  $\mathcal{A}(\mathbf{v}) = \lambda_i\mathbf{v}$ , si ottengono infine i tre autospazi

$$E(\lambda_1) = \{\mathbf{v} \in S_3, \text{ tale che } \mathbf{v} = \alpha(1, 1, -2)\}$$

$$E(\lambda_2) = \{\mathbf{v} \in S_3, \text{ tale che } \mathbf{v} = \alpha(1, 2, -1)\}$$

$$E(\lambda_3) = \{\mathbf{v} \in S_3, \text{ tale che } \mathbf{v} = \alpha(1, 1, -1)\}$$

tutti di dimensione pari a 1 perché la molteplicità algebrica di ciascun autovalore è pari a 1 e la dimensione dell'autospazio risulta allora in questo caso pari a  $1 \leq 1$ . Individuando ciascun autospazio mediante una base, ricaviamo gli autovettori

$$\mathbf{v}_1 = (1, 1, -2), \quad \mathbf{v}_2 = (1, 2, -1), \quad \mathbf{v}_3 = (1, 1, -1)$$

associati rispettivamente agli autovalori  $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ,  $\lambda_3 = -1$ .

Poiché gli autovettori di  $L$  costituiscono una base di  $S_3$ , allora è possibile ridefinire l'applicazione  $L$  sugli autovettori di base in modo da ottenere la matrice diagonale rispetto a tale base. Passando dunque dalla base naturale alla base degli autovettori, si ha la matrice del cambio di base  $\mathcal{T}$  data da

$$\mathcal{T} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ -2 & -1 & -1 \end{pmatrix}$$

tale che allora la matrice diagonale  $\mathcal{D}$  che rappresenta l'applicazione  $L$  nella nuova base degli autovettori si ottenga eseguendo

$$\mathcal{T}^{-1}\mathcal{A}\mathcal{T} = \mathcal{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Consideriamo ora il caso di applicazione lineare diagonalizzabile con autovalore di molteplicità algebrica 2 rappresentato nella base naturale dalla matrice

$$\mathcal{A} = \begin{pmatrix} 3 & -16 & 10 \\ -2 & 17 & -10 \\ -4 & 32 & -19 \end{pmatrix}$$

Il polinomio caratteristico  $\mathcal{P}_3(\lambda)$  di  $\mathcal{A}$  è  $\mathcal{P}_3(\lambda) = -\lambda^3 + \lambda^2 + \lambda - 1$  da cui scaturiscono gli autovalori  $\lambda_1 = -1$  di molteplicità algebrica 1 e  $\lambda_2 = \lambda_3 = 1$  di molteplicità algebrica 2.

Dall'autovalore  $\lambda_1 = -1$  discende l'autovettore  $\mathbf{v}_1 = (1, -1, -2)$ , ovvero l'autospazio di dimensione 1

$$E(\lambda_1) = \{\mathbf{v} \in S_3, \text{ tale che } \mathbf{v} = \alpha(1, -1, -2)\}$$

Nel caso dell'autovalore  $\lambda_2 = \lambda_3 = 1$  di molteplicità algebrica 2, abbiamo che la matrice  $\mathcal{A} - 1\mathcal{I}$  ha caratteristica 1, da cui segue che il sistema che fornisce l'autospazio corrispondente a  $\lambda_2 = \lambda_3 = 1$  ha, per il teorema di Rouché-Capelli, dimensione pari a 2 e coincide con il piano di equazione  $x - 8y + 5z = 0$ , ovvero

$$E(\lambda_2) = \{\mathbf{v}(x, y, z) \in S_3, \text{ tale che } \mathbf{v} = x - 8y + 5z = 0\}$$

In questo caso allora abbiamo che tutti gli autovettori relativi all'autovalore di molteplicità algebrica 2 appartengono non ad una retta, ma ad un piano sul quale possiamo scegliere due autovettori linearmente indipendenti qualsiasi, indicati con  $\mathbf{v}_2, \mathbf{v}_3$ , tali che i tre autovettori  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  costituiscano una base di  $S_3$  con la quale è possibile ridefinire l'applicazione  $L$  e ottenere come sua nuova rappresentazione la matrice diagonale.



Passando dunque dalla base naturale alla nuova base costituita da autovettori  $\mathcal{B} = \{\mathbf{v}_1 = (1, -1, -2), \mathbf{v}_2 = (8, 1, 0), \mathbf{v}_3 = (-5, 0, 1)\}$ , si ha la matrice del cambio di base  $\mathcal{T}$  data da

$$\mathcal{T} = \begin{pmatrix} 1 & 8 & -5 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

tale che allora la matrice diagonale  $\mathcal{D}$ , che rappresenta l'applicazione  $L$  nella nuova base degli autovettori, si ottenga, sempre in base alla relazione (10), eseguendo

$$\mathcal{T}^{-1}\mathcal{A}\mathcal{T} = \mathcal{D} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

## 2.5 Proprietà delle applicazioni simmetriche

**Definizione:** dato uno spazio euclideo  $V$ , un'applicazione lineare  $L : V \rightarrow V$  si dice *applicazione lineare simmetrica* se per ogni coppia di vettori  $\mathbf{v}_1, \mathbf{v}_2 \in V$  sussiste l'uguaglianza

$$\langle L(\mathbf{v}_1), \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, L(\mathbf{v}_2) \rangle$$

△

Omettendone la dimostrazione, enunciamo il teorema in base al quale, se nello spazio  $V$  è fissata una base ortonormale, allora un'applicazione  $L : V \rightarrow V$  è simmetrica se e solo se la matrice ad essa associata è simmetrica.

Per dimostrare alcune importanti proprietà delle applicazioni simmetriche, occorre estendere il concetto di prodotto interno al caso di spazio vettoriale sul campo dei numeri complessi, i cui elementi, indicati con  $z$ , sono della forma  $z = a + ib$ , dove  $a, b \in \mathcal{R}$  e il simbolo  $i$ , denominato *unità immaginaria*, è tale che  $i^2 = -1$ .

Per uno spazio vettoriale sul campo complesso  $V_{\mathcal{C}}$  (quando non sorgono equivoci sarà omesso l'indice  $\mathcal{C}$ ), le proprietà del prodotto interno debbono essere modificate perché in particolare le proprietà  $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle$ , per ogni coppia di vettori  $\mathbf{v}_1, \mathbf{v}_2 \in V$ , e  $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ , per ogni vettore  $\mathbf{v} \in V$  diverso dal vettore nullo, conducono in taluni casi ad una contraddizione, come si vede dal seguente esempio.

Dato il vettore  $\mathbf{v} = (1, i) \in S_2$ , diverso dal vettore nullo, si ha che la sua norma vale  $\langle \mathbf{v}, \mathbf{v} \rangle = 1^2 + i^2 = 0$ , sebbene il vettore  $\mathbf{v} = (1, i)$  non sia il vettore nullo.

Per rimuovere tali contraddizioni, la prima proprietà del prodotto interno nel caso di spazio euclideo complesso deve assumere la forma

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \overline{\langle \mathbf{v}_2, \mathbf{v}_1 \rangle}$$

per ogni  $\mathbf{v}_1, \mathbf{v}_2 \in V_{\mathcal{C}}$ , dove il simbolo di “soprasegnatura” rappresenta il numero *complesso coniugato* definito come da

$$\overline{a + ib} = a - ib$$

Dalla prima proprietà così modificata e dalla seconda si deduce che nel caso di prodotto interno su spazio euclideo complesso vale la relazione

$$\langle \mathbf{v}_1, \alpha \mathbf{v}_2 \rangle = \overline{\langle \alpha \mathbf{v}_2, \mathbf{v}_1 \rangle} = \overline{\alpha \langle \mathbf{v}_2, \mathbf{v}_1 \rangle} = \bar{\alpha} \overline{\langle \mathbf{v}_2, \mathbf{v}_1 \rangle} = \bar{\alpha} \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \quad (13)$$

dalla quale ricaviamo l'informazione che se dal prodotto interno *esce*, in virtù della seconda proprietà, il coefficiente del primo vettore, allora il coefficiente del secondo vettore *esce* dal prodotto interno diventando il complesso coniugato.

