# Una chiacchierata sull'Algebra Lineare

**Giuseppe Paxia**

March 16, 2020

# Chapter 1

# Elementi di base

## 1.1 I numeri complessi

Si suppone che il lettore abbia adeguata conoscenza del campo dei numeri reali $\mathbb{R}$ e delle sue principali proprietà.

Ricordiamo la definizione di **gruppo abeliano**

**Definizione 1.1.1.** Un insieme $G$ si dice un **gruppo abeliano** se è definita un'operazione $+$ che gode delle seguenti proprietà:

1) per ogni scelta di elementi in $G$ vale la proprietà associativa $a + (b + c) = (a + b) + c$;

2) esiste l'elemento 0 neutro per la somma, cioè: a+0=0+a=a per ogni $a \in G$;

3) per ogni $a \in G$ esiste un elemento $a'$ tale che $a + a' = a' + a = 0$;

4) vale la proprietà commutativa della somma.

Quando si parla di un **campo** si pensa soprattutto ai campi numerici a tutti noti: campo dei numeri razionali $\mathbb{Q}$, campo dei numeri reali $\mathbb{R}$ e campo dei numeri complessi $\mathbb{C}$. In effetti esistono altri campi tipo i campi finiti, il campo dei quaternioni e così via. In queste note ci riferiremo essenzialmente ai campi $\mathbb{R}$ e $\mathbb{C}$. E se non è necessario precisare quale scriveremo $K$.

Trattiamo adesso il campo dei **numeri complessi** $\mathbb{C}$ che è il campo immediatamente più ampio del campo $\mathbb{R}$, studiandone le principali proprietà.

**Definizione 1.1.2.** Un numero complesso $z$ è una coppia ordinata $(a, b)$ di numeri reali; si scrive $z = (a, b)$.

1) Due numeri complessi $(a, b)$ e $(a', b')$ sono **uguali** se $a = a'$ e $b = b'$.

2) La **somma** dei numeri complessi si definisce :

$$\boxed{\text{(a,b)+(a',b')=(a+a',b+b')}}.$$

Rispetto a questa operazione $\mathbb{C}$ risulta **un gruppo abeliano**. Lo **zero** è la coppia $(0,0)$. L'**opposto** di $(a,b)$ è la coppia $(-a,-b)$.

3) Il **prodotto** di due numeri complessi $(a,b)\cdot(a',b')$ si definisce:

$$\boxed{\text{(a,b)(a',b')=(aa'-bb',ab'+a'b)}}.$$

Rispetto a tale prodotto gli elementi $(a,b)\neq(0,0)$ formano un gruppo abeliano. L'**elemento neutro del prodotto** è il numero $(1,0)$.

.

**Osservazione** 1.1.3. ll sottocampo dei numeri complessi del tipo $(a,0)$ si identifica col campo dei numeri reali $\mathbb{R}$. Si dice che c'è un **isomorfismo** tra i numeri complessi del tipo $(a,0)$ e i numeri reali.

Due strutture algebriche sono **isomorfe** se esiste una corrispondenza biunivoca fra i loro elementi che "conserva" le operazioni.

Consideriamo il numero complesso $(0,1)$. Moltiplichiamolo per se stesso $(0,1)\cdot(0,1)$, usando la definizione di prodotto che abbiamo appena introdotto. Si ha $(0,1)\cdot(0,1) = (-1,0)$. **Otteniamo un risutato inatteso**. Cioè il quadrato del numero complesso $(0,1)$ risulta essere un numero reale negativo. **Tale numero è speciale. Si pone** $(0,1)=i$**, e si chiama l'unità immaginaria**.

È facile verificare tutte le proprietà formali delle operazioni che valgono per gli altri campi numerici. Il numero complesso $(a,b)$ si può scrivere nella forma $a+bi$, che si chiama **forma algebrica**.

**Osservazione** 1.1.4. Osserviamo che le operazioni coi numeri complessi in forma algebrica si effettuano usando le regole del calcolo algebrico note e tenendo conto che basta sostituire $-1$ ogni qualvolta si trova $i^2$.

*Esempio* 1.1.5. Troviamo l'inverso moltiplicativo del numero $a+bi\neq(0,0)$. Sarà un numero $x+yi$ che moltiplicato per $a+bi$ è tale che $(a+bi)\cdot(x+yi)=(1,0)$. Si ha $(ax-by)+(ay+bx)i$. Identificando le componenti si ha

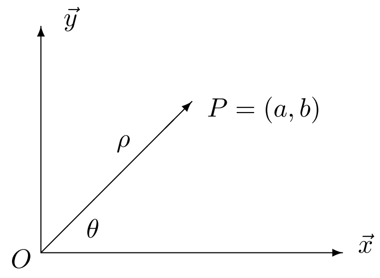
$$\begin{cases} ax-by &= 1 \\ bx+ay &= 0 \end{cases}$$

Risolvendo il sistema si ottiene

$$x=\frac{a}{a^2+b^2}, \qquad y=\frac{-b}{a^2+b^2}$$

che sono le componenti dell'inverso del numero dato.

Dato il numero complesso $a + ib$ si definisce suo **complesso coniugato** il numero $a - ib$. È facile osservare che la somma e il prodotto di due numeri complessi coniugati è sempre reale. Si definisce poi **modulo del numero complesso $a+ib$ il numero** $\sqrt{a^2 + b^2}$. Solitamente il modulo del numero complesso $z = a + ib$ si indica $|z|$.




Osservando la figura si vede che il numero complesso $(a, b)$ individua il punto $P$ e quindi il raggio vettore $\overrightarrow{OP}$. In effetti c'è una corrispondenza biunivoca tra i numeri complessi e i punti del piano, ed anche tra i numeri complessi e i vettori del piano applicati in $O$.

Da quanto sopra si deduce facilmente che $a = \rho \cos\theta$ e $b = \rho \sin\theta$. Quindi il numero complesso $a + bi = \rho(\cos\theta + i \sin\theta)$. Scritto in questo modo si dice che è in **forma trigonometrica**.
Due numeri $\rho(\cos\theta + i\sin\theta)$ e $\rho'(\cos\theta' + i\sin\theta')$ sono **uguali** se $\rho = \rho'$ e $\theta - \theta' = 2k\pi$ per qualche $k \in \mathbb{Z}$. Si suole anche dire che $\theta \equiv \theta', \quad \mod 2\pi$ e si dice che $\theta$ è congruo a $\theta'$ mod.$2\pi$
Effettuiamo il prodotto di due numeri complessi scritti in forma trigonometrica.

$$\begin{aligned}
&\rho(\cos\theta + i\sin\theta)\rho'(\cos\theta' + i\sin\theta') = \\
&\quad \rho\rho'[(\cos\theta\cos\theta' - \sin\theta\sin\theta') + i(\sin\theta\cos\theta' + \cos\theta\sin\theta') = \\
&\quad \rho\rho'[\cos(\theta + \theta') + i(\sin\theta + \theta')]
\end{aligned}$$

Si deduce facilmente che $[\rho(\cos\theta + i\sin\theta)]^2 = \rho^2(\cos 2\theta + i\sin 2\theta)$ e più in generale si dimostra, per induzione, che vale la seguente formula, detta di De Moivre

**Formula 1 (De Moivre).** *La potenza* $[\rho(\cos\theta + i\sin\theta)]^n = \rho^n(\cos n\theta + i\sin n\theta)$.

Facciamo adesso qualche applicazione. È molto importante la risoluzione dell'equazione $X^n = \alpha$, con $n \in \mathbb{N}$ e $\alpha \in \mathbb{C}$ numero complesso.

**Formula 2.** *L'equazione* $\boxed{X^n = \alpha}$ *ammette n radici complesse che sono dette le radici n-me del numero complesso* $\alpha$*.*

*Proof.* Sia $\alpha$ un numero complesso qualunque, eventualmente reale; scriviamolo in forma trigonometrica $\alpha = r(cos\varphi + i\sin\varphi)$. Cerchiamo numeri complessi incogniti $\rho(\cos\theta + i\sin\theta)$ che soddisfino l'equazione

$$[\rho(\cos\theta + i\sin\theta)]^n = \rho^n(\cos n\theta + i\sin n\theta) = r(cos\varphi + i\sin\varphi).$$

Ricordando come si esprime l'eguaglianza fra numeri complessi in forma trigonometrica si ha

$$\rho^n = r \qquad \text{e} \qquad n\theta - \varphi = 2\pi k, \text{ con } k \in \mathbb{Z}$$

Da ciò si deduce che $\rho = \sqrt[n]{r}$, radice $n$-ma aritmatica di $r$, mentre $\theta$ si ricava da

$$\theta = \frac{\varphi + 2\pi k}{n}, \qquad \text{con} \quad k \in \mathbb{Z}$$

Per ottenere soluzioni distinte non deve accadere che $\theta_1 - \theta_2$ (ottenuti dalla formula in corrispondenza di $k_1$ e $k_2$) sia un multiplo di $2\pi$. Calcoliamo

$$\frac{\varphi + 2\pi k_1}{n} - \frac{\varphi + 2\pi k_2}{n} = 2\pi h$$

dalla quale si deduce $k_1 - k_2 = hn$.

Coppie di valori di $k$ appartenenti all' intervallo $\{0, 1, \ldots, n-1\}$ risultano certamente a due a due incongrue mod $n$, perché la loro differenza è sempre minore di $n$, e quindi non può essere un multiplo di $n$. Allora possiamo dire che le $n$ radici $n$-me distinte di $\alpha = r(\cos\varphi + i\sin\varphi)$ si trovano con la formula:

$$X_i = \sqrt[n]{r}(\cos\frac{\varphi + 2k_i\pi}{n} + i\sin\frac{\varphi + 2k_i\pi}{n}) \qquad per \qquad k_i \in \{0, 1, , \ldots, n-1\}$$

□

Osserviamo che se si sceglie un qualunque numero $p \in \mathbb{Z}$ e lo si divide per $n$ si ha $p = qn + r$ con $0 \le r \le n-1$, da cui $p - r = qn$. Quindi $p$ è congruo ad $r$ mod $n$. Possiamo dire che le varie determinazioni ottenute si distribuiscono nelle $n$ classi dei resti mod $n$.

*Esempio* 1.1.6. Calcoliamo le radici cubiche del numero -1. Tale numero in forma trigonometrica si può scrivere $-1 = \cos\pi + i\sin\pi$. In tal caso la formula precedente si può scrivere

$$X_i = \cos\frac{\pi + 2\pi k_i}{3} + i\sin\frac{\pi + 2\pi k_i}{3} \qquad \text{per} \qquad i = 0, 1, 2$$

Si ottiene facilmente

$$\begin{aligned}
X_0 &= \cos\frac{\pi}{3} + i\sin\frac{\pi}{3} = \frac{1}{2} + i\frac{\sqrt{3}}{2} \\
X_1 &= \cos\frac{3\pi}{3} + i\sin\frac{3\pi}{3} = -1 \\
X_2 &= \cos\frac{\pi + 4\pi}{3} + i\sin\frac{\pi + 4\pi}{3} = \frac{1}{2} - i\frac{\sqrt{3}}{2}.
\end{aligned}$$

A questo punto è bene esercitarsi con espressioni contenenti numeri complessi in forma algebrica ed estrazioni di radici $n$-me di numeri complessi.

Per coloro che hanno delle conoscenze di Analisi voglio citare una formula molto interessante che si chiama la **Formula di Eulero** che dice :

$$\cos\theta + i\sin\theta = e^{i\theta}.$$

Per $\theta = \pi$ tale formula diventa: $\boxed{e^{\pi i} + 1 = 0}$ e viene detta **Idemtità di Eulero**.

Io non mi azzarderei a dire quello che ha detto il premio Nobel della Fisica Feynman e cioè che la formula di Eulero è la più straordinaria formula della matematica. Ma il legame tra **cinque** importanti numeri ( $e, \pi, i, 1, 0$) è veramente sorprendente.

È facile dimostrare tale formula senza fare uso degli sviluppi in serie. Basta ricordare che il numero complesso $z = a + ib$ sulla circonferenza unitaria si può scrivere : $z = \cos\theta + i\sin\theta$. Derivando rispetto a $\theta$ si ha;

$$\frac{dz}{d\theta} = -\sin\theta + i\cos\theta = i(i\sin\theta + \cos\theta) = iz.$$

Separando le variabili si ha $\frac{dz}{z} = i\theta$. Integrando ambo i membri si ha : $\log_e z = i\theta$, da cui si deduce : $e^{i\theta} = z = \cos\theta + i\sin\theta$.

Tale formula è molto usata nelle applicazioni dell' Ingegneria, ma è abbastanza agevole perché questa funzione esponenziale complessa si comporta formalmente come la funzione esponenziale reale.

## 1.2 Polinomi

In questa sezione trattiamo alcune proprietà dei polinomi che saranno utili nel seguito. Innanzitutto esaminiamo proprietà che valgono sia per i polinomi a coefficienti reali che quelli a coefficienti complessi. Quindi indichiamo con $K$ il campo $\mathbb{R}$ oppure $\mathbb{C}$.

La totalità dei polinomi a coefficienti in $K$ nella indeterminata $X$ si suole indicare $K[X]$. Un **polinomio** $p(X) \in K[X]$ è una espressione

$$p(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0. \tag{1.1}$$

Si può anche dire che un polinomio è una funzione $p : K \to K$. Se i coefficienti sono tutti nulli si ha il polinomio nullo 0, al quale, per convenzione, si attribuisce il grado $-\infty$. Se nella espressione (1) $a_n \neq 0$ si dice che il polinomio ha **grado** $n$.

Un elemento $\alpha \in K$ si dice una **radice** del polinomio $p(X)$ se accade che $p(\alpha) = 0$. Un elemento $\alpha \in K$ si dirà **radice multipla di molteplicità** $r$ se $p(X) = (X - \alpha)^r q(X)$, con $q(\alpha) \neq 0$. I polinomi di grado zero sono le costanti, quelli di grado uno sono del tipo $a_1 X + a_0$, e cosi via. Sussiste il seguente

**Theorema 1.2.1.** *Condizione necessaria e sufficiente perché $\alpha$ sia radice del polinomio $p(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ di grado $n \geq 1$ è che esista un polinomio $q(X)$ di grado $n - 1$ per cui si abbia identicamente*

$$p(X) = (X - \alpha) q(X) \tag{1.2}$$

*Proof.* In una direzione è ovvio. Supponiamo che

$$p(\alpha) = a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0. \tag{1.3}$$

Sottraendo la (3) da $p(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ si ottiene

$$p(X) = a_n (X^n - \alpha^n) + a_{n-1} (X^{n-1} - \alpha^{n-1}) + \cdots + a_1 (X - \alpha) \tag{1.4}$$

che si può ovviamente decomporre in $p(X) = (X - \alpha) q(X)$, con $q(X)$ polinomio di grado $n - 1$, in quanto il coefficiente di grado massimo è $a_n \neq 0$. Se ci si vuole convincere con maggiori dettagli basta ricordare che $X^i - \alpha^i = (X - \alpha)(X^{i-1} + X^{i-2}\alpha + \cdots + X\alpha^{i-2} + \alpha^{i-1})$. □

Dal teorema precedente segue

**Corollario 1.2.2.** *Ogni polinomio $p(X)$ di grado $n \geq 0$ ha al più $n$ radici distinte.*

*Proof.* Se $n = 0$ il polinomio $p(X) = a_0 \neq 0$; ovviamente non ha radici. Per $n = 1$ il polinomio è deltipo $p(X) = a_1 X + a_0$ con $a_1 \neq 0$. Questo polinomio ammette la radice $\alpha = -\frac{a_0}{a_1}$. Ragioniamo per induzione su $n$. Se $p(X)$ non ha radici il teorema è vero. Ammettiamo che la proprietà sia vera per tutti i polinomi di grado $n - 1$. Se $p(X)$ ammette la radice $\alpha$ per il teorema precedente si ha $p(X) = (X - \alpha) q(X)$, con $q(X)$ di grado $n - 1$. Allora ci sono al più $n$ radici distinte. □

Da quanto precede si ha la ovvia conseguenza: se il polinomio $p(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ è **identicamente nullo**, cioé è nullo per ogni $X$, allora tutti i coefficienti $a_i$ devono essere nulli. Altrimenti detto $k$ il massimo per cui $a_k \neq 0$, ci sarebbero al più $k$ radici distinte e non già infinite.

**Proposizione 1.2.3.** *(***Algoritmo di divisione***) Siano $f(X)$ e $g(X) \neq 0$ due polinomi, di gradi $m$ ed $n$ in $K[X]$. Allora esistono, e sono unici, due polinomi $q(X)$ e $r(X)$ tali che si abbia*

$$f(X) = q(X)g(X) + r(X)$$

*con* $\deg r(X) < \deg g(X)$*.*

*Proof.* Se $m < n$ il risultato è banale: basta scegliere $q(X) = 0$ e $r(X) = f(X)$. Supponiamo che sia $m \geq n$. Prendiamo un polinomio $q(X)$ tale che $f(X) - q(X)g(X)$ abbia il grado il minimo possibile. Ponendo $r(X) = f(X) - q(X)g(X)$ si vede che la prima richiesta è soddisfatta. Ci resta da provare che $\deg r(X) < \deg g(X)$. Per assurdo supponiamo che $\deg r(X) \geq \deg g(X)$.
Consideriamo

$$f(X) - (q(X) + hX^i)g(X) = r(X) - hX^i g(X). \tag{1.5}$$

Adesso per $i = \deg r(X) - \deg g(X)$, scegliamo $h$ in modo che i due polinomi all'ultimo membro abbiano lo stesso coefficiente in grado $\deg r(X)$. Ciò contraddice la nostra ipotesi perché la (5) fornisce i polinomi $q'(X) = (q(X) + hX^i)$ e $r'(X) = r(X) - hX^i g(X)$, e quest'ultimo con grado minore del grado che avevamo assunto nella nostra scelta.

Per l'unicità supponiamo che ci siano due polinomi $q'(X)$ ed $s'(X)$ tali che $f(X) = q'(X)g(X) + r'(X)$, con $\deg r'(X) < \deg q(X)$. Dalle due relazioni per differenza si deduce $(q(X) - q'(X))g(X) = r(X) - r'(X)$. Ma i gradi dei polinomi dei due menbri sono differenti, quindi l'eguaglianza può sussistere solo se i polinomi dei due membri sono nulli. da cui segue facilmente l'unicità. □

Consideriamo il polinomio $f(X) \in K[X]$, diremo che la radice $\alpha$ è una radice di **molteplicità** $m_\alpha$ per $f(X)$ se accade che $f(X) = (X - \alpha)^{m_\alpha} q(X)$, con $q(\alpha) \neq 0$.

A questo punto vogliamo ricordare un Teorema molto importante:

**Theorema 1.2.4. Teorema Fondamentale dell'Algebra***. Ogni polinomio non costante a coefficienti complessi ha una radice.*

*Proof.* **Per coloro che conoscono qualcosa di Analisi complessa**

Sia $p(z) \in \mathbb{C}[z]$ un polinomio non costante a coefficienti complessi. Supponiamo per assurdo che $p(z)$ non abbia alcuna radice. Possiamo considerare la funzione $\frac{1}{p(z)}$. Essa è una funzione analitica su tutto $\mathbb{C}$ e quindi è una funzione intera su $\mathbb{C}$. È facile vedere che $\lim_{z \to \infty} p(z) = \infty$ da cui segue $\lim_{z \to \infty} \frac{1}{p(z)} = 0$. Allora la funzione $\frac{1}{p(z)}$ è limitata. Per il teorema di Liouville (Ogni funzione analitica limitata è costante) ne segue che è costante. Quindi anche $p(z)$ è costante, contro l'ipotesi fatta. □

**Theorema 1.2.5.** *Ogni polinomio non costante a coefficienti complessi ha una decomposizione unica, a meno dell'ordine dei fattori, del tipo*

$$f(X) = h(X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_n)$$

*con* $h, \alpha_1, \alpha_2, \cdots \alpha_n \in \mathbb{C}$*.*

Ci occupiamo adesso della decomponibilità di un polinomio $f(X) \in \mathbb{R}$.
Sussiste l'importante Teorema.

**Theorema 1.2.6.** *Se* $f(X)$ *è un polinomio non costante a coefficienti reali allora sussiste la seguente decomposizione, unica a meno dell'ordine*

$$f(X) = h(X - \alpha_1) \cdots (X - \alpha_r)(X^2 + \beta_1 X + \gamma_1) \cdots (X^2 + \beta_s X + \gamma_s) \quad (1.6)$$

*dove* $h, \alpha_1, \cdots, \alpha_r \in \mathbb{R}$ *ed* $(\beta_i, \gamma_i) \in \mathbb{R}^2$ *tali che* ${\beta_i}^2 - 4\gamma_i < 0$ *per ogni* $i \in \{1, 2, \cdots, s\}$*.*

Si osservi che nella precedente formula $r$ oppure $s$ possono essere nulli. Bisogna notare altresì il diverso tipo di decomposizione cui si perviene a seconda che il polinomio $f(X)$ che si considera sia a coefficienti reali oppure complessi.

# Chapter 2

# Spazi vettoriali

## 2.1 Generalità sugli spazi vettoriali

**Definizione 2.1.1.** Un $K$-spazio vettoriale $V$ è un insieme di "vettori" su un dato campo base $K$ di "scalari"per cui sono verificate le seguenti proprietà:

1. $V$ è dotato di una operazione di somma $+$ per cui risulta essere un gruppo abeliano.

2. È definita una operazione da $K \times V \to V$, indicata solitamente $\cdot$ che associa alla coppia $(a, v)$ il vettore $(a \cdot v) \in V$ che se non c'è ambiguità si scrive $av$, che verifichi le seguenti proprietà:
   Per ogni scelta di scalari e vettori sia:
   A) $(ab)v = a(bv)$ ed 1v=v
   B) $a(v_1 + v_2) = av_1 + av_2$ e $(a_1 + a_2)v = a_1v + a_2v$.

Quando parliamo di uno spazio vettoriale viene naturale pensare allo spazio dei vettori dello spazio ordinario oppure allo spazio $\mathbb{R}^3$ delle loro componenti. Ma è bene sottolineare che gli spazi vettoriali che studieremo possono essere di natura molto diversa: avremo spazi vettoriali di matrici, di applicazioni lineari, di polinomi, di funzioni.

*Proprietà* 1. Consideriamo un $K$-spazio vettoriale $V$. Segue subito che:

- $a0_V = 0_V \quad \forall a \in K. \quad 0_K v = 0_V \quad \forall v \in V$
- Se $av = 0_V$ allora o $a = 0_K$ oppure $v = 0_V$.

*Proof.* Per esempio proviamo che $a0_V = 0_V$. Si può scrivere $a0_V = a(0_V + 0_V) = a0_V + a0_V$.Addizionando ad ambo i membri l'opposto di $a0_V$ si ottiene facilmente $0_V = a0_V$.
Se $av = 0_V$ o accade che è $a = 0_K$ oppure se $a \neq 0_K$ esiste il suo inverso moltiplicativo $a^{-1}$ per cui si ha $a^{-1}av = 1v = v = 0_V$. □

Dato un $K$-spazio vettoriale $V$. Diremo che $v \in V$ è **combinazione lineare** dei vettori $v_1, v_2, \ldots, v_n \in V$ se $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. La totalità delle combinazioni lineari dei vettori $v_1, v_2, \ldots, v_n$ si indica $L(v_1, v_2, \ldots, v_n)$.

**Definizione 2.1.2.** Diremo che i vettori $v_1, v_2, \ldots, v_n$ sono dei **generatori** dello spazio $V$ se accade che $V = L(v_1, v_2, \ldots, v_n)$.

**Definizione 2.1.3.** Diremo che i vettori $v_1, v_2, \ldots, v_n$ sono **linearmente dipendenti** se accade $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$ con coefficienti $a_i$ non tutti nulli.

Ovviamente i vettori $v_1, v_2, \ldots, v_n$ si dicono **linearmente indipendenti** se $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$ sussiste se e solo se i coefficienti $a_i$ sono tutti nulli.

È importante la seguente

**Proposizione 2.1.4.** *Se i vettori* $v_1, v_2, \ldots, v_n$ *sono* **linearmente indipendenti** *le loro combinazioni lineari si scrivono* **in modo unico**.

*Proof.* Infatti supponiamo che il vettore $v \in V$ si possa scrivere
$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Da ciò segue che
$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0$. Per l'ipotesi della indipendenza dei vettori segue che $a_i = b_i$, per ogni $i$. $\square$

**Definizione 2.1.5.** Si dice che $B = (v_1, v_2, \ldots, v_n)$ è **una base** per il $K$-spazio vettoriale $V$ se ogni vettore $v \in V$ si può scrivere **in modo unico** come combinazione lineare dei vettori $v_1, v_2, \ldots, v_n$.

Ne consegue che **una base è un insieme di generatori linearmente indipendenti**.
È importante il seguente risultato

*Criterio* 1. **Indipendenza lineare** Condizione necessaria e sufficiente perché i vettori $v_1, v_2, \ldots, v_n$ siano **linearmente indipendenti** è che

(1) $v_1 \neq 0_V$;

(2) $v_i \notin L(v_1, v_2, \ldots, v_{i-1})$ per ogni $2 \leq i \leq n$.

*Proof.* Se i vettori sono linearmente indipendenti non può accadere che sia $v_1 = 0$ altrimenti la combinazione lineare $1v_1 + 0_Kv_2 +, \cdots + 0_Kv_n = 0_V$ andrebbe contro l'indipendenza dei vettori dati perché il coefficiente di $v_1$ è $1 \neq 0_K$.
D'altro canto non può essere che sia $v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1}$ altrimenti la combinazione lineare $a_1v_1 + a_2v_2 + \cdots + a_{i-1}v_{i-1} - v_i + 0_Kv_{i+1} + \cdots + 0_Kv_n = 0_V$ e ciò contraddice l'indipendenza dei vettori dati. In tal caso il coefficiente di $v_i$ è $-1 \neq 0_K$.
Supponiamo viceversa che sia $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0_V$. Se fosse $a_n \neq 0$ si potrebbe esprimere $v_n$ come combinazione liare dei precedenti; e ciò è contro l'ipotesi (2). Quindi deve essere $a_n = 0_K$. Ripetendo il ragionamento saranno nulli tutti i coefficienti $a_i$ per $2 \leq i \leq n$. Si ottiene così $a_1v_1 = 0_V$. Poiché $v_1 \neq 0_V$ per ipotesi seguirà $a_1 = 0_K$ che prova l'indipendenza dei vettori dati. $\square$

Detto questo è facile provare la seguente

**Proposizione 2.1.6.** *Se un $K$-spazio vettoriale $V$ è generato da $(w_1, w_2, \ldots, w_m)$ da tali generatori è sempre possibile estrarre una base.*

*Proof.* Si comincia a considerare $w_1$; se tale vettore è nullo lo si scarta; se $w_1 \neq 0_V$ lo si prende. Supponiamo di prenderlo. Si passa a $w_2$. Se $w_2 \notin L(w_1)$ lo si prende altrimenti si scarta se $w_2 \in L(w_1)$. Supponiamo di scartarlo, ma si osservi che con tale scarto non si altera la generazione dello spazio. Si continua con lo stesso procedimento ed alla fine rimarranno generatori di $V$ ma che soddisfano al criterio di indipendenza lineare. Quindi costituiranno una base. □

**Proposizione 2.1.7.** *Un insieme di vettori $(v_1, v_2, \ldots, v_r)$ linearmente indipendenti può sempre essere completato ad una base di $V$.*

*Proof.* Se $V = L(w_1, w_2, \ldots w_s)$ sarà anche $V = L(v_1, \ldots, v_r, w_1, \ldots w_s)$. Applicando il precedente metodo degli scarti successivi si potrà estrarre una base che certamente conterrà come primi $r$ elementi i vettori $v_1, v_2, \ldots, v_r$ perché essendo indipendenti soddisfano il criterio di indipendenza lineare. □

**Lemma 2.1.8.** *(***Steinitz***) Sia $V$ un $K$-spazio vettoriale e $(v_1, v_2, \ldots, v_m)$ un insieme di vettori* **linearmente indipendenti**. *Se $(w_1, w_2, \ldots, w_n)$ sono dei generatori di $V$ allora deve necessariamente essere $m \leq n$.*

*Proof.* Per assurdo sia $m > n$. Ovviamente i vettori $(v_1, w_1, w_2, \ldots, w_n)$ costituiscono dei generatori di $V$ ed è $v_1 = a_1w_1 + a_2w_2 + \cdots + a_nw_n$. Non può accadere che tutti gli $a_i$ siano nulli altrimenti $v_1$ sarebbe nullo e ciò non può essere per l'indipendenza dei vettori $(v_1, v_2, \ldots, v_m)$. Supponiamo che sia $a_1 \neq 0$. Allora è possibile esprimere $w_1 = a_1^{-1}v_1 - a_1^{-1}a_2w_2 - \cdots - a_1^{-1}a_nw_n$. Quindi il vettore $w_1$ si può sopprimere e i vettori $(v_1, w_2, \ldots, w_n)$ saranno ancora generatori di $V$.
Anche $(v_1, v_2, w_2, \ldots, w_n)$ saranno generatori ma $v_2$ è combinazione lineare $v_2 = a_1v_1 + b_2w_2 + \cdots + b_nw_n$. Non può accadere che tutti i $b_i = 0$ altrimenti sarebbe $v_2 = a_1v_1$ contro l'indipendenza dei vettori dati. Per fissare le idee sia $b_2 \neq 0$. da cui segue $w_2 = b_2^{-1}(v_2 - a_1v_1 - b_3w_3 - \cdots - b_nw_n)$. Ne segue che $(v_1, v_2, w_3, \ldots, w_n)$ sono generatori.
Si continua ripetendo lo stesso ragionamento e alla fine si arriva alla conclusione che $(v_1, v_2, \ldots, v_n)$ sono dei generatori di $V$. Ma poiché abbiamo supposto che $m > n$ sarà $v_m = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Ma ciò è assurdo per l'indipendenza dei vettori $(v_1, v_2, \ldots, v_m)$. □

Una immediata conseguenza del Lemma di Steinitz è la seguente

**Proposizione 2.1.9.** *In uno stesso spazio vettoriale $V$ tutte le basi hanno lo stesso numero di elementi.*

*Proof.* Se $A = (v_1, v_2, \dots, v_n)$ e $B = (w_1, w_2, \dots, w_m)$ sono due basi, per il lemma di Steitnitz $n \leq m$, perché i vettori di $A$ sono indipendenti e quelli di $B$ sono generatori. D'altro canto, invertendo i ruoli accade che $m \leq n$. Quindi $n = m$. □

Diamo adesso la seguente

**Definizione 2.1.10.** La **dimensione** di un $K$-spazio vettoriale $V$ è il numero di vettori di una sua base.

**Osservazione** 2.1.11. In tutto quanto precedentemente visto ci siamo sempre riferiti a spazi vettoriali di dimensione finita.
Ma in approfondimenti successivi si studieranno anche spazi vettoriali di dimensione infinita.

Portiamo adesso alcuni esempi. Non staremo ad entrare in noiosi dettagli per la verifica di proprietà ovvie ed elementari. .

*Esempi* 1. 1) Con $\mathbb{R}^3$ su $\mathbb{R}$ si indica lo spazio vettoriale delle terne ordinate di numeri reali $(x_1 x_2, x_3)$ con somma e prodotto esterno così definito:

A) $(x_1, x_2, x_3)+(y_1, y_2, y_3) = (x_1+y_1, x_2+y_2, x_3+y_3)$

B) $a(x_1, x_2, x_3) = (ax_1, ax_2, ax_3)$
per ogni scelta di scalari e vettori. La verifica degli assiomi di spazio vettoriale è immediata e la omettiamo. Questo spazio ha dimensione 3 ed una sua base è costituita dai vettori $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$. Tale base si suole chiamare la **base canonica o base standard**.

2) In modo analogo si definisce lo spazio vettoriale $\mathbb{R}^n$ su $\mathbb{R}$ delle $n$-uple ordinate di numeri reali con le operazioni ovvie. Tale spazio vettoriale ha dimensione $n$.

3) Con $\mathbb{R}_n[X]$ si indica lo spazio vettoriale dei polinomi a coefficienti reali nella indeterminata $X$ di grado non superiore ad $n$. È uno spazio vettoriale con le operazioni ovvie una cui base è costituita dai polinomi $(1, X, X^2, \dots, X^{n-1}, X^n)$. Quindi $\dim \mathbb{R}_n[X] = n + 1$.

4) Se si considera lo spazio vettoriale $\mathbb{R}[X]$ di tutti i polinomi, senza limitazione sul grado, è ovvio che non può ammettere una base finita. Quindi questo è il primo esempio di spazio vettoriale di dimensione infinita. Nel seguito ed affrontando ulteriori studi ne incontreremo molti altri. Soprattutto se si pensa a spazi di funzioni.
Si può anche dimostrare in modo rigoroso che lo **spazio vettoriale delle successioni** ad elementi reali non ha una base finita, quindi la sua dimensione è infinita.