Questa modifica delle proprietà del prodotto interno conduce alla modifica della *regola* del prodotto scalare nello spazio vettoriale  $S_n$  quando le  $n$ -ple contengono anche numeri complessi e si ha in particolare per ogni coppia di vettori di  $S_n$ , indicati con  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  e  $\mathbf{v} = (v_1, v_2, \dots, v_n)$ , la nuova regola

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n u_i \bar{v}_i$$

che verifica le proprietà del prodotto interno in campo complesso.

La proprietà del passaggio al complesso coniugato in caso di scambio dei vettori è ora compatibile con la terza proprietà per cui il prodotto interno di un vettore  $\mathbf{v} \neq \mathbf{0}$  con se stesso è sempre un numero reale positivo ed è pari a zero solo se il vettore è  $\mathbf{v} = \mathbf{0}$ .

In virtù della regola del prodotto scalare complesso si ha infatti in particolare

$$\langle (1, i), (1, i) \rangle = 1 \cdot 1 + i \cdot \bar{i} = 1 \cdot 1 + i \cdot (-i) = 1 - i^2 = 1 + 1 = 2$$

A questo punto possiamo dimostrare le proprietà fondamentali sugli autovalori e autovettori delle applicazioni lineari simmetriche.

Anche se si è esteso il prodotto interno al caso di spazio euclideo complesso, continuiamo a considerare soltanto applicazioni lineari sul campo reale, a cui cioè sono associate matrici ad elementi reali, e l'estensione è dovuta soltanto all'esigenza di considerare il caso che una matrice, pur ad elementi reali, possa avere autovalori complessi.

**Teorema:** gli autovalori di un'applicazione simmetrica e reale  $L : V \rightarrow V$ , con  $V$  spazio euclideo, sono tutti reali.

**Dimostrazione:** data l'equazione agli autovalori  $L(\mathbf{v}) = \lambda \mathbf{v}$ , con  $L$  applicazione lineare simmetrica, dobbiamo dimostrare che  $\lambda$  è sempre un autovalore reale, ovvero coincidente con il suo complesso coniugato  $\bar{\lambda}$ . Sviluppando i due prodotti interni e utilizzando la (13) nella seconda catena di uguaglianze, si ottiene

$$\begin{cases} \langle L(\mathbf{v}), \mathbf{v} \rangle = \langle \lambda \mathbf{v}, \mathbf{v} \rangle = \lambda \langle \mathbf{v}, \mathbf{v} \rangle \\ \langle \mathbf{v}, L(\mathbf{v}) \rangle = \langle \mathbf{v}, \lambda \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle \end{cases}$$

da cui per confronto segue che, essendo uguali per la simmetria di  $L$  il primo membro nella prima e nella seconda catena di uguaglianze, allora saranno uguali anche

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle$$

cioè vale appunto  $\lambda = \bar{\lambda}$  e  $\lambda$  è allora un numero reale.

*C.d.d.*

**Teorema:** due autovettori di un'applicazione simmetrica reale  $L : V \rightarrow V$ , con  $V$  spazio euclideo, relativi a due autovalori distinti, sono ortogonali.

**Dimostrazione:** siano  $\mathbf{v}_1, \mathbf{v}_2 \in V$  due autovettori dell'applicazione simmetrica  $L$  relativi agli autovalori distinti rispettivamente  $\lambda$  e  $\mu$ , ovvero

$$L(\mathbf{v}_1) = \lambda \mathbf{v}_1 \quad \text{e} \quad L(\mathbf{v}_2) = \mu \mathbf{v}_2$$

con  $\lambda \neq \mu$  e  $\lambda, \mu$  numeri reali per il teorema appena dimostrato.

Sviluppando i due prodotti interni senza più dover considerare gli autovalori complessi coniugati perché essi sono reali, otteniamo

$$\begin{cases} \langle L(\mathbf{v}_1), \mathbf{v}_2 \rangle = \langle \lambda \mathbf{v}_1, \mathbf{v}_2 \rangle = \lambda \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \\ \langle \mathbf{v}_1, L(\mathbf{v}_2) \rangle = \langle \mathbf{v}_1, \mu \mathbf{v}_2 \rangle = \mu \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \end{cases}$$

da cui per confronto segue che, essendo uguali per la simmetria di  $L$  il primo membro nella prima e nella seconda catena di uguaglianze, allora saranno uguali anche

$$\lambda \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mu \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

ovvero  $(\lambda - \mu) \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$  e dunque segue l'ortogonalità di  $\mathbf{v}_1, \mathbf{v}_2$  perché per la relazione  $\lambda \neq \mu$  vale l'uguaglianza  $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ .

Il teorema appena dimostrato rende conto della relazione che intercorre fra due autovettori relativi ad autovalori distinti, ma non dice nulla sulla diagonalizzabilità o meno dell'applicazione lineare stessa.

**Teorema:** data un'applicazione simmetrica reale  $L : V \rightarrow V$ , con  $V$  spazio euclideo, e un suo autovettore  $\mathbf{v}$ , se il vettore  $\mathbf{w} \in V$  è ortogonale a  $\mathbf{v}$ , allora si ha che anche  $L(\mathbf{w})$  è ortogonale a  $\mathbf{v}$ .

**Dimostrazione:** se  $\mathbf{v}$  è autovettore di  $L$  relativo ad un certo autovalore  $\lambda$ , allora si ha  $L(\mathbf{v}) = \lambda \mathbf{v}$ . Se ora il vettore  $\mathbf{w} \in V$  è ortogonale a  $\mathbf{v}$ , allora segue

$$\langle L(\mathbf{w}), \mathbf{v} \rangle = \langle \mathbf{w}, L(\mathbf{v}) \rangle = \langle \mathbf{w}, \lambda \mathbf{v} \rangle = \lambda \langle \mathbf{w}, \mathbf{v} \rangle = \lambda 0 = 0$$

ovvero il vettore  $L(\mathbf{w})$  è appunto ortogonale al vettore  $\mathbf{v}$ .

*C.d.d.*

A questo punto possiamo dimostrare il teorema fondamentale sulle applicazioni simmetriche, denominato *teorema spettrale*. L'importanza fondamentale di tale teorema è dovuta al fatto che esso non solo ci assicura la diagonalizzabilità di un'applicazione lineare simmetrica reale  $L$ , ma ci assicura anche che la base rispetto alla quale l'applicazione  $L$  si rappresenta con una matrice diagonale, è una base ortonormale.

**Teorema spettrale:** data un'applicazione simmetrica reale  $L : V \rightarrow V$ , con  $V$  spazio euclideo, esiste sempre una base ortonormale di  $V$  costituita da autovettori dell'applicazione  $L$ .

**Dimostrazione:** nei testi di algebra lineare tale teorema si dimostra per induzione utilizzando il teorema precedente. Qui questo teorema sarà dimostrato

soltanto in due casi particolari dai quali comunque si capisce come esso può essere esteso al caso generale.

Se gli autovalori di  $L$  avessero tutti molteplicità algebrica 1, allora la tesi del teorema scaturirebbe immediatamente perché ad ogni autovalore corrisponderebbe un autospazio di dimensione 1 e quindi avremmo tanti autovettori quanto è la dimensione di  $V$  e tali autovettori sarebbero automaticamente ortogonali in quanto riferiti ad autovalori distinti.

Resta da dimostrare dunque la diagonalizzabilità dell'applicazione lineare  $L$  rispetto ad una base ortonormale nel caso in cui alcuni autovalori abbiano molteplicità algebrica maggiore di 1.