## 2.2 Sottospazi di uno spazio vettoriale

**Definizione 2.2.1.** Un sottoinsieme $W$ di un $K$- spazio vettoriale $V$ si dice un **sottospazio** se è non vuoto e risulta uno spazio vettoriale rispetto alle operazioni di somma e prodotto esterno ereditate da $V$.

Quanto sopra è equivalente a dire che: $0_V \in W$ e che per ogni coppia di elementi $a, b \in K$ e per ogni coppia di elementi di $v_1, v_2 \in V$ si ha $av_1 + bv_2 \in V$.

È ovvio che **l'intersezione** $V_1 \cap V_2$ di due sottospazi di $V$ è ancora un sottospazio. Ciò vale anche estendendo l'intersezione ad una famiglia, finita o infinita, di sottospazi.

**Proposizione 2.2.2.** *L'unione insiemistica di due sottospazi $V_1 \cup V_2$ di $V$ non è un sottospazio a meno che non sia $V_1 \subseteq V_2$ oppure $V_2 \subseteq V_1$.*

*Proof.* Negare $V_1 \subseteq V_2$ e $V_2 \subseteq V_1$ equivale a dire che esistono vettori tali che $v_1 \in V_1, v_1 \notin V_2$ e che $v_2 \in V_2, v_2 \notin V_1$. In tal caso si vede che il vettore $v_1 + v_2 \notin V_1 \cup V_2$. Perché se $v_1 + v_2$ fosse in $V_1$ sarebbe $v_1 + v_2 = u_1$ e quindi $v_2 \in V_1$ contro la nostra ipotesi. Analogamente non può essere $v_1 + v_2 = u_2$ altrimenti $v_1 \in V_2$, contro l'ipotesi. Mentre se $V_1 \subseteq V_2$ l'unione è $V_2$ che è un sottospazio. □

**Definizione 2.2.3.** Dati due sottospazi $U$ e $W$ di un $K$-spazio vettoriale $V$. Si dice che $V'$ è lo spazio generato dai due sottospazi e si indica $V' = U + W$ il sottospazio di $V$ costituito da tutti i vettori del tipo $v = u + w$, con $u \in U$ e $w \in W$.

Tale sottospazio si suole anche chiamare il **sottospazio minimo** di $V$ fra quelli che contengono $U$ e $W$. Tutti i sottospazi $(Z_i)_{i \in I}$ contenenti $U$e $W$ conterranno le somme dei loro elementi e quindi conterranno $U + W$. Quindi sarà $\bigcap_{i \in I} Z_i = U + W$.

La definizione di spazio somma si può estendere anche a un numero finito di sottospazi di $V$. Si dirà **spazio somma** di $n$ sottospazi $U_1 + U_2 + \cdots + U_n$ il sottospazio $U$ di $V$ costituito da tutti gli elementi del tipo $u = u_1 + u_2 + \cdots + u_n$. Potremo anche osservare che $U = U_1 + U_2 + \cdots + U_n$ è il sottospaio minimo di $V$ contenente $U_1, U_2, \ldots, U_n$.

**Theorema 2.2.4 (Formula di Grassmann).** *Siano $U$ e $W$ due sottospazi di un $K$-spazio vettoriale $V$. Allora sussiste la formula*

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

*Proof.* Siano $(e_1, e_2, \ldots, e_r)$ una base del sottospazio $U \cap W$. Come sappiamo possiamo completare tale insieme libero ad una base di $U$ e una di $W$ e siano $(e_1, e_2, \ldots, e_r, u_1, \ldots, u_s)$ e $(e_1, e_2, \ldots, e_r, w_1, \ldots, w_t)$ due basi per $U$ e $W$ rispettivamente. Proviamo che $(e_1, e_2, \ldots, e_r, u_1, \ldots, u_s, w_1, \ldots, w_t)$ è una base per lo spazio somma $U+W$. Ovviamente sono generatori di $U+W$. Ci resta da verificare che sono linearmente indipendenti. Se c'è una loro combinazione lineare nulla deve accadere che tutti i coefficienti siano nulli. Sia

$$a_1e_1 + a_2e_2 + \cdots + a_re_r + b_1u_1 + \cdots + b_su_s + c_1w_1 + \ldots c_tw_t = 0 \tag{2.1}$$

Scriviamo anche

$$-(a_1e_1 + a_2e_2 + \cdots + a_re_r + b_1u_1 + \cdots + b_su_s) = c_1w_1 + \ldots c_tw_t \tag{2.2}$$

Ma allora $c_1w_1 + c_2w_2 + \cdots + c_tw_t \in U \cap W$ e si può scrivere

$$c_1w_1 + c_2w_2 + \cdots + c_tw_t = a'_1e_1 + \cdots + a'_re_r \tag{2.3}$$

che si può anche scrivere

$$c_1w_1 + c_2w_2 + \cdots + c_tw_t - a'_1e_1 - \cdots - a'_re_r = 0, \tag{2.4}$$

che è una relazione lineare nulla fra vettori di una base ed allora tutti i coefficienti sono nulli. Dalla (1) segue anche che tutti i $b_i$ sono nulli.
Valutiamo la dim dello spazio somma $U+W$.

$$\dim(U+W) = r+s+t = r+s+t+r-r = \dim U + \dim W - \dim(U \cap W)$$

□

Si pone la seguente

**Definizione 2.2.5.** Diremo che un sottospazio $V' \subseteq V$ è **somma diretta** dei sottospazi $U$ e $W$ e si scriveremo $V' = U \oplus W$ se ogni vettore $v \in V'$ si può scrivere in **modo unico** $v = u + w$, con $u \in U$ e $w \in W$.

**Proposizione 2.2.6.** *È facile provare che $V' = U \oplus W$ se e solo che $V' = U+W$ e $U \cap W = (0)$.*

*Proof.* Lasciata al lettore per esercizio. □

**Osservazione** 2.2.7. Dalla formula di Grassmann Teorema 2.2.4 e dalla precedente proposizione si deduce che se $V' = U + W$ e l'intersezione $U \cap W = (0)$ allora segue che

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

.

**Definizione 2.2.8.** Sia $V$ un $K$-spazio vettoriale e $U_1, U_2, \ldots, U_n$ sottospazi tali che $V' = U_1 + U_2 + \cdots + U_n$. Se ogni vettore $v \in V'$ si può scrivere **in modo unico** $v = u_1 + u_2 + \cdots + u_n$, con $u_i \in U_i$, allora si dice che $V'$ è somma diretta dei sottospazi e si scrive $V' = U_1 \oplus U_2 \cdots \oplus U_n$. In tal caso si ha

$$\dim V' = \dim U_1 + \dim U_2 + \cdots + \dim U_n$$

**Proposizione 2.2.9.** *Se* $V' = U_1 + U_2 + \cdots + U_n$ *condizione necessaria e sufficiente perché tale somma sia diretta è che si verifichino le seguenti condizioni:*

$$\forall i \in \{1, 2, \ldots, n\} \qquad sia \qquad U_i \cap (U_1 + U_2 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n) = (0)$$

*Proof.* Se un vettore non nullo $w_i \in U_i \cap (U_1 + U_2 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_n)$ allora $w_i = u_1 + u_2 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_n$ da cui segue $u_1 + u_2 + \cdots + u_{i-1} - w_i + u_{i+1} + \cdots + u_n = 0_V = 0_V + 0_V + \cdots + 0_V$ che contraddice l'ipotesi che ogni vettore si scrive in modo unico.
Viceversa se valgono le condizione ogni vettore $v \in V'$ si scrive come somma in modo unico. Se $v = u_1 + u_2 + \cdots + u_n = v_1 + v_2 + \cdots + v_n$ allora segue, per esempio, che $(u_1 - v_1) = (v_2 - u_2) + \cdots + (v_n - u_n)$. Ciò dice che $u_1 - v_1 \in U_1 \cap (U_2 + U_3 = \cdots + U_n)$, contro l'ipotesi. Ovviamente almeno un indice $i$, per cui accade quanto sopra, dovrà esistere. □

Un altro importante risultato è il seguente

**Proposizione 2.2.10.** *Supponiamo che* $V$ *è somma dei sottospazi* $U_1, U_2, \ldots, U_n$ *e che* $\dim V = \dim U_1 + \dim U_2 + \cdots + \dim U_n$. *Allora segue che* $V$ *è somma diretta di* $U_1, U_2, \ldots, U_n$ *.*

*Proof.* Prendiamo una base per $U_1$, una per $U_2$ e così via e una per $U_n$. Tenuto conto della ipotesi segue che l'insieme costituito prendendo queste basi è un insieme di generatori di $V$. Ma tali generatori, per il fatto che
$\dim V = \dim U_1 + \dim U_2 + \cdots + \dim U_n$, sono indipendenti.
Se $0_V = u_1 + u_2 + \cdots + u_n$. Esprimiamo ogni $u_i$ come combinazione lineare della base suddetta di $U_i$. In tale combinazione lineare tutti i coefficienti devono essere nulli; quindi tutti gli $u_i$ saranno nulli. Quindi c'è un solo modo di scrivere lo $0_V$ come combinazione lineare di elementi degli $U_i$. Ed allora la somma è diretta. □

Vorrei porre qualche domanda

**Domanda 1.** *Nello spazio vettoriale dei vettori* $V(O)$ *applicati nella origine* $O$ *i sottospazi, oltre ai sottospazi banali* $(O)$ *e tutto* $V(O)$*, sono le rette per l'origine* $(O)$ *e i piani passanti per* $(O)$*.*

*1) Su vari esempi verificare la Formula di Grassmann.*

*2) Esaminare le varie possibilità che possono presentarsi nel confrontare le dimensioni di tre sottospazi.*

*3) È vero che sussiste la formula, per tre sottospazi* $V_1, V_2, V_3$ *di* $V(O)$*:*

$$\dim(V_1 + V_2 + V_3) = \dim V_1 + \dim V_2 + \dim V_3$$
$$- \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) + \dim(V_1 \cap V_2 \cap V_3)?$$

*4) portare esempi di sottospazi che sono somme dirette ed altri che non lo sono.*

**Domanda 2.** *Siano $V$ un $K$-spazio vettoriale e $v_1, v_2, v_3 \in V$. Dimostrare che $(v_1, v_2, v_3)$ è una base di $V$ se e solo se $(v_1 - v_2, v_2 - v_3, v_3)$ lo è.*
*Generalizzare al caso di $m$ vettori.*

## 2.3 Applicazioni Lineari

In questa sezione studiamo le applicazioni lineari fra spazi vettoriali.

**Definizione 2.3.1.** Una applicazione $f : V \to W$ fra $K$-spazi vettoriali si dice che è **lineare** se verifica le seguenti condizioni:

1) $f(v_1 + v_2) = f(v_1) + f(v_2), \quad \forall v_1, v_2 \in V.$

2) $\forall a \in K, \forall v \in V$accade che $f(av) = af(v)$.

Se una applicazione $f : V \to W$ è lineare allora
$$f(a_1v_1 + a_2v_2) = a_1f(v_1) + a_2f(v_2)$$

Per esempio l'applicazione $f : \mathbb{R}^3 \to \mathbb{R}^3$ definita dalla legge: $f(x, y, z) = (x+y+z, x-y, 2x-4z)$ è ovviamente lineare, mentre l'applicazione $g(x, y, z) = (1 - x, x + y - z, 2x - z)$ non lo è. Osserviamo che le applicazioni lineari sono definite da **equazioni lineari e omogenee.**

Consideriamo l'applicazione $d : \mathbb{R}_3[X] \to \mathbb{R}_2[X]$ derivazione, rispetto ad $X$, dallo spazio vettoriale dei polinomi in $X$ di grado $\leq 3$ allo spazio vettoriale dei polinomi in $X$ di grado $\leq 2$.
Per le ben note regole di derivazione si ha facilmente che tale applicazione è lineare.
Per quanto riguarda l'integrazione negli spazi vettoriali dei polinomi di un certo grado deve necessariamente accadere che la dimensione del dominio dev'essere almeno un grado minore del grado del codominio.
Se consideriamo l'applicazione integrazione $\int : \mathbb{R}[X] \to \mathbb{R}[X]$ nello spazio vettoriale dei polinomi a coefficienti in $\mathbb{R}$, senza condizioni sui gradi dei polinomi, si ha che anche l'integrazione è sempre una applicazione lineare.

Data una applicazione lineare fra $K$-spazi vettoriali $f : V \to W$. Si definisce **nucleo** della nostra applicazione, e si indica con $\boxed{Kerf}$, il sottospazio di $V = \{v \in V | f(v) = 0_W\}$.
Si dice **immagine** dell'applicazione e si indica con $\boxed{Imf}$ il sottospazio di $W = \quad \{w \in W| \ \ \exists v \in V \ \ \text{per cui} \ \ f(v) = w\}$.

Per **studiare** una applicazione lineare $f : V \to W$ bisogna determinare il nucleo e l'immagine dell'applicazione.

Ricordiamo che una applicazione lineare è **iniettiva** se elementi distinti vengono trasformati in elementi distinti. In altre parole se da $f(v_1) = f(v_2) \Rightarrow v_1 =$

$v_2$.

Una applicazione lineare è **suriettiva** se l'immagine coincide col codominio. Con le notazioni usate deve accadere che $Imf = W$.

**Theorema 2.3.2.** *Una applicazione lineare $f : V \to W$ è iniettiva se e solo se $Kerf = \{0_V\}$.*

*Proof.* Supponiamo che l'applicazione sia iniettiva. Sia $v \neq 0 \in Kerf$. Sappiamo che $f(0) = 0$, ma anche che $f(v) = 0$. Ciò contro l'ipotesi che $f$ è iniettiva. Viceversa supponiamo che $Kerf$ contenga solo lo zero di $V$ e sia $f(v_1) = f(v_2)$. Da ciò segue $f(v_1) - f(v_2) = 0$ e per linearità $f(v_1 - v_2) = 0$. quindi $v_1 = v_2$ e l'applicazione è iniettiva. □

Per studiare una applicazione lineare è molto importante il seguente risultato noto come teorema del confronto delle dimensioni.

**Theorema 2.3.3 (Confronto delle dimensioni).** *Sia $f : V \to W$ una applicazione lineare fra $K$-spazi vettoriali. Vale la seguente relazione*

$$\dim V = \dim Kerf + \dim Imf$$

*Proof.* Sia $(v_1, v_2, \ldots, v_r)$ una base di $kerf$. Possiamo completarla ad una base di $V$ aggiungendo i vettori $(w_1, w_2, \ldots, w_s)$. Il sottospazio $Imf$ è generato dalle immagini dei vettori di una base di $V$. E poiché i $v_i \in Kerf$, possiamo dire che $Imf$ è generata da $\{f(w_1), f(w_2), \ldots, f(w_s)\}$. Si prova facilmente che i vettori $f(w_1), f(w_2), \ldots, f(w_s)$ sono linearmemnte indipendenti e quindi costituiscono una base dell' immagine $Imf$.

Supponiamo che $b_1 f(w_1) + b_2 f(w_2) + \cdots + b_s f(w_s) = 0$. proviamo che i $b_i$ sono tutti nulli. Dalla precedente segue per linearità $f(b_1 w_1 + b_2 w_2 + \cdots + b_s w_s) = 0$. Allora $b_1 w_1 + b_2 w_2 + \cdots + b_s w_s \in Kerf$. Si deduce che tale elemento sta nel nucleo. Possiamo scriverlo in termini della base di $Kerf$. Allora abbiamo

$$b_1 w_1 + b_2 w_2 + \cdots + b_s w_s = a_1 v_1 + a_2 v_2 + \cdots + a_r v_r$$

da cui si ha una relazione lineare nulla di vettori linearmente indipendenti

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r - b_1 w_1 - b_2 w_2 - \cdots - b_s w_s = 0$$

equindi tutti i coefficienti sono nulli, in particolare i coefficienti $b_i$. Che è quanto volevamo dimostrare.

Ne segue

$$\dim V = r + s = \dim Kerf + \dim Imf$$

□

Adesso facciamo alcune applicazioni.

**Corollario 2.3.4.** *Se $V$ e $W$ sono due $K$-spazi vettoriali tali che* $\dim V > \dim W$*, non può esistere una applicazione lineare $f : V \to W$ che sia iniettiva.*

*Proof.*
$$\dim Kerf = \dim V - \dim Imf \geq \dim V - \dim W > 0$$
. Quindi l'applicazione non può essere iniettiva. □

**Corollario 2.3.5.** *Se $V$ e $W$ sono due $K$-spazi vettoriali tali che* $\dim V < \dim W$*, non può esistere una applicazione lineare $f : V \to W$ che sia suriettiva.*

*Proof.*
$$\dim Imf = \dim V - \dim Kerf \leq \dim V < \dim W.$$
□

*Esempio* 2.3.6. Ogni applicazione lineare $f : V \to V$da un $K$-spazio vettoriale $V$ di dimensione 1 in se è data dalla moltiplicazione per uno scalare.

*Proof.* Se $f$ è l'applicazione nulla il risultato è ovvio. Altrimenti per ogni $v \in V$ si ha $f(v) = vf(1) = h \cdot v$. □

*Esempio* 2.3.7. Sia $V$ un $K$-spazio vettoriale. Condizione necessaria e sufficiente perché esista un endomorfismo $f : V \to V$ tale che $\dim Kerf = \dim Imf$ è che $\dim V$ sia un numero pari.

*Proof.* Se un endomorfismo è tale che $\dim Kerf = \dim Imf$, dal teorema del confronto delle dimensioni si ha $\dim V = 2 \dim Kerf$.
Viceversa se la $\dim V = 2h$ è pari, detta $(v_1, \ldots, v_h, w_1, \ldots, w_h)$ una base di $V$, costruiamo un endomorfismo $f$ nel seguente modo: $(v_1, \ldots, v_h) \in Kerf$ e assegniamo $f(w_1) = v_1, f(w_2) = v_2, \ldots, f(w_r) = v_r$. Con tale scelta $f$ risponde a quanto voluto.
□

*Esempio* 2.3.8. Sia $V$ un $K$-spazio vettoriale di dimensione $n$ e $f, g : V \to K$ due applicazioni linearinon nulle tali che $Kerf = Kerg$. Provare che esiste un $h \in K$ per cui si ha: $f(v) = h \cdot g(v), \forall v \in V$.

*Proof.* Visto che $\dim Imf = 1$ (ed anche $\dim Img = 1$), dal teorema del confronto della dimensioni segue cha $\dim V = \dim Kerf + 1$. Detta $v_1, v_2, \ldots, v_{n-1}$ una base del $kerf$, completiamola ad una base di tutto $V$, aggiungendo un vettore $v_n \notin Kerf$. Intanto ogni vettore di $v \in V$ si può scrvere $v = a_1v_1 + a_2v_2 + \cdots + a_{n-1}v_{n-1} + a_nv_n$. Calcoliamo $f(v)$ e $g(v)$. Tenuto conto delle ipotesi si ha $f(v) = a_nf(v_n) = a_nh_1 \neq 0$. Analogamnete $g(v) = a_ng(v_n) = a_nh_2 \neq 0$. Ma $h_1$ e $h_2$ sono elementi non nulli di un campo $K$ e quindi $h_1 = h_2h$. Quindi $f(v) = a_nh_2h = h \cdot g(v)$. □

# Chapter 3

# Matrici e determinanti

## 3.1 Matrici

Una matrice $A = (a_{i,j}) \in K^{m,n}$ è una tabella di $m$ righe ed $n$ colonne dove $a_{i,j}$ è l'elemento che sta scritto nella $i$-ma riga e $j$-ma colonna. Le matrici del tipo $K^{n,n}$, che hanno lo stesso numero di righe e colonne, si dicono **matrici quadrate**.

Se si definisce in $K^{m,n}$ la somma di matrici elemento per elemento e il prodotto $aA$ la matrice ottenuta moltiplicando per $a$ tutti gli elementi di $A$, si ottiene uno spazio vettoriale. È facile fare una verifica.

Sotto particolari condizioni si può effettuare il prodotto di matrici, **righe per colonne**. Data la matrice $A = (a_{i,j}) \in K^{m,p}$ e la matrice $B = (b_{i,j}) \in K^{p,n}$ si definisce prodotto delle due matrici la matrice $C = (c_{i,j}) \in K^{m,n}$ il cui elemento $c_{i,j}$ è dato da

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,p}b_{p,j}.$$

Bisogna osservare subito che perchè tale prodotto sia possibile deve accadere che il numero di colonne della prima matrice deve essere uguale al numero delle righe della seconda matrice.

Ovviamente se trattiamo di matrici quadrate in $K^{n,n}$ il prodotto è sempre possibile. Il prodotto fra matrici gode delle proprietà associativa e distributiva. La matrice identica $I_n$ di ordine $n$ che ha tutti 1 sulla diagonale principale e ovunque 0 è elemento neutro nella moltiplicazione fra matrici.

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Il prodotto fra matrici in generale non è commutativo. Come vedremo in seguito alcune matrici quadrate sono invertibili, nel senso che se $A$ è invertibile esiste una matrice $A^{-1}$ tale che $AA^{-1} = I_n$.

*Esempio* 3.1.1. Consideriamo le due matrici $A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & 1 \\ -1 & -1 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & -1 \end{pmatrix}$. La matrice prodotto $C \in \mathbb{R}^{3,2}$ ed è uguale a $C = AB = \begin{pmatrix} 3 & 3 \\ 3 & 1 \\ -2 & 0 \end{pmatrix}$.

Il prodotto $BA$ non si può effettuare per questioni di dimensione. Il numero delle colonne di $B$ non è uguale al numero delle righe di $A$. Ma anche quando il prodotto si può effettuare si vede che in generale non è commutativo. Provarlo sulle matrici $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} 2 & 0 \\ 1 & -2 \end{pmatrix}$.

Osserviamo che questa è la prima volta che in una struttura moltiplicativa non vale la proprietà commutativa. Ed anche che possono esistere matrici $A$ e $B$ non nulle il cui prodotto è nullo. Si chiamano divisori dello zero.

### 3.1.1 Applicazioni lineari e matrici

In questa sezione mettiamo in relazione le applicazioni lineari $L(V,W)$ con le matrici $K^{m,n}$. Sia $E = (e_1, e_2, \cdots, e_n)$ una base di $V$ e $F = (f_1, f_2, \cdots, f_m)$ una base di $W$.
Sia $f : V \to W$ una applicazione lineare. Si definisce matrice associata a questa applicazione, relativamente alla base $E$ in $V$ e alla base $F$ in $W$ la matrice $M^{E,F}(f)$ che ha come colonne ordinatamente le componenti di

$$f(e_1), f(e_2), \cdots, f(e_n)$$

relativamente alla base $F$ di $W$. Ricordando che le componenti di $v$ rispetto alla base $E$ si possono scrivere $[v]_E = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ allora la matrice $M(f)$ associata ad $f$ si può scrivere

$$M^{E,F}(f) = \begin{pmatrix} [f(e_1]_F & [f(e_2)]_F \cdots [f(e_n)]_F \end{pmatrix}$$

. Una osservazione importante. Se consideriamo in $\mathbb{R}^n$ il vettore $(x_1, x_2, \cdots, x_n)$ le sue componenti rispetto alla base canonica sono esattamente $(x_1, x_2, \cdots, x_n)$. Infatti il vettore $(x_1, x_2, \cdots, x_n)$ si potrà scrvere

$$(x_1, x_2, \cdots, x_n) = x_1(1, 0, \cdots, 0) + x_2(0, 1, \cdots, 0) + \cdots + x_n(0, 0, \cdots, 1).$$

Se per esempio abbiamo $f : \mathbb{R}^3 \to \mathbb{R}^3$ data da$f(x, y, z) = (x-2y, x+y+2z, -y+z)$ allora la matrice $M(f)$, rispetto alle basi canoniche, è la matrice che ha come colonne $f(e_1) = f(1, 0, 0) = (1, 1, 0), f(e_2) = f(0, 1, 0) = (-2, 1, -1), f(e_3) =$

$f(0,0,1) = (0,2,1)$. Quindi abbiamo la matrice $M(f) = \begin{pmatrix} 1 & -2 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix}$.

Adesso partiamo da una matrice $A \in K^{m,n}$ e supponiamo come prima che $E$ sia una base di $V$ e che $F$ sia una base di $W$. Si può costruire una applicazione lineare $\varphi_A^{E,F} : V \to W$ nel seguente modo: detto $v$ un qualunque vettore in $V$, esso ha componenti $[v]_E = \underline{\mathrm{X}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, facciamo il prodotto

$$A\underline{\mathrm{X}} = \begin{pmatrix} y_1 = a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ y_2 = a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ y_m = a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{pmatrix}.$$

Infine il vettore $\varphi_A^{E,F}(v) = w = y_1 f_1 + y_2 f_2 + \cdots + y_m f_m$.

Vale la seguente **importante propietà**. Sempre con le solite notazioni si può provare che:

$$Se \quad f : V \to W \quad e \quad A = M(f) \quad allora \qquad \varphi_A = f$$

$$Se \quad A \in K^{m,n} \qquad e \quad \varphi_A : V \to W, \qquad allora \quad M(\varphi_A) = A.$$

Questo significa che gli operatori che fanno passare dalle applicazioni lineari alle matrici e dalle matrici alle applicazioni lineari sono l'uno inverso dell'altro. In altre parole questi due spazi sono in corrispondenza biunivoca. Fra l'altro vengono "conservate" le operazioni di spazio vettoriale e quindi $L(V,W)$ e $M \in K^{m,n}$ sono strutture algebriche isomorfe e quindi mantengono le proprietà di struttura, tipo la dimensione e così via.

Ricordiamo che si definisce la **composizione delle applicazioni lineari** nel seguente modo.

Siano date le applicazioni lineari $\varphi : V \to W$ e $\psi : W \to Z$. Allora la composizione delle due applicazioni lineari, e si indica $\psi \circ \varphi$, l'applicazione definita nel seguente modo $(\psi \circ \varphi)(v) = \psi(\varphi(v))$, in altre parole l'applicazione successiva delle due applicazioni.Un importante risultato è il seguente

**Theorema 3.1.2.** *Siano date le applicazioni lineari $\varphi : V \to W$ e $\psi : W \to Z$ e la composizione $(\psi \circ \varphi) : V \to Z$. Siano $E = (e_1, e_2, \cdots, e_n), F = (f_1, f_2, \cdots, f_m), G = (g_1, g_2, \cdots, g_p)$ basi fissate rispettivamente in $V, W, Z$. Allora si ha $M^{E,G}(\psi \circ \varphi) = M^{F,G}(\psi) \cdot M^{E,F}(\varphi)$. In parole la matrice associata alla composizione è il prodotto delle matrici purché siano riferite alle basi che a loro competono.*

*Proof.* Per fissare le idee siano $A = (a_{ij}) = M^{E,F}(\varphi) \in K^{m,n}$, $B = (b_{ij}) = M^{F,G}(\psi) \in K^{p,m}$. Sia $C \in K^{p,n}$ La matrice $M^{E,G}(\psi \circ \varphi)$. In base alla definizione di matrice associata e per linearità possiamo scrivere:

$$\begin{aligned}
(\psi \circ \varphi)(e_1) &= \psi(\varphi(e_1)) \\
&= \psi(a_{11}f_1 + a_{21}f_2 + \cdots + a_{m1}f_m) \\
&= a_{11}\psi(f_1) + a_{21}\psi(f_2) + \cdots + a_{m1}\psi(f_m) \\
&= a_{11}(b_{11}g_1 + b_{21}g_2 + \cdots + b_{p1}g_p) + \\
&\quad\ a_{21}(b_{12}g_1 + b_{22}g_2 + \cdots + b_{p2}g_p) + \cdots + \\
&\quad\ a_{m1}(b_{1m}g_1 + b_{2m}g_2 + \cdots + b_{pm}g_p) \\
&= (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1m}a_{m1})g_1 + \cdots
\end{aligned}$$

Da quanto scritto sopra è facile vedere che l'elemento $c_{11}$ si ottiene moltiplicando (riga per colonna) la prima riga di $B$ per la prima colonna di $A$. Calcolando ulteriormente ci si rende conto che la matrice $C = B \cdot A$. ☐

### 3.1.2 Il rango di una matrice

Consideriamo una matrice $A \in K^{m,n}$. Indichiamo con $R_1, R_2, \cdots, R_m$ le righe di $A$ e con $C_1, C_2, \cdots, C_n$ le colonne di $A$. Ogni riga $R_i$ è una $n$-upla e ogni colonna $C_j$ è una $m$-upla. Indichiamo con $R$ lo spazio generato dalle righe $R = \mathcal{L}(R_1, R_2, \cdots, R_m)$ e con $C = \mathcal{L}(C_1, C_2, \cdots, C_n)$. Sussiste l'importante risultato

**Proposizione 3.1.3.** *In una stessa matrice $A$ la* $\dim R = \dim C$.

*Proof.* Siano $r = \dim R$ e $s = \dim C$. Esisteranno $r$ vettori

$$\begin{aligned}
v_1 &= (c_{11}, c_{12}, \cdots, c_{1n}) \\
v_2 &= (c_{21}, c_{22}, \cdots, c_{2,n} \\
&\vdots \\
v_r &= (c_{r1}, c_{r2}, \cdots, c_{r,n})
\end{aligned}$$

che costituiscono una base per lo spazio delle righe $R$. Possiamo allora scrivere

$$\begin{aligned}
R_1 &= d_{11}v_1 + d_{12}v_2 + \cdots + d_{1r}v_r \\
R_2 &= d_{21}v_1 + d_{22}v_2 + \cdots + d_{2r}v_r \\
&\vdots \\
R_m &= d_{m1}v_1 + d_{m2}v_2 + \cdots + d_{mr}v_r
\end{aligned}$$

Sostituendo i valori delle $v_i$ si ottiene

$$
\begin{aligned}
R_1 &= d_{11}(c_{11}, c_{12}, \cdots, c_{1n}) + d_{12}(c_{21}, c_{22}, \cdots, c_{2n}) + \cdots + d_{1r}(c_{r1}, c_{r2}, \cdots, c_{rn}) \\
R_2 &= d_{21}(c_{11}, c_{12}, \cdots, c_{1n}) + d_{22}(c_{21}, c_{22}, \cdots, c_{2n}) + \cdots + d_{2r}(c_{r1}, c_{r2}, \cdots, c_{rn}) \\
\vdots & \\
R_m &= d_{m1}(c_{11}, c_{12}, \cdots, c_{1n}) + d_{m2}(c_{21}, c_{22}, \cdots, c_{2n}) + \cdots + d_{mr}(c_{r1}, c_{r2}, \cdots, c_{rn})
\end{aligned}
$$

□

È facile vedere che la prima colonna $C_1$ della matrice $A$ si può scrivere

$$
\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = c_{11} \begin{pmatrix} d_{11} \\ d_{21} \\ \vdots \\ d_{m1} \end{pmatrix} + c_{21} \begin{pmatrix} d_{12} \\ d_{22} \\ \vdots \\ d_{m2} \end{pmatrix} + \cdots + c_{r1} \begin{pmatrix} d_{1r} \\ d_{2r} \\ \vdots \\ d_{mr} \end{pmatrix}
$$

Ed ancora la seconda colonna $C_2$ della matrice $A$ si può scrivere

$$
\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} = c_{12} \begin{pmatrix} d_{11} \\ d_{21} \\ \vdots \\ d_{m1} \end{pmatrix} + c_{22} \begin{pmatrix} d_{12} \\ d_{22} \\ \vdots \\ d_{m2} \end{pmatrix} + \cdots + c_{r2} \begin{pmatrix} d_{1r} \\ d_{2r} \\ \vdots \\ d_{mr} \end{pmatrix}
$$

E così via. In sostanza le colonne $C_i$ della matrice $A$ sono combinazioni lineari delle $r$ colonne $D_j$ della matrice

$$
D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1r} \\ d_{21} & d_{22} & \cdots & d_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ d_{m1} & d_{m2} & \cdots & d_{mr} \end{pmatrix}.
$$

Quindi $\mathcal{L}(C_1, C_2, \cdots, C_n) \subseteq \mathcal{L}(D_1, D_2, \cdots, D_r)$. Ne segue $\dim C = s \leq r = \dim R$. Tale proprietà vale per ogni matrice. Se adesso applichiamo il precedente risultato alla matrice ${}^tA$ segue $r \leq s$. In conclusione $r = s$.

Poniamo adesso l'importante definizione

**Definizione 3.1.4.** Il **rango** di una matrice $A \in K^{m,n}$ è la dimensione dello spazio delle righe, o dello spazio delle colonne, visto che i due numeri coincidono.

In seguito vedremo che il rango si può calcolare in svariati modi.

**Definizione 3.1.5.** Una matrice $A \in K^{m,n}$ si dice **ridotta per righe** se in ogni riga non nulla c'è almeno un **elemento speciale**, cioè un elemento non nullo al di sotto del quale sulla stessa colonna ci sono tutti zeri.

Facciamo qualche esempio.

*Esempio* 3.1.6.