A tale scopo consideriamo il caso di applicazione  $L$  sullo spazio  $S_3$  avente autovalori  $\lambda, \lambda, \mu$ , ovvero un autovalore  $\mu$  di molteplicità algebrica 1 e un autovalore  $\lambda$  di molteplicità algebrica 2. All'autovalore  $\mu$  corrisponde un autospazio di dimensione 1, cioè un autovettore che indichiamo con  $\mathbf{u}$ ; dobbiamo dimostrare allora che all'autovalore  $\lambda$  corrisponde autospazio di dimensione 2, perché in tal caso allora si possono scegliere in tale autospazio due autovettori ortogonali fra loro e automaticamente ortogonali all'autovettore  $\mathbf{u}$ .

Per dimostrare che  $E(\lambda)$  ha dimensione 2, consideriamo un autovettore, indicato con  $\mathbf{v}$ , relativo all'autovalore  $\lambda$ , ricordando che per definizione ciascun autovalore di molteplicità algebrica maggiore di 1 fornisce sempre almeno un autovettore: tale autovettore  $\mathbf{v}$  risulta ortogonale all'autovettore  $\mathbf{u}$  perché  $\lambda \neq \mu$ .

A questo punto, poiché siamo in  $S_3$ , esiste il complemento ortogonale del sottospazio di dimensione 2 individuato dai vettori  $\mathbf{u}, \mathbf{v}$  e tale complemento ortogonale, avendo dimensione  $1 = 3 - 2$ , è individuato da un vettore che chiamiamo  $\mathbf{w}$ .

Poiché dunque si ha  $\langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle = 0$ , allora per il teorema precedente si ha anche

$$\langle L(\mathbf{w}), \mathbf{u} \rangle = \langle L(\mathbf{w}), \mathbf{v} \rangle = 0$$

da cui si deduce che  $L(\mathbf{w})$  è un vettore ortogonale agli autovettori  $\mathbf{u}, \mathbf{v}$  e pertanto risulta appartenente alla retta di  $\mathbf{w}$  perché il complemento ortogonale del sottospazio individuato dagli autovettori  $\mathbf{u}, \mathbf{v}$  ha, come detto, dimensione 1.

Essendo dunque  $L(\mathbf{w})$  linearmente dipendente da  $\mathbf{w}$ , allora si ha la relazione di proporzionalità  $L(\mathbf{w}) = c\mathbf{w}$ , dove però si identificano  $c = \lambda$  perché se  $c$  fosse diverso da  $\lambda$ , allora si avrebbe la conclusione assurda che  $c$  sarebbe il quarto autovalore di  $L$  dopo i tre autovalori già trovati  $\lambda, \lambda, \mu$ .

Possiamo concludere quindi che all'autovalore  $\lambda$  di molteplicità algebrica 2 corrisponde un autospazio di dimensione 2 individuato dagli autovettori  $\mathbf{v}, \mathbf{w}$  ortogonali e che dunque esiste una base di  $S_3$  costituita dai tre autovettori  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  di  $L$  a due a due ortogonali.

Per completare la dimostrazione consideriamo il caso in cui l'applicazione lineare sia  $L : S_4 \rightarrow S_4$  con due autovalori di molteplicità algebrica 2 che, per semplicità e senza perdita di generalità, supponiamo essere  $\lambda_1 = \lambda_2 = 5$  e  $\lambda_3 = \lambda_4 = 7$ , ottenuti quindi come zeri del polinomio caratteristico  $\mathcal{P}_4(\lambda) = (\lambda - 5)^2 (\lambda - 7)^2$ .

Dobbiamo dimostrare che a ciascun autovalore corrisponde un autospazio di

dimensione 2. A tale scopo consideriamo un autovettore  $\mathbf{u}$  relativo a  $\lambda = 5$  e un autovettore  $\mathbf{v}$  relativo a  $\lambda = 7$ , ricordando che per definizione ciascun autovalore di molteplicità algebrica maggiore di 1 fornisce sempre almeno un autovettore.

Poichè lo spazio vettoriale è  $S_4$ , allora potremo sempre trovare due suoi vettori, indicati con  $\mathbf{w}, \mathbf{z}$ , che siano ciascuno ortogonale ad entrambi gli autovettori  $\mathbf{u}, \mathbf{v}$  già determinati. Per il teorema precedente avremo che anche i vettori immagine  $L(\mathbf{w})$  e  $L(\mathbf{z})$  sono ortogonali ad entrambi gli autovettori  $\mathbf{u}, \mathbf{v}$  e quindi apparterranno al medesimo sottospazio individuato dai vettori  $\mathbf{w}, \mathbf{z}$ , complemento ortogonale del sottospazio individuato dagli autovettori  $\mathbf{u}, \mathbf{v}$ .

Ridefinendo a questo punto l'applicazione  $L$  sulla base ortonormale (supponendo di aver normalizzato i vettori)  $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ , avremo le leggi di trasformazione dei vettori di base

$$\begin{cases} L(\mathbf{u}) = 5\mathbf{u} \\ L(\mathbf{v}) = 7\mathbf{v} \\ L(\mathbf{w}) = a\mathbf{w} + b\mathbf{z} \\ L(\mathbf{z}) = b\mathbf{w} + c\mathbf{z} \end{cases}$$

dalle quali, per l'ortonormalità della base, incolonnando i coefficienti, si otterrà la matrice  $\mathcal{A}$  simmetrica associata ad  $L$  (rispetto a tale base) data da

$$\mathcal{A} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & b & c \end{pmatrix}$$

Sviluppando il determinante  $\det(\mathcal{A} - \lambda\mathcal{I})$ , si ottiene il polinomio caratteristico della matrice  $\mathcal{A}$  avente la forma  $\mathcal{P}_4(\lambda) = (5 - \lambda)(7 - \lambda)[(a - \lambda)(c - \lambda) - b^2]$ .

Poiché il polinomio caratteristico di un'applicazione lineare è, come dimostrato in precedenza, invariante per cambio di base, allora il fattore del polinomio caratteristico in parentesi quadra avrà come zeri i numeri 5 e 7, in modo che l'insieme degli autovalori contenga sempre il 5 e il 7 con le rispettive molteplicità.

Nella matrice  $\mathcal{A}$  abbiamo che il minore  $\mathcal{M}$  di ordine 2 dato da

$$\mathcal{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

può essere considerato come la *restrizione* dell'applicazione  $L$  al solo sottospazio individuato dai vettori  $\mathbf{w}, \mathbf{z}$  e possiede, come detto, gli autovalori 5 e 7.

Gli autovettori di tale minore  $\mathcal{M}$  forniranno pertanto le due combinazioni lineari dei vettori  $\mathbf{w}, \mathbf{z}$  che sono autovettori ortogonali di  $L$ , indicati con  $\mathbf{w}', \mathbf{z}'$ , riferiti agli autovalori 5 e 7 e ortogonali agli autovettori  $\mathbf{u}, \mathbf{v}$ , in modo dunque che l'applicazione  $L$  abbia i due autovettori  $\mathbf{u}, \mathbf{w}'$  relativi all'autovalore 5 e i due autovettori  $\mathbf{v}, \mathbf{z}'$  relativi all'autovalore 7.

L'insieme dei quattro autovettori  $\mathbf{u}, \mathbf{v}, \mathbf{w}', \mathbf{z}'$ , a due a due ortogonali, di  $L$  forma in definitiva una base ortonormale di  $S_4$ .

Dato ora una generica applicazione lineare definita su uno spazio vettoriale di qualunque dimensione e avente autovalori di qualsiasi molteplicità algebrica maggiore di 1, si calcolano i vettori ortogonali a tutti gli autovettori trovati uno per ciascun autovalore.