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

La matrice $A$ non è ridotta per righe, mentre la $B$ lo è. Infatti $b_{11} = 1$ è un elemento speciale sulla prima riga. La seconda riga è nulla. Sulla terza riga $b_{32} = 2$ è speciale. Infine $b_{43} = 2$ è speciale.

È facile provare che **in una matrice ridotta per righe le righe non nulle sono linearmente indipendenti** e quindi costituiscono una base per lo spazio delle righe. Quindi in una matrice ridotta per righe **il numero delle righe non nulle da' il rango della matrice**.

Sia data una qualunque matrice $A \in K^{m,n}$. C'è un procedimento che permette di **ridurre** la matrice, ottenendo una matrice $A'$ che sia **equivalente** ad $A$, cioè che sia ridotta per righe ma che abbia lo stesso spazio $R$ generato dalle righe.
È evidente che lo scambio di righe non altera lo spazio $R$.
Lo spazio $R$ non si altera se ad una riga $R_i$ si aggiunge una combinazione lineare delle rimanenti righe. Dimostriamo il seguente risultato

**Proposizione 3.1.7.** *Sia $R = \mathcal{L}(R_1, R_2, \cdots, R_m)$. Consideriamo $R' = \mathcal{L}(R_1, R_2 + kR_1, \cdots, R_m)$. Proviamo che $R = R'$.*

*Proof.* Sia $a_1R_1 + a_2(R_2 + kR_1) + \cdots + a_mR_m$ un qualunque elemento in $R'$. Questo si piò scrivere $(a_1 + a_2k)R_1 + a_2R_2 + \cdots + a_mR_m$. Da ciò segue che $R' \subseteq R$. Viceversa prendiamo un qualunque elemento in $R$. E sia $a_1R_1 + a_2R_2 + \cdots + a_mR_m$. Aggiungiamo e togliamo $a_2kR_1$; si ottiene $(a_1 - a_2k)R_1 + a_2(R_2 + kR_1) + \cdots + a_mR_m$. da ciò segue $R \subseteq R'$. Quindi $R = R'$. $\square$

Ovviamente reiterando tale risultato si ha quanto affermato e cioè che **addizionando ad una riga di una matrice una combinazione lineare delle rimanenti righe lo spazio generato dalle righe non si altera**.
Di fatto per calcolare il rango di una matrice $A$ si riduce la matrice per righe e si contano le righe non nulle.
Facciamo qualche esempio.
Consideriamo la matrice

$$A = \begin{pmatrix} h & -1 & 0 \\ -1 & h & 0 \\ 5-2h & 1 & 2 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 5-2h & 1 & 2 \\ -1 & h & 0 \\ h & -1 & 0 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_3 + hR_2}$$

$\begin{pmatrix} 5-2h & 1 & 2 \\ -1 & h & 0 \\ 0 & h^2-1 & 0 \end{pmatrix}$. La matrice $A$ ha rango 3 per $h \neq \pm 1$. Mentre per $h = \pm 1$ ha rango 2. Ricordiamo che gli elementi speciali devono essere diversi

da 0 e che gli elementi dell'ultima riga diversi da 0 si considerano speciali.
Calcolare il rango della seguente matrice, al variare del parametro $h$.

$$A = \begin{pmatrix} 1 & -h & 0 & 1 \\ -h & 1 & 1 & 0 \\ 0 & 1-h & 1 & 0 \\ h+1 & -2h & 0 & 1 \end{pmatrix} \xrightarrow[R_3 \leftrightarrow R_3 - R_2]{R_4 \leftrightarrow R_4 - R_1} \begin{pmatrix} 1 & -h & 0 & 1 \\ -h & 1 & 1 & 0 \\ h & -h & 0 & 0 \\ h & -h & 0 & 0 \end{pmatrix}$$

Si vede subito che per $h = 0$ la matrice ha rango 2, mentre per $h \neq 0$ il rango è 3.

## 3.2 Sistemi di equazioni lineari

Supponiamo come al solito che $K$ indichi un campo. In generale sarà il campo dei numeri reali $\mathbb{R}$ o dei numeri complessi $\mathbb{C}$. Una **equazione di primo grado (lineare)** nelle variabili $x_1, x_2, \cdots, x_n$ è una equazione del tipo $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b_1$, con coefficienti $a_i \in K$. Una **soluzione** di tale equazione sarà una $n$-upla $\underline{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in K^n$ tale che $a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = b_1$. Se poi si hanno $m$ equazioni di questo tipo si ha un **sistema di equazioni lineari**. Una soluzione del sistema sarà una soluzione di tutte le equazioni del sistema.

Introduciamo qualche notazione. Consideriamo il sistema

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

Vengono evidenziati la matrice dei coefficienti $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, il vettore colonna $\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$ e il vettore colonna $B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}$ dei termini noti.
In modo compatto il precedente sistema (1) si può anche scrivere $A\underline{X} = B$.

Se si vuole il sistema (1) si può anche scrivere

$$(2) \qquad \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}.$$

È chiaro da ciò che **il sistema ammette soluzioni** se e solo se la colonna $B$ è combinazione lineare delle colonne di $A$.

Quindi sussiste il teorema

**Theorema 3.2.1** (Rouchè-Capelli)**.** *Condizione necessaria e sufficiente perchè il sistema* (1) *abbia soluzioni è che il rango $r(A)$ di $A$ sia uguale al rango $r(A|B)$ della matrice completa.*

Supponiamo che il rango $r(A) = r(A|B) = p \leq n$. Ci saranno $p$ colonne linearmente indipendenti di $A$, per semplicità supponiamo che siano le prime $p$ colonne. Il sistema può essere riscritto

$$(2) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = b_1 - a_{1,p+1}x_{p+1} - \cdots - a_{1,n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = b_2 - a_{2,p+1}x_{2,p+1} - \cdots - a_{2,n}x_n \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mp}x_n = b_m - a_{m,p+1}x_{p+1} - \cdots - a_{m,n}x_n \end{cases}$$

Se si assegnano valori arbitrari alle incognite $x_{p+1}, x_{p+2}, \cdots, x_n$, a secondo membro si ottiene un elemento dello spazio delle colonne della matrice $(A|B)$. Ma, per le ipotesi fatte le prime $p$ colonne di $A$ sono indipendenti e quindi costituiscono una base per lo spazio delle colonne di $(A|B)$. Segue quindi che per ogni scelta delle incognite a secondo membro si determinano univocamente i valori delle incognite $x_1, x_2, \cdots, x_p$. Ciò accade per ogni scelta delle incognite a secondo membro, che vengono dette **incognite libere**. Quando accade questo si suol dire che il sistema ammette $\infty^{n-p}$ soluzioni. Quando $n = p$ il sistema ammette una sola soluzione.

Nella pratica si manipola il sistema $A\underline{X} = B$ fino ad ottenere un sistema ad esso equivalente $A'\underline{X}' = B'$, che abbia la matrice $A'$ ridotta per righe. Si può provare rigorosamente, ma noi lo omettiamo, che la procedura di riduzione che fa passare da $(A|B)$ a $(A'|B')$ è tale che il sistema $A'\underline{X}' = B'$ è equivalente al sistema dato, nel senso che l'uno ha tutte e sole le soluzioni dell'altro.
È importante notare che per un sistema ridotto per righe il Teorema di Rouchè Capelli si può riformulare nel seguente modo.

**Theorema 3.2.2** (Rouchè-Capelli)**.** *Condizione necessaria e sufficiente perché un sistema ridotto sia risolubile è che in corrispondenza ad ogni riga di zeri ci sia come termine noto uno zero.*

**Proposizione 3.2.3.** *I sistemi che hanno tutti i termini noti nulli si dicono* **omogenei**. *Per il Teorema di Rouchè-Capelli tali sistemi sono sempre risolubili.*

*Esempio* 3.2.4. Discutere al variare del parametro $h$ il seguente sistema di equazioni lineari.

$$\begin{cases} (h-1)y = -1 \\ -x + (h+1)z = 1 \\ hx - 2z = 2 \end{cases}$$

Scriviamo la matrice completa del sistema

$$\begin{pmatrix} 0 & h-1 & 0 & | & -1 \\ -1 & 0 & h+1 & | & 1 \\ h & 0 & -2 & | & 2 \end{pmatrix}$$

Se $h-1=0$ la matrice diventa

$$\begin{pmatrix} 0 & 0 & 0 & | & -1 \\ -1 & 0 & 2 & | & 1 \\ 1 & 0 & -2 & | & 2 \end{pmatrix}$$

e il sistema è incompatibile perché in corrispondenza alla prima riga tutta nulla il termine noto è non nullo. Si prosegue nella riduzione con l'idea di scegliere -1 come elemento speciale sulla seconda riga.

$$\begin{pmatrix} 0 & h-1 & 0 & | & -1 \\ -1 & 0 & h+1 & | & 1 \\ h & 0 & -2 & | & 2 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_3 + hR_2}$$

$$\xrightarrow{R_3 \leftrightarrow R_3 + hR_2} \begin{pmatrix} 0 & h-1 & 0 & | & -1 \\ -1 & 0 & h+1 & | & 1 \\ 0 & 0 & -2+h(h+1) & | & 2+h \end{pmatrix}$$

L'elemento di posto $(3,3)$ è $h^2+h-2$. Questo è zero per $h=1,-2$. Il caso $h=1$ lo abbiamo già esaminato; per $h=-2$ la terza riga della matrice dei coefficienti è nulla, ma anche il termine noto è nullo.Per il teorema di Rouchè Capelli il sistema ha rango due ed ammette $\infty^1$ soluzioni. Per $h=-2$ il sistema diventa $\begin{cases} -3y=-1 \\ -x-z=1 \end{cases}$. Le soluzioni del sistema sono $y=\frac{1}{3}, x=-z-1$, con $z$ variabile libera.

Si possono anche considerare sistemi ad **incognite vettoriali**. Il metodo di riduzione e il Teorema di Rouchè Capelli sussistono anche in questo caso. Una importante applicazione è la determinazione dell'inversa $A^{-1}$ di una matrice invertibile $A$. Cioè una matrice tale che $AA^{-1}=A^{-1}A=I$, la matrice identica. Sussiste la seguente

**Proposizione 3.2.5.** *Condizione necessaria e sufficiente perché la matrice $A \in K^{n,n}$ sia invertibile è che $A$ sia di rango massimo. Se $A^{-1}$ è inversa a destra lo sarà anche a sinistra. Si avrà $AA^{-1}=A^{-1}A=I_n$.*

*Proof.* Consideriamo l'equazione matriciale $A\underline{X}=I_n$. Il rango della matrice completa $(A|I_n)=n$. Perché il sistema ad incognite vettoriali sia possibile deve accadere che $r(A)=n$; anzi la soluzione è unica, diciamola $A^{-1}$. Consideriamo adesso l'equazione $\underline{Y}A=I_n$. Passando alle trasposte si ha ${}^tA{}^t\underline{Y}=I_n$. Il rango di ${}^tA$ è $n$ e quindi $\underline{Y}A=I_n$. Da $\underline{Y}A=I_n$ segue $\underline{Y}A\underline{A}^{-1}=\underline{A}^{-1}$. Ma $A\underline{A}^{-1}=I_n$ e quindi $\underline{Y}=\underline{A}^{-1}$. □

*Esempio* 3.2.6. Calcoliamo l'inversa della matrice quadrata $A=\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 1 & -1 & 3 \end{pmatrix}$.
Dobbiamo trovare una matrice $\underline{X} \in \mathbb{R}^{3,3}$ tale che sia $A\underline{X}=I_3$. La matrice $\underline{X}$

si può scrivere $\underline{X} = \begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \underline{X}_3 \end{pmatrix}$, dove $\underline{X}_i = (x_{i,1}, x_{i,2}, x_{i,3})$ le pensiamo incognite vettoriali. Consideriamo quindi la matrice completa

$$\begin{pmatrix} 1 & 3 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \\ 1 & -1 & 3 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_3 - R_1} \begin{pmatrix} 1 & 3 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \\ 0 & -4 & 1 & | & -1 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_2}$$

$$\xrightarrow{R_3 \leftrightarrow R_2} \begin{pmatrix} 1 & 3 & 2 & | & 1 & 0 & 0 \\ 0 & -4 & 1 & | & -1 & 0 & 1 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \end{pmatrix}$$

La matrice è ridotta e si ha $\underline{X}_2 = (0,1,0)$; $\quad \underline{X}_3 = (-1,0,1) + 4\underline{X}_2 = (-1,4,1)$; $\underline{X}_1 = (1,0,0) - 3\underline{X}_2 - 2(-2,4,-1) = (3,-11,-2)$. In definitiva l'inversa della matrice $A$ è la matrice $A^{-1} = \begin{pmatrix} 3 & -11 & -2 \\ 0 & 1 & 0 \\ -1 & 4 & 1 \end{pmatrix}$.

## 3.3 Matrici e determinanti

Prima di parlare dei determinanti delle matrici quadrate dobbiamo ricordare alcune proprietà delle **permutazioni** su $n$ oggetti, per esempio prendiamo $\{1, 2, \cdots, n\}$. Ricordiamo che una permutazione è un nuovo gruppo degli $n$ numeri riscritto in ordine diverso, senza ripetizioni. Indichiamo con $\mathcal{P}$ la totalità delle permutazione sugli $n$ numeri $\{1, 2, \cdots, n\}$. Dal calcolo combinatorio si sa che ci sono $n! = n(n-1)\cdots 3 \cdot 2 \cdot 1$ permutazioni distinte. Per esempio se si considera $\{1, 2, 3\}$ ci sono 6 permutazioni distinte:

$$(1) \qquad \{1,2,3\}, \{1,3,2\}, \{2,1,3\}, \{2,3,1\}, \{3,1,2\}, \{3,2,1\}.$$

Inoltre le permutazioni vengono distinte in permutazioni di classe pari e classe dispari. Sono di classe pari quelle che hanno un numero totale di inversioni che sia un numero pari e di classe dispari quelle il cui numero totale di inversioni sia un numero dispari. Per esempio la permutazione $\{2, 1, 3\}$ è di classe dispari, perché 2 è in inversione con 1, ma è nell'ordine naturale con 3; 1 non è in inversione con 3. C'è una sola inversione e quindi la permutazione è di classe dispari. Mentre $\{2, 3, 1\}$ è di classe pari in quanto 2 è in inversione con 1, e 3 è in inversione con 1; quindi un numero complessivo pari di inversioni. Le permutazioni su $n$ oggetti si distribuiscono metà di classe pari e metà di classe dispari.

Consideriamo la matrice quadrata $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Vogliamo dare la definizione di determinante $\det(A)$ della matrice $A$. Talvolta si scrive $|A|$ per indicare $\det(A)$. Ebbene $\det(A)$ è la **somma algebrica** di tutti i possibili **prodotti dedotti dalla matrice** $(A)$. Ricordiamo che un prodotto dedotto

dalla matrice è un prodotto di $n$ fattori presi dalla matrice , con la proprietà che i fattori non stiano ne' sulla stessa riga ne' sulla stessa colonna. La proprietà che i fattori dei prodotti dedotti non stiano ne' sulla stessa riga ne' sulla stessa colonna è assicurata se attribuiamo ai primi e ai secondi indici degli elementi della matrice due qualunque permutazioni. Il segno che daremo al prodotto dedotto dipenderà dalla classe delle permutazioni dei primi e dei secondi indici. Due classi della stessa parità danno un segno +, due classi di parità diverse danno un segno -. Prendiamo un prodotto dedotto dalla matrice $3 \times 3$ considerata. Per esmpio: $a_{21}a_{32}a_{13}$. Siccome gli elementi $a_{i,j} \in K$ vale la proprietà commutativa e quindi lo stesso prodotto dedotto si può riscrivere : $a_{13}a_{21}a_{32}$. Tale prodotto dedotto ha lo stesso segno del precedente, e siccome i primi indici si seguono nell'ordine naturale, il segno è determinato dalla classe della permutazione sui secondi indici. In definitiva avremo $3! = 3 \cdot 2 \cdot 1 = 6$ prodotti dedotti distinti.
Consideriamo i prodotti: $a_{1,-}a_{2,-}a_{3,-}$. Sui secondi indici sostituiamo i valori delle 6 permutazioni (1). Si hanno i seguenti prodotti dedotti cui viene attribuito un segno in base alla parità della classe della permutazione sui secondi indici

$$det(A) = a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}$$

Per una matrice quadrata di ordine $n$ il determinante si definisce:

$$det(A) = \sum_{(j_1, j_2, \cdots, j_n) \in \mathcal{P}} (-1)^s a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

dove $s$ è la classe della permutazione dei secondi indici $(j_1, j_2, \cdots, j_n)$ rispetto alla permutazione fondamentale. C'è da dire che ci sono regole e tecniche per il calcolo dei determinanti delle matrici, che ora andremo ad esporre.