Ridefinendo quindi l'applicazione lineare su tale base ortonormale formata da autovettori e da vettori appartenenti al loro complemento ortogonale, si calcolano a questo punto gli autovettori dei minori relativi a quegli autovalori di cui deve essere *completato* l'autospazio, determinando in tal modo tanti autovettori ortonormali in ciascun autospazio quant'è la molteplicità algebrica dell'autovalore corrispondente.

*C.d.d.*

## 2.6 Diagonalizzazione di forme quadratiche

**Definizione:** dato uno spazio euclideo  $V$  di dimensione  $n$  e un'applicazione  $L : V \rightarrow V$  simmetrica, l'applicazione  $Q : V \rightarrow V$  definita per ogni vettore  $\mathbf{v} \in V$  come

$$Q(\mathbf{v}) = \langle \mathbf{v}, L(\mathbf{v}) \rangle$$

prende il nome di *forma quadratica*.  $\triangle$

Se fissiamo una base ortonormale  $\mathcal{B}$  rispetto a cui la matrice  $\mathcal{A}$  associata all'applicazione  $L$  risulti simmetrica, allora la forma quadratica può essere espressa attraverso il prodotto righe per colonne

$$Q(\mathbf{v}_{\mathcal{B}}) = (\mathbf{v}_{\mathcal{B}})^T \mathcal{A} \mathbf{v}_{\mathcal{B}}$$

dove  $\mathbf{v}^T$  rappresenta il vettore  $\mathbf{v}$  disposto in riga a sinistra della matrice e  $\mathbf{v}$  rappresenta il vettore colonna disposto a destra della matrice  $\mathcal{A}$ .

Se ora in  $V$  passiamo dalla base iniziale  $\mathcal{B}$  alla base nuova  $\mathcal{B}'$  con una matrice di cambio di base  $\mathcal{T}$ , allora avremo che le componenti dei vettori di  $V$  cambieranno secondo l'usuale regola

$$\mathbf{v}_{\mathcal{B}'} = \mathcal{T}^{-1} \mathbf{v}_{\mathcal{B}} \quad \text{e l'inversa} \quad \mathbf{v}_{\mathcal{B}} = \mathcal{T} \mathbf{v}_{\mathcal{B}'}$$

in conseguenza delle quali la forma quadratica diventa

$$\tilde{Q}(\mathbf{v}_{\mathcal{B}'}) = (\mathbf{v}_{\mathcal{B}'})^T \mathcal{A} \mathbf{v}_{\mathcal{B}'} = (\mathcal{T} \mathbf{v}_{\mathcal{B}'})^T \mathcal{A} \mathcal{T} \mathbf{v}_{\mathcal{B}'} = (\mathbf{v}_{\mathcal{B}'})^T (\mathcal{T}^T \mathcal{A} \mathcal{T}) \mathbf{v}_{\mathcal{B}'}$$

che ha sempre la stessa struttura di prodotto righe per colonne di un vettore trasposto per una matrice al centro per un vettore a destra.

Affinchè la forma quadratica relativa alla nuova base contenga soltanto i quadrati delle variabili e non anche i cosiddetti *termini misti*, allora la matrice  $\mathcal{T}^T \mathcal{A} \mathcal{T}$  deve essere una matrice diagonale. Quello che sappiamo fino ad ora è che se eseguiamo il prodotto  $\mathcal{T}^{-1} \mathcal{A} \mathcal{T}$ , come prescritto dalla relazione (10), con  $\mathcal{T}$  matrice del cambio di base verso la base di autovettori di  $\mathcal{A}$ , allora la matrice  $\mathcal{A}$ , sempre diagonalizzabile perché simmetrica, si trasformerebbe nella matrice diagonale  $\mathcal{D}$ .

Ma nel nostro caso abbiamo il prodotto  $\mathcal{T}^T \mathcal{A} \mathcal{T}$  e dunque quanto detto a proposito della diagonalizzazione di applicazioni lineari non vale più del tutto.

Se però, data l'ortogonalità degli autovettori della matrice simmetrica  $\mathcal{A}$ , assicurata dal teorema spettrale, consideriamo il cambio di base verso la base  $\mathcal{B}'$  contenente non generici autovettori dell'autospazio di appartenenza, come andrebbe bene invece per diagonalizzare un'applicazione lineare attraverso la (10), ma contenente tutti gli autovettori con norma unitaria, allora la matrice  $\mathcal{T}^T$  coincide, come sarebbe facile verificare, con la matrice  $\mathcal{T}^{-1}$ , e dunque la matrice della forma quadratica associata a tale base costituita da autovettori di norma unitaria assume la forma diagonale  $\mathcal{D}$  data dal prodotto  $\mathcal{T}^T \mathcal{A} \mathcal{T} = \mathcal{T}^{-1} \mathcal{A} \mathcal{T} = \mathcal{D}$ .

### 3 Riduzione a forma canonica di Jordan

Data un'applicazione lineare  $L : V \longrightarrow V$ , con  $V$  spazio vettoriale di dimensione  $n$ , se almeno un autovalore  $\lambda_i$  di molteplicità algebrica  $m_a(\lambda_i)$  maggiore di 1 fornisce l'autospazio corrispondente con dimensione minore di  $m_a(\lambda_i)$ , allora l'applicazione lineare  $L$  non è diagonalizzabile, ovvero non esiste una base di  $V$  costituita da autovettori di  $L$ .

In questo caso allora si aggiungono agli autovettori linearmente indipendenti dell'applicazione  $L$  dei vettori in modo da completare la base (costituita però non da soli autovettori) di  $V$  e ottenere che rispetto ad essa la matrice  $\mathcal{A}'$ , associata all'applicazione  $L$ , assuma quella che prende il nome di *forma canonica di Jordan*.

Per completare la base di  $V$ , si dovranno pertanto aggiungere agli autovettori della matrice una quantità di altri vettori pari alla differenza tra la molteplicità algebrica di ciascun autovalore e la dimensione dell'autospazio che l'autovalore genera.

In particolare si procederà con il completamento separato di ciascun autospazio la cui dimensione è minore della molteplicità algebrica dell'autovalore corrispondente

Per illustrare la procedura della determinazione dei vettori da aggiungere agli autovettori di  $L$  al fine di completare la base di  $V$ , considereremo soltanto degli esempi che saranno in ogni caso estensibili a tutti i casi generali.

#### 3.1 Forma canonica di Jordan per matrici di ordine 2

Consideriamo in  $S_2$  un'applicazione lineare la cui matrice associata  $\mathcal{A}$  rispetto ad una base iniziale  $\mathcal{B}$  è data da

$$\mathcal{A} = \begin{pmatrix} 5 & 2 \\ -2 & 1 \end{pmatrix}$$

avente  $k = 3$  come autovalore di molteplicità algebrica 2 e l'autospazio di dimensione 1 dato dall'autovettore

$$\mathbf{u} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Poiché il solo autovettore  $\mathbf{u}$  non costituisce una base per lo spazio vettoriale  $S_2$ , allora la matrice  $\mathcal{A}$  è non diagonalizzabile.

Quando non si può ridurre dunque tale matrice alla forma diagonale, la si potrà comunque ridurre sempre alla forma canonica di Jordan.

Nel caso di matrice di ordine 2 non diagonalizzabile dovremo aggiungere all'autovettore  $\mathbf{u}$  un solo altro vettore perché si ha che la differenza fra la molteplicità algebrica 2 dell'autovalore  $k = 3$  e la dimensione dell'autospazio è appunto  $2 - 1 = 1$ .

Il vettore  $\mathbf{v}$  da aggiungere all'autovettore  $\mathbf{u}$  si otterrà risolvendo un sistema lineare analogo a quello da cui abbiamo ricavato l'autovettore, in cui però inseriamo come termini noti le componenti dell'autovettore. Abbiamo

$$\begin{pmatrix} 5-3 & 2 \\ -2 & 1-3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

dove, come appunto per calcolare l'autovettore, è stato sottratto l'autovalore 3 dagli elementi della diagonale principale.

Il sistema diventa

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

che ha  $\infty^1$  soluzioni fra le quali se ne può scegliere una qualsiasi per formare il vettore  $\underline{\mathbf{v}}$ . Una di queste soluzioni è

$$x = \frac{1}{2}, \quad y = 0$$

da cui si ottiene il vettore

$$\mathbf{v} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$$

A questo punto possiamo scrivere la matrice  $\mathcal{T}$  del cambiamento di base

$$\mathcal{T} = \begin{pmatrix} 1 & 1/2 \\ -1 & 0 \end{pmatrix}$$

da cui segue che la matrice inversa è

$$\mathcal{T}^{-1} = \begin{pmatrix} 0 & -1 \\ 2 & 2 \end{pmatrix}$$

A questo punto otteniamo, attraverso la relazione (10), la forma canonica  $\mathcal{J}$  di Jordan della matrice  $\mathcal{A}$  data da

$$\mathcal{J} = \mathcal{T}^{-1} \mathcal{A} \mathcal{T} = \begin{pmatrix} 0 & -1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$



### 3.2 Forma canonica di Jordan per matrici di ordine 3

Le matrici di ordine 3 non diagonalizzabili sono soltanto quelle che hanno

- un autovalore di molteplicità algebrica 2 con autospazio di dimensione 1;
- un autovalore di molteplicità algebrica 3 con autospazio di dimensione 1;
- un autovalore di molteplicità algebrica 3 con autospazio di dimensione 2.

e la procedura che illustreremo per ridurre tali matrici a forma canonica di Jordan sarà facilmente estensibile a matrici non diagonalizzabili di tutti gli ordini.

#### 3.2.1 Matrice con autovalore di molteplicità algebrica 2

Consideriamo in  $S_3$  un'applicazione lineare la cui matrice associata  $\mathcal{A}$  rispetto ad una base iniziale  $\mathcal{B}$  è data da

$$\mathcal{A} = \begin{pmatrix} -5 & -1 & 2 \\ 2 & 1 & 2 \\ -7 & -1 & 4 \end{pmatrix}$$

Tale matrice ha autovalore  $\lambda = 2$  con molteplicità algebrica 1 a cui corrisponde l'autovettore  $\mathbf{u} = (0, 2, 1)$  e autovalore  $\lambda = -1$  con molteplicità algebrica 2 che fornisce autospazio con dimensione 1 individuato dall'autovettore  $\mathbf{v} = (1, -2, 1)$ .

Poiché i soli autovettori  $\mathbf{u}$  e  $\mathbf{v}$  non formano una base di  $S_3$ , allora dobbiamo aggiungerne un altro per completare la base: quest'ultimo vettore dovrà "colmare" la differenza fra la molteplicità algebrica 2 dell'autovalore  $\lambda = -1$  e la dimensione 1 dell'autospazio corrispondente.

Poiché l'autospazio con dimensione inferiore alla molteplicità algebrica del relativo autovalore è l'autospazio dell'autovalore  $\lambda = -1$ , allora determineremo il vettore da aggiungere all'unico autovettore relativo a  $\lambda = -1$  risolvendo il sistema

$$[\mathcal{A} - (-1)\mathcal{I}] \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

ovvero

$$\begin{pmatrix} -5 - (-1) & -1 & 2 \\ 2 & 1 - (-1) & 2 \\ -7 & -1 & 4 - (-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

in cui il primo membro coincide con quello del sistema da cui abbiamo ottenuto l'autovettore  $\mathbf{v}$  e il secondo membro è dato dall'autovettore stesso  $\mathbf{v}$  già trovato.

Questo sistema ha la due matrici associate di caratteristica 2 e per il teorema di Rouché-Capelli esso ha allora  $\infty^1$  soluzioni che si possono trovare per esempio dal minore di ordine 2 corrispondente alle prime due incognite nelle prime due equazioni

$$x = \alpha, \quad y = -1 - 2\alpha, \quad z = \alpha$$

da cui, per  $\alpha = 0$  che è la scelta più “comoda”, ricaviamo appunto il terzo vettore della nuova base  $\mathbf{w} = (0, -1, 0)$ .

A questo punto costruiamo la matrice  $\mathcal{T}$  del cambio di base dalla base  $\mathcal{B}$  alla base  $\mathcal{B}' = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$  costituita dai due autovettori  $\mathbf{u}, \mathbf{v}$  e dal vettore  $\mathbf{w}$

$$\mathcal{T} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & -2 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

in modo che la matrice  $\mathcal{J}$  associata all'applicazione  $L$  rispetto alla base  $\mathcal{B}'$  si ottenga eseguendo il prodotto dato dalla (10)

$$\begin{aligned} \mathcal{J} = \mathcal{T}^{-1}A\mathcal{T} &= \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ -4 & -1 & 2 \end{pmatrix} \begin{pmatrix} -5 & -1 & 2 \\ 2 & 1 & 2 \\ -7 & -1 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 2 & -2 & -1 \\ 1 & 1 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

### 3.2.2 Autovalore di molteplicità 3 con un solo autovettore

Consideriamo in  $S_3$  un'applicazione lineare la cui matrice associata  $\mathcal{A}$  rispetto ad una base iniziale  $\mathcal{B}$  è data da

$$\mathcal{A} = \begin{pmatrix} 20 & -31 & -11 \\ 11 & -17 & -7 \\ -3 & 5 & 3 \end{pmatrix}$$

Tale matrice ha autovalore  $\lambda = 2$  con molteplicità algebrica 3 a cui corrisponde un autospazio di dimensione 1 individuato dall'autovettore di base  $\mathbf{u} = (-8, -5, 1)$

Poiché il solo autovettore  $\mathbf{u}$  non forma una base di  $S_3$ , allora per avere una base dobbiamo aggiungere ad esso altri due vettori indicati con  $\mathbf{v}, \mathbf{w}$  che dovranno “colmare” la differenza fra la molteplicità algebrica dell'autovalore  $\lambda = 2$  e la dimensione dell'autospazio corrispondente.

Poiché all'autovalore  $\lambda = 2$  di molteplicità algebrica 3 corrisponde un autospazio di molteplicità geometrica 1, l'unico autovettore che si trova viene *utilizzato* per ricavare il secondo vettore e il secondo vettore viene *utilizzato* per ricavare il terzo vettore (catena di Jordan).

Determineremo il vettore  $\mathbf{v}$  scrivendo l'autovettore  $\mathbf{u}$  come termine noto nel sistema

$$\begin{pmatrix} 20-2 & -31 & -11 \\ 11 & -17-2 & -7 \\ -3 & 5 & 3-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -8 \\ -5 \\ 1 \end{pmatrix}$$

che riscriviamo nella forma

$$\begin{cases} 18x - 31y - 11z = -8 \\ 11x - 19y - 7z = -5 \\ -3x + 5y + z = 1 \end{cases}$$

Questo sistema ha entrambe le matrici associate con caratteristica 2 e quindi per il teorema di Rouché-Capelli possiede  $\infty^1$  soluzioni: risolvendo rispetto al minore di ordine 2 relativo alla prima e terza incognita nella prima e terza equazione, si ottiene la soluzione particolare

$$x = -\frac{1}{5}, \quad y = 0, \quad z = \frac{2}{5}$$

che scriviamo sotto forma di vettore

$$\mathbf{v} = \left( -\frac{1}{5}, 0, \frac{2}{5} \right)$$

Per ottenere  $\mathbf{w}$ , dobbiamo risolvere il sistema di equazioni algebriche

$$\begin{cases} 18x - 31y - 11z = -1/5 \\ 11x - 19y - 7z = 0 \\ -3x + 5y + z = 2/5 \end{cases}$$

una cui soluzione particolare (tra le  $\infty^1$  che il sistema possiede!) è

$$x = -\frac{7}{25}, \quad y = 0, \quad z = -\frac{11}{25}$$

che scriviamo sotto forma di vettore

$$\mathbf{w} = \left( -\frac{7}{25}, 0, -\frac{11}{25} \right)$$

A questo punto costruiamo la matrice  $\mathcal{T}$  del cambio di base dalla base naturale alla base formata dall'autovettore  $\mathbf{u}$  e dagli altri due vettori  $\mathbf{v}, \mathbf{w}$

$$\mathcal{T} = \begin{pmatrix} -8 & -1/5 & -7/25 \\ -5 & 0 & 0 \\ 1 & 2/5 & -11/25 \end{pmatrix}$$