*Proprietà* 2. Sia data una matrice $A \in K^{nn}$.

1. Se in $A$ c'è tutta una riga o una colonna formata da zeri allora il determinante di $A$ è zero.

2. Se in una matrice si scambiano comunque due righe o colonne il determinante cambia di segno.

3. Una matrice che abbia due righe o colonne uguali ha determinante zero.

4. Se tutti gli elementi di una riga o colonna vengono moltiplicati per una costante $k$ il determinante della matrice viene moltiplicato per $k$.

5. Se ad una riga di una matrice viene addizionata una combinazione lineare di altre righe il determinante della matrice resta invariato.

*Proof.* La (1) è ovvia in quanto ogni prodotto dedotto contiene uno zero.
La (2) è conseguenza del fatto che dopo lo scambio tutti i prodotti dedotti vengono cambiati di segno.

(3) Se si scambiano due righe o colonne uguali la matrice non cambia, ma per la (2) il determinante cambia segno. Si ha quindi $det(A) = -det(A)$. Quindi $det(A) = 0$.
(4) e (5) sono abbastanza ovvie. □

**Definizione 3.3.1.** Sia $A \in K^{mn}$. Una **sottomatrice** di ordine $p$ è una matrice ottenuta intersecando $p$ righe e $p$ colonne di $A$. Un **minore** di ordine $p$ è il determinante di una sottomatrice di ordine $p$.
In una matrice quadrata si dice **complemento algebrico** dell'elemento $a_{i,j}$ e si indica $A_{i,j} = (-1)^{i+j}$ per il determinante della matrice che si ottiene sopprimendo in $A$ la riga $i$-ma e la colonna $j$-ma.

**Theorema 3.3.2 (Laplace).** *In una matrice quadrata $A \in K^{n,n}$ si ha*

$$a_{i,1}A_{j,1} + a_{i,2}A_{j,2} + \cdots + a_{i,n}A_{j,n} = \begin{cases} det(A) & se \quad i = j \\ 0 & se \quad i \neq j \end{cases}$$

*Cioè il determinante di $A$ si può calcolare facendo la somma dei prodotti degli elementi di una riga o colonna per i rispettivi complementi algebrici. Mentre facendo la somma dei prodotti degli elementi di una riga per i complementi algebrici di una riga parallela si ottiene 0.*

*Proof.* Se si considerano tutti i prodotti dedotti che contengono $a_{i,j}$ si vede che a fattore di questo elemento si ha lo sviluppo del minore complementare ottenuto in $A$ sopprimendo la riga i-ma e la j-ma colonna. In effetti a fattore c'è il complemento algebrico perché bisogna tenere conto del segno del prodotto dedotto.
Quando si fa $a_{i,1}A_{j,1} + a_{i,2}A_{j,2} + \cdots + a_{i,n}A_{j,n}$, con $i \neq j$ è come se calcolassimo il determinante di una matrice con le righe $i$-ma e $j$-ma uguali. Come sappiamo tale determinante vale 0. □

Adesso ricordiamo alcuni importanti teoremi, senza dare la dimostrazione, ma che sono utili nella pratica e che comunque è bene conoscere.

**Theorema 3.3.3 (Binet).**

Il determinante del prodotto di due matrici $det(AB) = det(A)det(B)$.

**Theorema 3.3.4 (Kronecker).** *Se in una matrice $A \in K^{m,n}$ esiste un minore di ordine $p$ a determinante non nullo e tutti i minori di ordine maggiore di $p$ che lo contengono hanno determinante nullo allora il rango della matrice $A$ è $p$.*

**Theorema 3.3.5.** *Condizione necessaria e sufficiente perché una matrice quadrata $(A) \in K^{nn}$ sia* **invertibile** *è che $det(A) \neq 0$. In tal caso l'inversa di $(A)$ è la matrice $(A)^{-1} = \frac{1}{det(A)}{}^t(A_{i,j})$. La matrice $(A_{ij})$ si dice l'aggiunta di $(A)$.*

*Proof.* Se $(A)$ è invertibile sarà $(A)(A^{-1}) = I_n$. Prendendo i determinanti, per il Teorema di Binet, si ha $\det(A)\det(A^{-1}) = 1$. Quindi $\det(A) \neq 0$. Se

$\det(A) \neq 0$ consideriamo la matrice $\frac{1}{\det(A)} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$. Se si effettua il prodotto righe per colonne di $(A)$ per questa matrice, applicando i Teoremi di Laplace si ottiene facilmente la matrice identica $I_n$. □

## 3.4 Studio delle applicazioni lineari

Come abbiamo detto in precedenza studiare una applicazione lineare fra spazi vettoriali vuol dire determinare il nucleo ($Kerf$) e l'immagine ($Imf$) dell'applicazione $f : V \to W$. In verità il modo di procedere per tale studio dipende dal modo come viene data l'applicazione e dalla conoscenza delle basi a cui conviene riferirsi. Ciò detto siano $E = (e_1, e_2, \cdots, e_n)$ ed $(F = f_1, f_2, \cdots, f_m)$ due basi, una del dominio $V$ e l'altra del codominio $W$. Consideriamo la matrice $A = M^{E,F}(f)$. Ricordando come è definita la matrice associata ad $f$ possiamo dire che $[f(e_1)]_F, [f(e_2)]_F, \cdots, [f(e_n)]_F$, che sono le colonne della matrice associata, **danno le componenti, rispetto ad $F$, dell'immagine dell'applicazione**. Quindi basterà prendere le colonne indipendenti per trovare una base dell'immagine. Il nucleo, come sappiamo, è costituito da tutti i vettori $v \in V$ tali che $f(v) = 0_W$. Detto $\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ il vettore delle componenti di $v$ rispetto alla base $E$, deve accadere $A\underline{X} = \underline{0}$. Cioè si deve risolvere il sistema lineare omogeneo. Il numero di soluzioni dipende dal rango della matrice. In effetti si hanno $\infty^{n-r(A)}$ soluzioni, dove $n-r(A)$ indica il numero delle incognite libere. **Le soluzioni indipendenti danno le componenti, rispetto ad $E$, dei vettori che formano una base del nucleo**.
A chiarimento di ciò portiamo qualche esempio.

*Esempio* 3.4.1. Sia $A = M^{B,E}(f) = \begin{pmatrix} -1 & -2 & h+1 \\ -1 & -1 & h+1 \\ 0 & -1 & h \end{pmatrix}$ la matrice associata all'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$, relativamente alla base $B = (v_1 = (1,1,0), v_2 = (0,1,-1), v_3 = (0,0,1)$ del dominio e alla **base canonica** $E$ del codominio. Riduciamo per righe la matrice; sottraiamo alla seconda la prima riga e addizioniamo alla terza riga la seconda ottenuta. si ottiene la matrice $\begin{pmatrix} -1 & -2 & h+1 \\ 0 & 1 & 0 \\ 0 & 0 & h \end{pmatrix}$. Questa è ridotta per righe. È di rango tre per $h \neq 0$; in tal caso abbiamo un isomorfismo e quindi nucleo $0_V$ e immagine tutto il codominio $\mathbb{R}^3$. Se invece $h = 0$ si ottiene la matrice $\begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Questa matrice è di rango due; la prima e la seconda colonna sono indipendenti e una base dell'immagine si ottiene prendendo **dalla matrice data** $A$ la prima e la seconda colonna per il valore $h = 0$. Notare che le colonne ottenute sono componenti rispetto alla base canonica e quindi si prendono le colonne stesse. Per trovare il nucleo basta risolvere il sistema $A\underline{X} = \underline{0}$. Nel nostro caso si ha $\begin{cases} y & = 0 \\ x - z & = 0 \end{cases}$. Si hanno $\infty^1$ soluzioni $(x, 0, x)$, con $x$ variabile libera. Possiamo prendere come base per le soluzioni la terna $(1, 0, 1)$, che da' le componenti rispetto alla base $B$ di **una base del nucleo**. In definitiva una base del nucleo sarà: $1 \cdot (1,1,0) + 0 \cdot (0,1,-1) + 1 \cdot (0,1,-1) = (1,2,-1)$.

*Esempio* 3.4.2. Si consideri l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ definito da

$$f(x,y,z) = (hx - y + 3z, -x + hy + 3z, -x - y + (h+4)z)$$

e si calcoli, al variare di $h$, la controimmagine $f^{-1}(-1,1,0)$.

*Proof.* La matrice associata ad $f$, rispetto alle basi canoniche sia nel dominio che nel codominio, è $A = \begin{pmatrix} h & -1 & 3 \\ -1 & h & 3 \\ -1 & -1 & h+4 \end{pmatrix}$. Riduciamo per righe per determinare il suo rango. Addizionando alla seconda riga la prima moltiplicata per $h$ e sottraendo alla terza la prima riga si ha che l'elemento $-1$ sulla prima è speciale. Si ha

$$\begin{pmatrix} h & -1 & 3 \\ h^2-1 & 0 & 3(h+1) \\ -1-h & 0 & h+1 \end{pmatrix}.$$

Per $h = -1$ la matrice diventa $\begin{pmatrix} -1 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Il rango è uno e una base dell'immagine si trova prendendo una qualunque colonna della matrice data per $= -1$. Il nucleo si trova risolvendo il sistema lineare omogeneo $x + y - 3z = 0$, da cui $x = -y + 3z$. Il generico vettore del nucleo è $(-y + 3z, y, z)$. Come base del nucleo si può prendere $(-1,1,0); \quad (3,0,1)$.

Se invece $h + 1 \neq 0$ si continua a ridurre la matrice dove avevamo lasciato e con la sostituzione $R_3 \leftrightarrow R_3 - 3R_2$ si ottiene la matrice

$$\begin{pmatrix} h & -1 & 3 \\ -1-h & 0 & h+1 \\ h^2+3h+2 & 0 & 0 \end{pmatrix}$$

Si trova che $h^2 + 3h + 2 = 0$ per $h = -1$ e per $h = -2$.
Per $h = -2$ la matrice diventa $\begin{pmatrix} -2 & -1 & 3 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ e il suo rango è due. Una base dell'immagine è data da due colonne della matrice di partenza per $h = -2$ e una base del nucleo si ottiene dal sistema lineare omogeneo costruito su quest'ultima matrice. Si può prendere $(1,1,1)$.
Ovviamente per $h \neq -1, -2$ si ha un **isomorfismo**.

La controimmagine di $(-1,1,0)$ si ottiene dal sistema lineare $A\underline{X} = B$, dove $A$ è la matrice associata ad $f$ e $B = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Si riduce la matrice completa $(A|B)$ utilizzando le stesse riduzioni fatte nello studio della $f$ nel punto precedente.

Si deduce subito che per $h-1$ il sistema è incompatibile; mentre per $h+1 \neq 0$ riducendo ulteriormente si ha

$$\begin{pmatrix} h & -1 & 3 & | & -1 \\ -1-h & 0 & h+1 & | & 1 \\ h^2+3h+2 & 0 & 0 & | & -h-2 \end{pmatrix}.$$

In conclusione si vede che per $h=-2$ il sistema è compatibile ed ammette $\infty^1$ soluzioni, mentre per $h \neq -1, -2$ il sistema ammette una e una sola soluzione. I dettagli sono lasciati per esercizio. $\square$

# Chapter 4

# Autovalori Autovettori Forme Quadratiche

## 4.1 Autovalori

**Matrice del cambio di base in uno spazio vettoriale**

Sia $V$ un $K$-spazio vettoriale di dimensione $n$ e siano $\mathcal{E} = \{e_1, e_2, \cdots, e_n\}$ ed $\mathcal{F} = \{f_1, f_2, \cdots, f_n\}$ due sue basi. Un vettore $v \in V$ si può scrivere

$$v = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = y_1 f_1 + y_2 f_2 + \cdots + y_n f_n. \tag{4.1}$$

Vengono così determinati i due insiemi di componenti dello stesso vettore $v$, uno sarà $[v]_{\mathcal{E}} = \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e l'altro $[v]_{\mathcal{F}} = \underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

Ricordiamo che con $[v]_{\mathcal{E}}$ si denota il vettore colonna formato dalle componenti del vettore $v$ rispetto alla base $\mathcal{E}$.
Adesso tenuto conto che $\mathcal{E}$ è una base esprimiamo i vettori della base $\mathcal{F}$ in termini dei vettori della base $\mathcal{E}$.
Si ha

$$\begin{cases} f_1 = p_{11}e_1 + p_{21}e_2 + \cdots + p_{n1}e_n \\ f_2 = p_{12}e_1 + p_{22}e_2 + \cdots + p_{n2}e_n \\ \qquad \vdots = \vdots \\ f_n = p_{1n}e_1 + p_{2n}e_2 + \cdots + p_{nn}e_n \end{cases} \tag{4.2}$$

Resta individuata la matrice $P$ avente per colonne le componenti $[f_1]_{\mathcal{E}}, [f_2]_{\mathcal{E}}, \cdots, [f_n]_{\mathcal{E}}$.
Tale matrice si definisce **la matrice del cambio di base dalla base $\mathcal{E}$ alla base $\mathcal{F}$.**

Talvolta è bene evidenziare l'ordine delle basi e la matrice del cambio di base da $\mathcal{E}$ a $\mathcal{F}$ si indica $P^{\mathcal{E},\mathcal{F}}$.

Dalle (4.1) e dalle (4.2) si deduce che

$$\begin{aligned} v = & \; y_1 f_1 + y_2 f_2 + \cdots + y_n f_n = \\ & y_1(p_{11}e_1 + p_{21}e_2 + \cdots + p_{n1}e_n) + y_2(p_{12}e_1 + p_{22}e_2 + \cdots + p_{n2}e_n) \\ & \cdots + y_n(p_{1n}e_1 + p_{2n}e_2 + \cdots + p_{nn}e_n) = \\ & (p_{11}y_1 + p_{12}y_2 + \cdots + p_{1n}y_n)e_1 + (p_{21}y_1 + p_{22}y_2 + \cdots + p_{2n}y_n)e_2 + \\ & \cdots + (p_{n1}y_1 + p_{n2}y_2 + \cdots + p_{nn}y_n)e_n \end{aligned} \tag{4.3}$$

Tenendo conto delle (4.1) e (4.3) e del fatto che le componenti di un vettore rispetto ad una data base sono univocamente determinate segue che

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_{11}y_1 + p_{12}y_2 + \cdots + p_{1n}y_n \\ p_{21}y_1 + p_{22}y_2 + \cdots + p_{2n}y_n \\ \vdots \\ p_{n1}y_1 + p_{n2}y_2 + \cdots + p_{nn}y_n \end{pmatrix} \tag{4.4}$$

Identificando le componenti si può anche scrivere, in forma compatta

$$\underline{x} = P\underline{y} \tag{4.5}$$

dove con $\underline{y}$ si è indicato $\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

Le (4.5) si dicono le formule del **cambiamento di coordinate nello spazio vettoriale** dalla base $\mathcal{E}$ alla base $\mathcal{F}$.

Dalle (4.2) è facile convincersi che la matrice $P^{\mathcal{E},\mathcal{F}}$ coincide con la matrice $M^{\mathcal{F},\mathcal{E}}(i_V)$; dove $i_V$ è l'applicazione identica su $V$.

È anche facile verificare che in $V$, qualunque sia la base $\mathcal{E}$ scelta , la matrice $M^{\mathcal{E},\mathcal{E}}(i_V) = I_n$, cioè la matrice identica di ordine $n$.

Consideriamo adesso la composizione di applicazioni identiche su $V$

$$V^{\mathcal{E}} \xrightarrow{i_V} V^{\mathcal{F}} \xrightarrow{i'_V} V^{\mathcal{E}}$$

Tenuto conto che la matrice associata alla composizione $i'_V \circ i_V$ è il prodotto delle matrici associate alle applicazioni componenti $i'_V$ e $i_V$, con le basi che a loro competono si ha

$$I_n = M^{\mathcal{E},\mathcal{E}}(i'_V \circ i_V) = M^{\mathcal{F},\mathcal{E}}(i'_V) \cdot M^{\mathcal{E},\mathcal{F}}(i_V) \tag{4.6}$$

Da ciò si deduce che le matrici $P^{\mathcal{E},\mathcal{F}}$ e $P^{\mathcal{F},\mathcal{E}}$ sono l'una inversa dell'altra.

### Matrici Simili

Sia $V$ uno spazio vettoriale su $K$ di dimensione $n$ e $\varphi : V \longrightarrow V$ un endomorfismo su $V$. Se $\mathcal{E}$ è una base di $V$ allora si può costruire la matrice $M^{\mathcal{E},\mathcal{E}}(\varphi) = A$.