La matrice  $\mathcal{J}$  di Jordan associata all'applicazione  $L$  relativamente alla nuova base  $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$  si determina infine mediante la (10)

$$\begin{aligned} \mathcal{J} &= \mathcal{T}^{-1} A \mathcal{T} = \\ &= \begin{pmatrix} 0 & -1/5 & 0 \\ -11/5 & 19/5 & 7/5 \\ -2 & 3 & -1 \end{pmatrix} \begin{pmatrix} 20 & -31 & -11 \\ 11 & -17 & -7 \\ -3 & 5 & 3 \end{pmatrix} \begin{pmatrix} -8 & -1/5 & -7/25 \\ -5 & 0 & 0 \\ 1 & 2/5 & -11/25 \end{pmatrix} = \\ &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

### 3.2.3 Autovalore di molteplicità 3 con due autovettori

Consideriamo in  $S_3$  un'applicazione lineare la cui matrice associata  $\mathcal{A}$  rispetto ad una base iniziale  $\mathcal{B}$  è data da

$$\mathcal{A} = \begin{pmatrix} 27 & -40 & -8 \\ 15 & -22 & -5 \\ -3 & 5 & 4 \end{pmatrix}$$

Tale matrice ha autovalore  $\lambda = 3$  con molteplicità algebrica 3 a cui corrisponde un autospazio che si ottiene risolvendo il sistema

$$\begin{pmatrix} 27-3 & -40 & -8 \\ 15 & -22-3 & -5 \\ -3 & 5 & 4-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Tale sistema ha le caratteristiche delle due matrici associate pari a 1 e dunque, per il teorema di Rouché-Capelli, esso avrà  $\infty^2$  soluzioni ottenibili dal minore di ordine 1 corrispondente alla  $z$  nella terza equazione

$$z = 3x - 5y$$

da cui segue

$$\begin{cases} x = \alpha \\ y = \beta \\ z = 3\alpha - 5\beta \end{cases}$$

Riscriviamo la soluzione nella forma vettoriale

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix}$$

da cui ricaviamo gli autovettori

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix}$$

Poiché i due autovettori  $\mathbf{u}, \mathbf{v}$  non formano una base di  $S_3$ , allora per avere una base dobbiamo aggiungere ad essi un solo altro vettore indicato con  $\mathbf{w}$  che “colmi” la differenza fra la molteplicità algebrica dell’autovalore  $\lambda = 3$  e la dimensione 2 dell’autospazio corrispondente.

Gli autovettori relativi all’autovalore  $\lambda = 3$  formano un sottospazio di dimensione 2 di cui  $\mathbf{u}$  e  $\mathbf{v}$  sono una base. Come già detto, qualunque coppia di vettori linearmente indipendenti che siano combinazione lineare di  $\mathbf{u}$  e  $\mathbf{v}$  è ancora una coppia di autovettori di  $\mathcal{A}$  relativi all’autovalore  $\lambda = 3$ .

Possiamo concludere dunque che gli autovettori ottenuti sono solo due possibili autovettori del piano di autovettori da essi individuato, ma allora c'è la possibilità di scegliere una qualsiasi coppia di autovettori come base dell'autospazio stesso.

In questo caso però si presenta il problema di trovare l'autovettore da scrivere come termine noto nel sistema di equazioni algebriche avente come matrice dei coefficienti la matrice

$$\begin{pmatrix} 27-3 & -40 & -8 \\ 15 & -22-3 & -5 \\ -3 & 5 & 4-3 \end{pmatrix} = \begin{pmatrix} 24 & -40 & -8 \\ 15 & -25 & -5 \\ -3 & 5 & 1 \end{pmatrix}$$

Allora non possiamo fare nessuna scelta dell'autovettore da porre come termine noto e dobbiamo considerare il sistema con l'espressione generale dell'autospazio

$$\begin{pmatrix} 24 & -40 & -8 \\ 15 & -25 & -5 \\ -3 & 5 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ 3\alpha - 5\beta \end{pmatrix}$$

Come detto, la matrice dei coefficienti ha caratteristica 1 e allora dobbiamo determinare i valori di  $\alpha$  e  $\beta$  in modo che anche la matrice completa del sistema abbia caratteristica 1: a tale scopo dobbiamo rendere nulli tutti i determinanti dei suoi minori di ordine 2, in particolare il determinante del minore formato dalle ultime due righe e colonne

$$\det \begin{pmatrix} -5 & \beta \\ 1 & 3\alpha - 5\beta \end{pmatrix} = 24\beta - 15\alpha$$

Abbiamo  $24\beta - 15\alpha = 0$  per  $\alpha = 8$  e  $\beta = 5$ , da cui segue che l'autovettore

$$\mathbf{v}' = \begin{pmatrix} \alpha \\ \beta \\ 3\alpha - 5\beta \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \\ -1 \end{pmatrix}$$

scritto come termine noto, rende risolubile il sistema

$$\begin{pmatrix} 24 & -40 & -8 \\ 15 & -25 & -5 \\ -3 & 5 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \\ -1 \end{pmatrix}$$

Tale sistema possiede, per il teorema di Rouché-Capelli,  $\infty^2$  soluzioni, una delle quali è data da

$$x = 0, \quad y = 0, \quad z = -1$$

che fornisce appunto il terzo vettore  $\mathbf{w} = (0, 0, -1)$

A questo punto costruiamo la matrice  $\mathcal{T}$  del cambio di base dalla base iniziale  $\mathcal{B}$  alla nuova base  $\mathcal{B}' = \{\mathbf{u}, \mathbf{v}', \mathbf{w}\}$ .

La matrice  $\mathcal{T}$  è allora

$$\mathcal{T} = \begin{pmatrix} 1 & 8 & 0 \\ 0 & 5 & 0 \\ 3 & -1 & -1 \end{pmatrix}$$

La matrice  $\mathcal{J}$  di Jordan associata all'applicazione  $L$  relativamente alla nuova base  $\mathcal{B} = \{\mathbf{u}, \mathbf{v}', \mathbf{w}\}$  si determina infine mediante la (10)

$$\begin{aligned} \mathcal{J} &= \mathcal{T}^{-1} A \mathcal{T} = \\ &= \mathcal{T}^{-1} A \mathcal{T} = \begin{pmatrix} 1 & -8/5 & 0 \\ 0 & 1/5 & 0 \\ 3 & -5 & -1 \end{pmatrix} \begin{pmatrix} 27 & -40 & -8 \\ 15 & -22 & -5 \\ -3 & 5 & 4 \end{pmatrix} \begin{pmatrix} 1 & 8 & 0 \\ 0 & 5 & 0 \\ 3 & -1 & -1 \end{pmatrix} = \\ &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \end{aligned}$$

## 4 Esercizio risolto

**Testo:** nello spazio vettoriale  $\mathcal{R}^4$ , riferito alla base naturale  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ , sono assegnati i vettori  $\mathbf{u} = (2, 0, 0, -1)$ ,  $\mathbf{v} = (0, 0, 1, -1)$ ,  $\mathbf{w} = (2, 0, 1, -1)$  e l'applicazione lineare  $L : \mathcal{R}^4 \rightarrow \mathcal{R}^4$ .

Sapendo che  $L(\mathbf{w}) \in N(L)$  e che vale inoltre

$$L(\mathbf{u}) = (-1, -1, 1, 2), \quad L(\mathbf{e}_2) = (1, 1, 5, -2), \quad L(\mathbf{v}) = (0, 0, -4, 0)$$

1. scrivere, rispetto alla base naturale, la matrice  $\mathcal{A}'$  associata all'applicazione lineare  $L : \mathcal{R}^4 \rightarrow \mathcal{R}^4$
2. determinare il nucleo  $N(L)$  e l'immagine  $Im(L)$ , precisando se  $N(L) + Im(L)$  è una somma diretta oppure no;
3. trovare una base del sottospazio somma  $N(L) + Im(L)$  e del sottospazio intersezione  $N(L) \cap Im(L)$ ;
4. stabilire quali vettori del sottospazio  $S = \{(x, y, z, t) \in \mathcal{R}^4, \text{ con } x - z = 0\}$  appartengono a  $Im(L)$  e quali vettori di  $\mathcal{R}^4$  hanno immagine appartenente al sottospazio  $S$ ;
5. determinare la proiezione del vettore  $\phi = (0, 0, 0, 1) \in \mathcal{R}^4$  sul sottospazio vettoriale somma  $N(L) + Im(L)$ .

**Soluzione:** 1) le relazioni assegnate possono essere scritte nella forma

$$\begin{cases} L(\mathbf{u}) = (-1, -1, 1, 2) = -\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 + 2\mathbf{e}_4 \\ L(\mathbf{e}_2) = (1, 1, 5, -2) = \mathbf{e}_1 + \mathbf{e}_2 + 5\mathbf{e}_3 - 2\mathbf{e}_4 \\ L(\mathbf{v}) = (0, 0, -4, 0) = 0\mathbf{e}_1 + 0\mathbf{e}_2 - 4\mathbf{e}_3 + 0\mathbf{e}_4 \\ L(\mathbf{w}) = (0, 0, 0, 0) = 0\mathbf{e}_1 + 0\mathbf{e}_2 + 0\mathbf{e}_3 + 0\mathbf{e}_4 \end{cases}$$

da cui ricaviamo che nello spazio  $V$  come spazio di partenza si considera la base iniziale  $\mathcal{B}' = \{\mathbf{u}, \mathbf{e}_2, \mathbf{v}, \mathbf{w}\}$  e nello spazio  $V$  come spazio di arrivo si considera la base naturale perché l'applicazione è assegnata attraverso le leggi di trasformazione dei vettori della base  $\mathcal{B}$  e i risultati sono espressi come combinazione lineare dei vettori della base naturale.

Rispetto a tali basi, all'applicazione  $L$  è associata la matrice  $\mathcal{A}$  ottenuta incolonnando i coefficienti delle immagini

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

Per ottenere la matrice  $\mathcal{A}'$  associata a  $L$  rispetto alla base naturale in  $V$  sia come spazio di partenza che come spazio d'arrivo, occorre effettuare il cambio di base soltanto in  $V$  come spazio di partenza, perché in  $V$  come spazio d'arrivo i risultati sono già espressi come combinazione lineare dei vettori della base naturale.

Costruiamo allora la matrice  $\mathcal{T}_1$  del cambio di base dalla base  $\mathcal{B}'$  alla base naturale  $\mathcal{B}$  in  $V$  come spazio di partenza. A tale scopo esprimiamo i vettori della base naturale (che è la nuova base!) come combinazione lineare dei vettori della (vecchia) base  $\mathcal{B}'$

$$\begin{cases} \mathbf{e}_1 = 0\mathbf{u} + 0\mathbf{e}_2 - \frac{1}{2}\mathbf{v} + \frac{1}{2}\mathbf{w} \\ \mathbf{e}_2 = 0\mathbf{u} + 1\mathbf{e}_2 + 0\mathbf{v} + 0\mathbf{w} \\ \mathbf{e}_3 = -1\mathbf{u} + 0\mathbf{e}_2 + 0\mathbf{v} + 1\mathbf{w} \\ \mathbf{e}_4 = -1\mathbf{u} + 0\mathbf{e}_2 - 1\mathbf{v} + 1\mathbf{w} \end{cases}$$

da cui, incolonnando i coefficienti dei risultati, otteniamo la matrice  $\mathcal{T}_1$

$$\mathcal{T}_1 = \begin{pmatrix} 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ -1/2 & 0 & 0 & -1 \\ 1/2 & 0 & 1 & 1 \end{pmatrix}$$

e quindi, utilizzando la (11) in cui si riconosce che la matrice  $\mathcal{T}_2$  è la matrice identità, ricaviamo la matrice  $\mathcal{A}'$  associata ad  $L$  relativamente alla base naturale sia in  $V$  come spazio di partenza che in  $V$  come spazio di arrivo

$$\mathcal{A}' = \mathcal{A}\mathcal{T}_1 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 5 & -1 & 3 \\ 0 & -2 & -2 & -2 \end{pmatrix}$$

2) Ricordando che il nucleo di un'applicazione lineare  $L$  è l'insieme dei vettori  $(x, y, z, t)$  dello spazio di partenza che hanno come immagine il vettore nullo dello

spazio di arrivo, dobbiamo risolvere il sistema

$$\mathcal{A}' \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ovvero

$$\begin{cases} y + z + t = 0 \\ y + z + t = 0 \\ 2x + 5y - z + 3t = 0 \\ -2y - 2z - 2t = 0 \end{cases}$$

La matrice dei coefficienti di tale sistema ha caratteristica 2 e quindi le soluzioni, per il teorema di Rouché-Capelli, saranno  $\infty^2$ , potendosi ottenere con il minore relativo alle incognite  $x, y$  nella terza e quarta equazione

$$\begin{cases} 2x + 5y = \alpha - 3\beta \\ 2y = -2\alpha - 2\beta \end{cases}$$

A questo punto si determinano le  $\infty^2$  soluzioni

$$\begin{cases} x = 3\alpha + \beta \\ y = -\alpha - \beta \\ z = \alpha \\ t = \beta \end{cases}$$

che, poste nella forma vettoriale, costituiscono il nucleo di  $L$

$$N(L) = \{(x, y, z, t) \in \mathcal{R}^4, \quad \text{con } (x, y, z, t) = \alpha(3, -1, 1, 0) + \beta(1, -1, 0, 1)\}$$

individuato dai due vettori di base  $(3, -1, 1, 0)$ ,  $(1, -1, 0, 1)$ .

Per quanto riguarda l'immagine, ricordiamo che un vettore  $(a, b, c, d)$  appartiene all'immagine di  $L$  se si ottiene come risultato di qualche vettore  $(x, y, z, t)$ , ovvero se il sistema

$$\begin{cases} y + z + t = a \\ y + z + t = b \\ 2x + 5y - z + 3t = c \\ -2y - 2z - 2t = d \end{cases}$$

è risolubile e il vettore  $(x, y, z, t)$  ha appunto come immagine il vettore  $(a, b, c, d)$ .

Poiché la matrice dei coefficienti, come detto, ha caratteristica 2, allora perché il sistema sia risolubile, anche la matrice completa deve avere caratteristica 2, ovvero si debbono annullare tutti i suoi minori di ordine 3.

In questo caso particolare possiamo anche procedere confrontando le equazioni e osservando che i termini noti  $a, b$  debbono essere uguali e che il termine noto  $d$  deve essere pari a  $-2a$ .



Abbiamo dunque che l'immagine di  $L$  è data dall'insieme

$$Im(L) = \{(a, b, c, d) \in \mathcal{R}^4, \quad \text{con } (a, b, c, d) = (\gamma, \gamma, \delta, -2\gamma)\}$$

ed è un sottospazio individuato dai due vettori  $(1, 1, 0, -2)$ ,  $(0, 0, 1, 0)$ .