Ovviamente $A \in K^{n,n}$. Se si considera un'altra base $\mathcal{F}$ di $V$ si può considerare la matrice $M^{\mathcal{F},\mathcal{F}}(\varphi) = B$. Le due matrici $A$ e $B$ sono associate allo stesso endomorfismo $\varphi$ su $V$, ma sono riferite a basi diverse.

Due tali matrici vengono dette **simili**. Sussiste la seguente

**Proposizione 4.1.1.** *Due matrici $A$ e $B$ sono simili se e solo se $B = P^{-1}AP$, dove $P = P^{\mathcal{E},\mathcal{F}}$.*

*Proof.* Siano $A = M^{\mathcal{E},\mathcal{E}}(\varphi)$ e $B = M^{\mathcal{F},\mathcal{F}}(\varphi)$ due matrici associate alla stessa applicazione lineare $\varphi$ relativamente alle basi $\mathcal{E}$ ed $\mathcal{F}$. Per definizione le matrici sono simili; consideriamo la composizione delle applicazioni

$$V^{\mathcal{F}} \xrightarrow{i_V} V^{\mathcal{E}} \xrightarrow{\varphi} V^{\mathcal{E}} \xrightarrow{i'_V} V^{\mathcal{F}} \tag{4.7}$$

Per il teorema che dice che la matrice associata alla composizione di applicazioni lineari è la matrice prodotto delle matrici associate alle singole applicazioni, relativamente alle basi che loro competono, si ha

$$M^{\mathcal{F},\mathcal{F}}(i'_V \circ \varphi \circ i_V) = M^{\mathcal{E},\mathcal{F}}(i'_V) \cdot M^{\mathcal{E},\mathcal{E}}(\varphi) \cdot M^{\mathcal{F},\mathcal{E}}(i_V) \tag{4.8}$$

Da ciò segue $B = P^{-1}AP$.
Viceversa se $B = P^{-1}AP$, sia $A \in \mathbb{R}$ la matrice associata ad un endomorfismo $\varphi : \mathbb{R}^n \to \mathbb{R}^n$, rispetto alla base canonica $\mathcal{E}$. Sia $P = P^{\mathcal{E},\mathcal{F}}$ la matrice del cambio di base da $\mathcal{E}$ ad una nuova base $\mathcal{F}$; detta $B' = M^{\mathcal{F},\mathcal{F}}(\varphi)$, per la parte precedente si ha $B' = P^{-1}AP$. Ma allora $B = B'$ e la dimostrazione è completa. $\square$

**Autovalori e Autovettori**

Sia $V$ un $K$-spazio vettoriale di dimensione finita $n$ e $\varphi : V \to V$ un endomorfismo su $V$.

**Definizione 4.1.2.** Uno scalare $\lambda \in K$ si dice un **autovalore** di $\varphi$ se esiste un vettore $v \neq 0$ tale che $\varphi(v) = \lambda v$. Un tale vettore si definisce un **autovettore**.

Ovviamente se c'è un autovettore $v$ associato a $\lambda$ ce ne sono infiniti: tutti quelli del tipo $av$, con $a \neq 0$.

**Definizione 4.1.3.** Si definisce autospazio $V_\lambda$ il sottospazio costituito dai vettori $v$ tali che $\varphi(v) = \lambda v$.

È bene osservare che in $V_\lambda$ ci stanno tutti gli autovettori associati a $\lambda$ più lo zero di $V$.
Vedremo che tali nozioni si riveleranno di fondamentale importanza nello studio dell' algebra lineare.
Mostriamo adesso come si procede per il **calcolo degli autovalori** di un endomorfismo.

Consideriamo l'endomorfismo $\varphi : V \to V$. Fissiamo una base $\mathcal{E}$ in $V$ e diciamo $A = M^{\mathcal{E},\mathcal{E}}(\varphi)$. Per quanto visto in precedenza la relazione vettoriale $\varphi(v) = \lambda v$ si può esprimere, in forma matriciale, $A\underline{X} = \lambda\underline{X}$, dove $\underline{X}$ è il vettore colonna delle componenti di $v$ rispetto alla base $\mathcal{E}$. Ovviamente $\underline{X} \neq \underline{0}$ perché componenti di un autovettore che per definizione è diverso da zero. Da $A\underline{X} = \lambda\underline{X}$ si ha che si devono cercare soluzioni non nulle del sistema lineare omogeneo $(A - \lambda I)\underline{X} = \underline{0}$. Pertanto la matrice $(A - \lambda I)$ deve avere rango $< n$, e quindi il determinante $|A - \lambda I| = 0$.
Ci si rende subito conto che si ottiene un polinomio di grado $n$ in $\lambda$ eguagliato a zero, che si chiama il **polinomio caratteristico** di $\varphi$. Gli autovalori sono le radici del polinomio caratteristico che appartengono al campo base $K$ di $V$.

**Osservazione** 4.1.4. Da come è stato definito, potrebbe sembrare che il polinomio caratteristico dipenda dalla scelta della base in $V$. In effetti si prova subito che il polinomio caratteristico è indipendente dalla scelta della base. Se ci riferiamo ad un'altra base $\mathcal{F}$ e consideriamo la matrice $B = M^{\mathcal{F},\mathcal{F}}(\varphi)$, per quanto visto nella Proposizione 4.1.1 si ha $B = P^{-1}AP$. Quindi calcolando $|B - \lambda I| = |P^{-1}AP - \lambda I| = |P^{-1}AP - P^{-1}(\lambda I)P| = |P^{-1}(A - \lambda I)P| = |P^{-1}||A - \lambda I||P|$. Il precedente calcolo tiene conto del fatto che la matrice identica è permutabile con ogni matrice e poi applicando il teorema di Binet.

Sia $\lambda^* \in K$ un autovalore. Per calcolare l'autospazio $V_{\lambda^*}$ associato a $\lambda^*$, per quanto adesso visto, si deve risolvere il sistema lineare omogeneo $(A-\lambda^* I)\underline{X} = \underline{0}$; l'insieme delle soluzioni ci da' proprio $V_{\lambda^*}$. Si può dire che la $\dim V_{\lambda^*} = n - r(A-\lambda^* I)$. Tale numero si chiama la **molteplicità geometrica** dell'autovalore $\lambda*$

Dimostriamo adesso l'importante teorema

**Theorema 4.1.5 (Indipendenza degli autovettori).** *Siano* $\lambda_1, \lambda_2, \ldots, \lambda_s$ *gli autovalori distinti di un endomorfismo* $\varphi : V \to V$. *Consideriamo gli autovettori* $v_1, v_2, \ldots, v_s$ *rispettivamente negli autospazi* $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_s}$. *Essi sono linearmente indipendenti.*

*Proof.* Applichiamo il criterio di indipendenza lineare e operiamo ragionando per assurdo. Negando la tesi deve esistere un vettore $v_i \in \mathcal{L}(v_1, v_2, \ldots, v_{i-1})$, ed anzi sia $v_i$ il primo vettore che sia combinazione lineare dei precedenti. Si ha pertanto

$$v_i = a_1 v_1 + a_2 v_2 + \cdots + a_{i-1} v_{i-1} \tag{4.9}$$

Applicando $\varphi$ ad ambo i membri e tenendo conto che tutti i $v_i$ sono autovettori si ha

$$\lambda_i v_i = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_{i-1} \lambda_{i-1} v_{i-1} \tag{4.10}$$

Sostituendo a $v_i$ il valore di 4.9 e raggruppando i termini simili si ha

$$a_1(\lambda_i - \lambda_1)v_1 + a_2(\lambda_i - \lambda_2)v_2 + \cdots + a_{i-1}(\lambda_i - \lambda_{i-1})v_{i-1} = 0 \tag{4.11}$$

Poiché $v_i$ è il primo vettore ad essere combinazione lineare dei precedenti, segue che i vettori $v_1, v_2, \ldots, v_{i-1}$ sono linearmente indipendenti, da cui segue che i coefficienti della 4.11 devono essere nulli. Quindi $a_1 = 0, a_2 = 0, \ldots, a_{i-1} = 0$. Ed allora tenuto conto della 4.9 se ne deduce che $v_i = 0$. Ciò è assurdo perché gli autovettori, per definizione, sono non nulli. Quindi gli autovettori soddisfano il criterio di indipendenza lineare. □

Dal precedente risultato si deduce l'importante

**Corollario 4.1.6.** *La somma di autospazi* $V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_k}$ *è sempre* **una somma diretta**.

*Proof.* Supponiamo, per assurdo, che in $V$ ci sia un vettore $v$ che si possa scrivere con due diverse rappresentazioni in termini di autovettori:

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = b_1 v_1 + b_2 v_2 + \cdots + b_k v_k \tag{4.12}$$

la precedente da' luogo alla relazione

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_k - b_k)v_k = 0 \tag{4.13}$$

Ma gli autovettori sono linearmente indipendenti e quindi ne segue $a_i = b_i$ per ogni $i$. □

Adesso trattiamo un argomento molto importante e cioè la diagonalizzazione delle matrici. **Diagonalizare una matrice** quadrata $A \in K^{nn}$ significa daterminare , quando esiste, una matrice invertbile $P$ tale che $P^{-1}AP = D$, dove $D$ è una matrice diagonale, cioè una matrice che ha elementi tutti nulli fuori dalla diagonale principale. In altre parole una matrice è diagonalizzabile se è **simile** ad una matrice diagonale.

**Definizione 4.1.7.** Un endomorfismo $\varphi : V \to V$ è **semplice** se esiste una base $\mathcal{A}$ di $V$ formata da autovettori. In tal caso $M^{\mathcal{A}}(\varphi) = D$, dove $D$ è una matrice diagonale sulla cui diagonale ci stanno gli autovalori di $\varphi$.

È importante la seguente

**Proposizione 4.1.8.** *Sia $\varphi : V \to V$ un endomorfismo. Detto $\overline{\lambda}$ un qualunque autovalore di $\varphi$; sarà sempre*

$$1 \leq \dim V_{\overline{\lambda}} \leq m_{\overline{\lambda}}.$$

*Quindi la molteplicità geometrica è sempre minore o uguale a quella algebrica.*

*Proof.* Poiché un autovettore associato a $\overline{\lambda}$ è diverso da zero sarà sempre $1 \leq \dim V_{\overline{\lambda}}$. Adesso diciamo $(v_1, v_2, \cdots, v_s)$ una base di $V_{\overline{\lambda}}$. Completiamola ad una base $B$ di tutto $V$ con l'aggiunta dei vettori: $v_{s+1}, \cdots, v_n$. La matrice $A^B(\varphi)$ è del tipo seguente:

$$\begin{pmatrix} \overline{\lambda} & 0 & \cdots & 0 & a_{1,s+1} & \cdots & a_{1,n} \\ 0 & \overline{\lambda} & \cdots & 0 & a_{2,s+1} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \overline{\lambda} & a_{s,s+1} & \cdots & a_{s,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & 0 & b_{n,s+1} & \cdots & b_{n,n} \end{pmatrix}$$

per cui $|A - TI)| = (\overline{\lambda} - T)^s \cdot g(T)$, con $g(T)$ polinomio di grado $n - s$. Allora per definizione si ha che se $g(\lambda_i) \neq 0$ la molteplicità algebrica di $\lambda_i = s$. Se invece $g(\overline{\lambda}) = 0$ la molteplicità algebrica sarà maggiore di $s$. □

**Theorema 4.1.9.** *Condizione necessaria e sufficiente perché $\varphi$ sia semplice è che*

1. *tutti gli autovalori $\lambda_i$ appartengano al campo base $K$;*
2. *per ogni $\lambda_i$ accada $\dim V_{\lambda_i} = m_{\lambda_i}$, cioè la molteplicità geometrica $g_{\lambda_i}$ sia uguale alla molteplicità algebrica $m_{\lambda_i}$.*

*Proof.* Se $\varphi$ è semplice esiste una base di autovettori. Si considerano tutti gli autospazi $V_{\lambda_i}$ raggruppando tutti gli autovettori associati allo stesso autovalore $\lambda_i$. Poiché per ogni $i$ si ha $\dim V_{\lambda_i} \leq m_{\lambda_i}$, per potere coprire $V$ mediante $\sum V_{\lambda_i}$ deve necessariamente essere $\dim V_{\lambda_i} = m_{\lambda_i}$, per ogni $i$. Ovvio che tutti gli autovalori appartengano al campo base $K$. Viceversa se sono vere 1. e 2. abbiamo $\dim V_{\lambda_i} = m_{\lambda_i}$, per ogni $i$. Quindi $\sum \dim V_{\lambda_i} = n$, poiché la somma di autospazi è diretta e $\sum m_{\lambda_i} = n$. Ne segue che l'endomorfismo è semplice. □

**Corollario 4.1.10.** *Sia $\varphi : V \to V$ un endomorfismo. Se accade che tutti gli autovalori sono radici semplici del polinomio caratteristico e stanno nel campo base allora l'endomorfismo è semplice.*

*Esempio* 4.1.11. Si consideri l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ che ha come matrice associata rispetto alla base $\mathcal{B} : (v_1 = (1,1,0); v_2 = (0,1,-1); v_3 = (0,0,1))$ e alla base canonica $\mathcal{E}$ la matrice $M^{\mathcal{B},\mathcal{E}}(f) = \begin{pmatrix} -1 & -2 & h+1 \\ -1 & -1 & h+1 \\ 0 & -1 & h \end{pmatrix}$.

1. Studiare l'applicazione $f$, al variare di $h$, determinando in ogni caso $Ker(f)$ e $Im(f)$.

2. Dopo aver trovato la matrice $M^{\mathcal{B},\mathcal{B}}(f)$ studiare la semplicità di $f$, al variare di $h$, e trovare una base di autovettori, quando $f$ non è un isomorfismo.

3. Determinare il sottospazio $V \subseteq \mathbb{R}^3$ tale che $f(V) \subseteq W$, dove $W = \{(y_1, y_2, y_3)| \quad y_1 = y_2 = 0\}$. Qual'è il valore di $h$ per cui $V = \mathcal{L}(e_3)$.

*Proof.* 1. Per studiare l'applicazione lineare si può usare la matrice $M^{\mathcal{B},\mathcal{E}}(f) = \begin{pmatrix} -1 & -2 & h+1 \\ -1 & -1 & h+1 \\ 0 & -1 & h \end{pmatrix}$. Riducendo per righe si ottiene facilmente la matrice ridotta $\begin{pmatrix} -1 & -2 & h+1 \\ 0 & 1 & 0 \\ 0 & 0 & h \end{pmatrix}$.
Per $h \neq 0$ si ha un isomorfismo.
Per $h = 0$ la nostra matrice diventa $\begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Quindi la dimensione dell'immagine è due e una base dell'immagine è data dalle colonne $C_1$ e $C_2$ della matrice data, per $h = 0$. Cioè $C_1 = (-1,-1,0)$ e $C_2 = (-2,-1,,-1)$. Per il nucleo si ha $y = 0;\ x = z$. E quindi $(1,0,1)_{\mathcal{B}}$ danno le componenti di una base del nucleo, rispetto alla base $\mathcal{B}$; quindi una base del nucleo è $(1,1,1)$.
2. Per studiare la semplicità viene suggerito di usare la matrice $M^{\mathcal{B},\mathcal{B}}(f)$. Si devono scrivere sulle colonne le componenti , rispetto alla base $\mathcal{B}$, delle immagini dei vettori della base $\mathcal{B}$.
Troviamo le componenti del generico vettore $(a,b,c)$, rispetto alla base $\mathcal{B}$.
Scriviamo $(a,b,c) = x(1,1,0) + y(0,1,-1) + z(0,0,1)$. Eguagliando le componenti si ha il sisstema $\begin{cases} x & = & a \\ x+y & = & b \\ -y+z & = & c \end{cases}$. Da questo si deducono le componenti cercate: $x = a;\ y = b - a; z = b - a + c$. Dalla matrice $M^{\mathcal{B},\mathcal{E}}(f)$ si ha $f(v_1) = (-1,-1,0);\ f(v_2) = (-2,-1,-1); f(v_3) = (h+1, h+1, h)$. La matrice cercata è quindi la matrice $A = M^{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} -1 & -2 & h+1 \\ 0 & 1 & 0 \\ 0 & 0 & h \end{pmatrix}$. La matrice caratteristica è $(A - TI) = \begin{pmatrix} -1-T & -2 & h+1 \\ 0 & 1-T & 0 \\ 0 & 0 & h-T \end{pmatrix}$ e gli autovalori sono $T_1 = -1;\ T_2 = 1;\ T_3 = h$.