Osserviamo che la somma della dimensione 2 del nucleo e della dimensione 2 dell'immagine di  $L$  è pari alla dimensione 4 di  $V$ .

Per stabilire se il sottospazio somma del nucleo e dell'immagine di  $L$ , che sono entrambi sottospazi appartenenti a  $\mathcal{R}^4$ , è una somma diretta oppure no, occorre stabilire se il nucleo e l'immagine di  $L$  hanno in comune soltanto il vettore nullo o anche altri vettori, cioè dobbiamo verificare se esistono, oppure non esistono, vettori ottenibili contemporaneamente con una particolare scelta di  $\alpha, \beta$  (come vettori del nucleo) e di  $\gamma, \delta$  (come vettori dell'immagine).

Uguagliando il generico vettore del nucleo e il generico vettore dell'immagine, si ottiene allora il sistema

$$\begin{cases} 3\alpha + \beta - \gamma & = 0 \\ -\alpha - \beta - \gamma & = 0 \\ \alpha & - \delta = 0 \\ \beta + 2\gamma & = 0 \end{cases}$$

che possiede  $\infty^1$  soluzioni date da  $(\alpha, \beta, \gamma, \delta) = k(1, -2, 1, 1)$ .

Segue dunque che il vettore  $(1, 1, 1, -2)$  appartiene al nucleo perché si ottiene dall'espressione del nucleo con  $\alpha = 1$  e  $\beta = -2$  e appartiene all'immagine perché si ottiene dall'espressione dell'immagine con  $\gamma = \delta = 1$ .

Concludiamo pertanto che il sottospazio somma  $N(L) + Im(L)$  non è una somma diretta perché il sottospazio intersezione  $N(L) \cap Im(L)$  contiene vettori diversi dal vettore nullo.

3) Il sottospazio intersezione  $N(L) \cap Im(L)$  ha dunque come base il vettore comune ai due sottospazi  $(1, 1, 1, -2)$ ; il sottospazio somma  $N(L) + Im(L)$  ha come base i tre vettori  $(3, -1, 1, 0)$ ,  $(1, 1, 0, -2)$ ,  $(1, 1, 1, -2)$ , i primi due dei quali appartengono rispettivamente soltanto al nucleo e soltanto all'immagine di  $L$  e il terzo è il vettore in comune al nucleo e all'immagine di  $L$ .

4) Il sottospazio  $S$  ha rappresentazione parametrica data dal vettore  $(a, b, c, a)$  in cui compaiono tre parametri perché  $S$  è un sottospazio di dimensione 3, come si vede risolvendo il sistema con la sola equazione  $x - z = 0$  in quattro variabili.

Per trovare i vettori di  $S$  che appartengono all'immagine di  $L$ , ovvero l'intersezione  $S \cap Im(L)$ , uguagliamo il generico vettore di  $Im(L)$  dato da  $(\gamma, \gamma, \delta, -2\gamma)$  al generico vettore di  $S$  dato da  $(a, b, c, a)$ . Tale uguaglianza si scrive allora sotto forma di sistema

$$\begin{cases} a & - \gamma & = 0 \\ b & - \gamma & = 0 \\ c & - \delta & = 0 \\ a & + 2\gamma & = 0 \end{cases}$$

che, per il teorema di Rouché-Capelli, possiede  $\infty^1$  soluzioni ottenibili con il minore di ordine 4 relativo alle incognite  $a, b, c, \gamma$ , rappresentate dal vettore  $k(0, 0, 1, 0)$ .

Poiché  $S \cap Im(L)$  è un sottospazio di dimensione 1, allora il vettore  $(0, 0, 1, 0)$ , che rappresenta le  $\infty^1$  soluzioni del sistema, è il vettore di base di  $S \cap Im(L)$ .

Per determinare i vettori di  $\mathcal{R}^4$  che hanno immagine appartenente a  $S$ , osserviamo che, per quanto appena detto, non tutti i vettori di  $S$  sono immagine di qualche vettore di  $\mathcal{R}^4$  come spazio di partenza, ma soltanto quelli appartenenti a  $S \cap Im(L)$ . Poiché i vettori appartenenti a  $S \cap Im(L)$  sono solo quelli del tipo  $2k(0, 0, 1, 0)$  (dove per comodità si è scritto il coefficiente con  $2k$ ), allora cerchiamo i vettori di  $\mathcal{R}^4$  che hanno immagine data da  $(0, 0, 2k, 0)$ , ovvero risolviamo il sistema

$$\begin{cases} y + z + t = 0 \\ y + z + t = 0 \\ 2x + 5y - z + 3t = 2k \\ -2y - 2z - 2t = 0 \end{cases}$$

che, attraverso il minore relativo alle incognite  $x, y$  nella terza e quarta equazione, fornisce la soluzione  $(x, y, z, t) = k(1, 0, 0, 0) + \alpha(3, -1, 1, 0) + \beta(1, -1, 0, 1)$  in cui si riconosce che la struttura del generico vettore che ha immagine in  $S$ , cioè nell'intersezione  $S \cap Im(L)$ , è data dalla somma di un multiplo del vettore  $(1, 0, 0, 0)$ , che ha immagine in  $S$ , e di un qualsiasi vettore del nucleo, che ha immagine nulla sotto azione di  $L$ .

5) Per trovare la proiezione, indicata con  $\psi$ , del vettore  $\phi$  sul sottospazio  $N(L) + Im(L)$ , utilizziamo i vettori di base di  $N(L) + Im(L)$  già determinati  $\psi_1 = (3, -1, 1, 0)$ ,  $\psi_2 = (1, 1, 0, -2)$ ,  $\psi_3 = (1, 1, 1, -2)$  in modo da esprimere tale proiezione come combinazione lineare  $\psi = \alpha_1\psi_1 + \alpha_2\psi_2 + \alpha_3\psi_3$  dei vettori di base mediante coefficienti incogniti  $\alpha_1, \alpha_2, \alpha_3$ .

I coefficienti  $\alpha_1, \alpha_2, \alpha_3$  si ottengono risolvendo il sistema  $\langle \phi - \psi, \psi_i \rangle = 0$ , ovvero il sistema  $\langle \phi - \alpha_1\psi_1 - \alpha_2\psi_2 - \alpha_3\psi_3, \psi_i \rangle = 0$ , che, con la separazione delle incognite e dei termini noti, assume la forma  $\alpha_1\langle \psi_1, \psi_i \rangle + \alpha_2\langle \psi_2, \psi_i \rangle + \alpha_3\langle \psi_3, \psi_i \rangle = \langle \phi, \psi_i \rangle$ , con  $i = 1, 2, 3$ . Eseguendo i prodotti scalari  $\langle \psi_j, \psi_i \rangle$  e  $\langle \phi, \psi_i \rangle$ , il sistema diventa

$$\begin{cases} 11\alpha_1 + 2\alpha_2 + 3\alpha_3 = 0 \\ 2\alpha_1 + 6\alpha_2 + 6\alpha_3 = -2 \\ 3\alpha_1 + 6\alpha_2 + 7\alpha_3 = -2 \end{cases}$$

la cui soluzione è  $\alpha_1 = 1/14$ ,  $\alpha_2 = -2/7$ ,  $\alpha_3 = -1/14$ . Concludiamo quindi che la proiezione del vettore  $\phi = (0, 0, 0, 1)$  sul sottospazio somma  $N(L) + Im(L)$  è data dal vettore

$$\psi = \alpha_1\psi_1 + \alpha_2\psi_2 + \alpha_3\psi_3 = \left(-\frac{1}{7}, -\frac{3}{7}, 0, \frac{5}{7}\right)$$

## Riferimenti bibliografici

1. Alessandro Blasi, *Matematica per le Applicazioni Economiche e Finanziarie*, Ed. Kappa, 1998;
2. Serge Lang, *Algebra Lineare*, Bollati Boringhieri, 1970.