Per $h \neq \pm 1$ gli autovalori sono reali e distinti e quindi l'endomorfismo è semplice.
Per $h = -1$ si ha $T_1 = T_3 = -1$; la matrice caratteristica per questi valori diventa $\begin{pmatrix} 0 & -2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$; essa ha rango 1 e quindi la $\dim V_{-1} = 3 - 1 = 2$. Cioè la molteplicità geometrica è uguale a quella algebrica e l'endomerfismo è semplice.
Per $h = 1$ si ha $T_2 = T_3 = 1$. Anche in questo caso l'endomorfismo è semplice.
L'endomorfismo non è un isomorfismo per $h = 0$. La matrice diventa
$\begin{pmatrix} -1-T & -2 & 1 \\ 0 & 1-T & 0 \\ 0 & 0 & -T \end{pmatrix}$; per $T = 0$ si ha $y = 0; x = z$. quindi un autovettore associato a $T = 0$ è $(1, 0, 1)_{\mathcal{B}}$. Un autovettore associato a $T = -1$ è dato da $(1, 0, 0)_{\mathcal{B}}$. Infine un autovettore associato a $T = 1$ ha componenti $(1, -1, 0)_{\mathcal{B}}$.
3. Siano $(x, y, z)$ le componenti di un vettore generico $v \in \mathbb{R}^3$. Usando la matrice $M^{\mathcal{B},\mathcal{E}}(f)$ si trovano le componenti del vettore immagine. Si trova il vettore
$\begin{pmatrix} y_1 & = & -x - 2y + (h+1)z \\ y_2 & = & -x - y + (h+1)z \\ y_3 & = & -y + hz \end{pmatrix}$. Da cui si ha $\begin{cases} y = 0 \\ x = (h+1)z \end{cases}$. Quindi il sottospazio $V$ cercato è dato dal generico vettore $((h+1)z, 0, z)$.
Se si vuole ottenere $V = \mathcal{L}(e_3)$ si deve porre $h = -1$. $\square$

## 4.2 Spazi con prodotto scalare

Come è noto per i vettori dello spazio ordinario si definisce il **prodotto scalare**. Se uno dei due vettori è nullo il prodotto scalare è zero; altrimenti $v \bullet w = |v||w| \cos\theta$, dove $|v|$ e $|w|$ sono i moduli di $v$ e $w$ e $\theta$ è l'angolo formato dai due vettori.

Si deduce facilmente che due vettori sono ortogonali se e solo se il loro prodotto scalare è zero, e si scrive $v \perp w$.

Conosciamo molti altri esempi di spazi con prodotto scalare. Ricordiamo lo spazio euclideo reale $\mathbb{R}^3 = \{(x_1, x_2, x_3) | x_i \in \mathbb{R}\}$ dove il prodotto scalare è definito da

$$(x_1, x_2, x_3) \bullet (y_1, y_2, y_3) = x_1y_1 + x_2y_2 + x_3y_3.$$

Mentre **nel caso complesso** il prodotto scalare viene definito come segue:

$$(u_1, u_2, u_3) \bullet (z_1, z_2, z_3) = u_1\overline{z}_1 + u_2\overline{z}_2 + u_3\overline{z}_3$$

Le proprietà di cui godono i prodotti scalari sopra definiti, ci permetteranno di dare una definizione astratta di spazio vettoriale, reale o complesso, con prodotto scalare, secondo le definizioni che seguono.

**Definizione 4.2.1.** Uno spazio $V$ su $\mathbb{R}$ è uno spazio con prodotto scalare se è data una applicazione $V \times V \to \mathbb{R}$, che ad ogni coppia ordinata di vettori $v, w \in V$ associa un numero reale $v \bullet w \in \mathbb{R}$ che soddisfa le seguenti condizioni

1. $v \bullet w = w \bullet v$, per ogni scelta dei vettori.
2. per ogni $v$ si ha $v \bullet v \geq 0$; l'eguaglianza si ha se e solo se $v = 0$;
3. $(u + v) \bullet w = u \bullet w + (v \bullet w)$, per ogni scelta di vettori;
4. $(av) \bullet w = a(v \bullet w)$ per ogni scelta di $a, v, w \in \mathbb{R}$

**Definizione 4.2.2.** Uno spazio $V$ su $\mathbb{C}$ è uno spazio complesso con prodotto scalare se è data una applicazione $V \times V \to \mathbb{C}$ che ad ogni coppia ordinata di vettori $v, w \in V$ associa un numero complesso $v \bullet w \in \mathbb{C}$ che soddisfa le seguenti condizioni

1. $v \bullet w = \overline{w \bullet v}$, per ogni scelta dei vettori.
2. per ogni $v$ si ha $v \bullet v \geq 0$; l'eguaglianza si ha se e solo se $v = 0$;
3. $(u + v) \bullet w = u \bullet w + (v \bullet w)$, per ogni scelta di vettori;
4. $(av) \bullet w = a(v \bullet w)$ per ogni scelta di $a, v, w \in \mathbb{C}$.

Gli esempi all'inizio del paragrafo ovviamente soddisfano le condizioni suddette.

**Osservazione** 4.2.3. **A)** Dalla 1. segue facilmente che $v \bullet v$ è un numero reale, il che permette di giustificare la proprietà 2.

**B)** Accade che $v \bullet (aw) = \overline{((aw) \bullet v)} = \overline{a}\overline{(w \bullet v)} = \overline{a}(v \bullet w)$.

**C)** Si prova facilmente che: $u \bullet (v + w) = u \bullet v + u \bullet w$.

Nello spazio euclideo reale $\mathbb{R}^n$, pensando alle $n$-uple come vettori colonna, il prodotto scalare si può indicare nel modo seguente:

$$(x_1, x_2, \cdots, x_n) \bullet (y_1, y_2, \cdots, y_n) = {}^t\underline{X} \cdot \underline{Y}$$

Nello spazio complesso $\mathbb{C}^n$ si avrà

$$(x_1, x_2, \cdots, x_n) \bullet (y_1, y_2, \cdots, y_n) = {}^t\underline{X} \cdot \overline{\underline{Y}}$$

Questa notazione sarà utile nel seguito.
Ora introduciamo il concetto di **norma** di un vettore che "corrisponde" al concetto di modulo di un vettore di uno spazio euclideo.

**Definizione 4.2.4.** Sia $V$ un $\mathbb{R}$-spazio vettoriale. Una **norna** in $V$ è una applicazione $|| \; || : V \to \mathbb{R}$ che gode delle seguenti proprietà (dove $a \in \mathbb{R};\ v, w \in V$):

1. $||v|| \geq 0$ e $||v|| = 0$ se e solo se $v = 0$;
2. $||(a \cdot v)|| = |a| ||v||$; $|a|$ vuol dire valore assoluto di $a$;
3. $||v + w|| \leq ||v|| + ||w||$; questa è detta **disuguaglianza triangolare**.

È bene osservare che in uno spazio con prodotto scalare possono esserci diverse norme. Ciò verrà approfondito in uno studio più avanzato.

In uno spazio con prodotto scalare normato si chiama **versore** un vettore di norma uno. E per normalizzare un vettore $v \neq 0$ basta dividere il vettore per la sua norma. Quindi in $V$ con prodotto scalare $\bullet$ il vettore $v \neq 0$ si normalizza prendendo il vettore $\frac{v}{||v||}$ che è evidentemente di norma 1.

Se $V$ è uno spazio con prodotto scalare $\bullet$ la **norma associata** al prodotto scalare è data da $||v|| = \sqrt{v \bullet v}$. Inoltre due vettori $v$ e $w$ sono ortogonali se il loro prodotto scalare $v \bullet w = 0$. In tal caso si scrive $v \perp w$.

**Definizione 4.2.5.** Una base di vettori $(e_1, e_2, \cdots, e_n)$ di uno spazio $V$ con prodotto scalare si dice **una base ortonormale** se

$$e_i \bullet e_j = \delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases},$$

dove $\delta_{ij}$ è il delta di Kronecker.

È interessante osservare che se $(e_1, e_2, \cdots, e_n)$ è una base ortonormale di $V$ le componenti di un vettore $v \in V$ si trovano mediante prodotti scalari. Precisamente $v = (v \bullet e_1)e_1 + (v \bullet e_2)e_2 + \cdots + (v \bullet e_n)e_n$.

Adesso proviamo come sia sempre possibile trovare una base ortonormale.

**Proposizione 4.2.6.** *Se $(v_1, v_2, \cdots, v_n)$ è una base di uno spazio $V$ con prodotto scalare $\bullet$ è possibile a partire da essa di costruire una base ortonormale, con un procedimento detto di* **ortonormalizzazione di Gram-Schmidt**.

*Proof.* Si parte dal prima vettore $v_1$ e lo si normalizza dividendolo per la sua norma. Si ottiene il versore $e_1 = \frac{v_1}{||v_1||}$. Consideriamo il vettore $v_2 - (v_2 \bullet e_1)e_1$ e si vede subito che esso è ortogonale ad $e_1$. Facendo il prodotto scalare $[v_2 - (v_2 \bullet e_1)e_1] \bullet e_1 = v_2 \bullet e_1 - (v_2 \bullet e_1)(e_1 \bullet e_1) = 0$. Per rendere questo vettore un versore dobbiamo normalizzarlo. Si ottiene il versore

$$e_2 = \frac{v_2 - (v_2 \bullet e_1)e_1}{||v_2 - (v_2 \bullet e_1)e_1||}$$

si continua; sarà

$$e_3 = \frac{v_3 - [(v_3 \bullet e_1)e_1 + (v_3 \bullet e_2)e_2]}{||v_3 - [(v_3 \bullet e_1)e_1 + (v_3 \bullet e_2)e_2]||}$$

e così via fino all'ultimo vettore $v_n$. $\square$

**Proposizione 4.2.7.** *Una matrice quadrata $P \in \mathbb{R}^{nn}$ si dice* **ortogonale** *se $P \cdot {}^tP = I_n$. Ne segue che*

1. *$P$ è invertibile e la sua inversa è $P^{-1} = {}^tP$.*

2. *Anche la trasposta ${}^tP$ è ortogonale.*

**Proposizione 4.2.8.** *Se $V$ è uno spazio reale con prodotto scalare, $\mathcal{E}$ ed $\mathcal{F}$ due basi di $V$ e $P$ la matrice del cambio di base da $\mathcal{E}$ ad $\mathcal{F}$. Allora $P$* **trasforma basi ortonormali in basi ortonormali** *se e solo se $P$ è una matrice ortogonale.*

*Proof.* Come sappiamo la matrice del cambio di base è la matrice che ha per colonne ordinatamente le componenti della base $\mathcal{F}$ rispetto alla base $\mathcal{E}$. Si ha quindi

$$\begin{array}{lcl} f_1 & = & p_{11}e_1 + p_{21}e_2 + \cdots + p_{n1}e_n \\ f_2 & = & p_{12}e_1 + p_{22}e_2 + \cdots + p_{n2}e_n \\ \cdots & \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n & = & p_{1n}e_1 + p_{2n}e_2 + \cdots + p_{nn}e_n \end{array} \tag{4.14}$$

Facendo tutti i prodotti e tenendo conto che $\mathcal{E}$ è ortonormale si ha:

$$f_i \bullet f_j = p_{1i}p_{1j} + p_{2i}p_{2j} + \cdots + p_{ni}p_{nj}.$$

Da ciò si deduce ${}^tP \cdot P = (f_i \bullet f_j)$. Si deduce quindi che la matrice $P$ è ortogonale, cioè ${}^tP \cdot P = I_n$, se e solo se la base $\mathcal{F}$ è ortonormale. □

**Corollario 4.2.9.** *Sia $P \in \mathbb{R}^{n,n}$. Sono equivalenti le seguenti condizioni:*

1. *La matrice $P$ è ortogonale.*
2. *Le colonne di $P$ formano una base ortonormale di $\mathbb{R}^n$.*
3. *Le righe di $P$ formano una base ortonormale di $\mathbb{R}^n$.*

La dimostrazione è lasciata al lettore.

**Corollario 4.2.10.** *Il determinante di una matrice ortogonale $P$ è $\pm 1$. Le matrici con determinante +1 si dicono* **ortogonali di tipo speciale***; le altre non speciali.*

*Esempio* 4.2.11. La matrice $P = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$ è una matrice ortogonale di tipo speciale. Le sue colonne formano una base ortonormale per lo spazio euclideo $\mathbb{R}^2$. Questa matrice è importante perchè, come vedremo, è la matrice dei cambiamenti di riferimento che sono rotazioni antiorarie.

I Teoremi che seguono li dimostreremo per spazi vettoriali $V$ sui complessi $\mathbb{C}$. Le dimostrazioni nel caso di spazi $V$ sui reali $\mathbb{R}$ sono analoghe e anche più facili.

**Theorema 4.2.12 (Teorema di Pitagora-2500 anni fa).** *Sia $V$ uno spazio vettoriale complesso con prodotto scalare. Dati due qualunque vettori $v$ e $w$ ortogonali si ha*

$$||v + w||^2 = ||v||^2 + ||w||^2$$

*Proof.* La dimostrazione lasciata al lettore. □

**Proposizione 4.2.13 (Disegueglianza di Cauchy-Schwartz).** *Sia $V$ uno spazio vettoriale complesso con prodotto scalare. Dati due qualunque vettori $u$ e $v$, sussiste la seguente diseguaglianza*

$$|u \bullet v| \leq ||u|| \cdot ||v||$$

*dove $|u \bullet v|$ denota il modulo del numero complesso $u \bullet v$ e $||u||$ la norma associata al prodotto scalare.*

*Proof.* Sussiste la seguente decomposizione $u = av + u - av$, con $a$ da determinare. Scegliamo $a$ in modo che il vettore $u - av$ sia ortogonale a $v$. Cioè deve accadere che $(u - av) \bullet v = 0$. Quindi $a = \frac{u \bullet v}{||v||^2}$. Allora avremo

$$u = \frac{u \bullet v}{||v||^2} \cdot v + \Big(u - \frac{u \bullet v}{||v||^2} \cdot v\Big)$$

Allora visto che $w = \Big(u - \frac{u \bullet v}{||v||^2} \cdot v\Big)$ è stato scelto ortogonale a $v$, per il Teorema di Pitagora 4.2.12, avremo

$$||u||^2 = \left\|\frac{(u \bullet v)v}{||v||^2}\right\|^2 + ||w||^2 \geq \frac{|u \bullet v|^2}{\|v\|^2}$$

Da ciò segue facilmente

$$|u \bullet v| \leq \|u\| \cdot \|v\|.$$

$\square$

### 4.2.1 Sottospazi ortogonali

Sia $V$ uno spazio vettoriale con prodotto scalare. Si definisce il sottospazio ad esso ortogonale $W^{\perp} = \{v \in V : \ v \bullet w = 0 \ \forall w \in W\}$. È evidente da tale relazione che $W \subseteq (W^{\perp})^{\perp}$.

Due sottospazi $Z$ e $W$ sono **mutuamente ortogonali** se accade che $z \bullet w = 0$ per ogni $z \in Z, w \in W$. È bene ricordare che **si lavora sempre con spazi vettoriali di dimensione finita**. La proposizione seguente ci da' proprietà dei sottospazi ortogonali.

**Proposizione 4.2.14.** *Sia $W \subseteq V$ un sottospazio di uno spazio $V$ con prodotto scalare.*

*a)* $V = W \oplus W^{\perp}$;

*b)* $(W^{\perp})^{\perp} = W$.

*Proof.* a) Consideriamo una base ortonormale $(e_1, e_2, \cdots, e_r)$ di $W$. La si può completare ad una base ortonormale $(e_1, e_2, \cdots .e_r, e_{r+1}, \cdots, e_n)$ di tutto $V$. Sia $V' = \mathcal{L}(e_{r+1}, e_{r+2}, \cdots, e_n)$. Si vede subito che $V' \subseteq W^{\perp}$, in quanto ogni generatore di $V'$ è ortogonale ad ogni vettore della base di $W$. Consideriamo adesso un qualunque vettore $v \in W^{\perp}$. il vettore $v$ si può scrivere $v = (v \bullet e_1)e_1 +$

$(v \bullet e_2)e_2 + \cdots + (v \bullet e_n)e_n$. Ma $v \in W^\perp$ e quindi $v \bullet e_1 = 0; v \bullet e_2 = 0; \cdots ; v \bullet e_r = 0$. Ne segue $W^\perp = \mathcal{L}(e_{r+1}, \cdots , e_n)$. Essendo $W \cap W^\perp = 0$ si ha $V = W \oplus W^\perp$.

b) Intanto sappiamo che $W \subseteq (W^\perp)^\perp$. Si hanno le relazioni

$$\begin{aligned} \dim V &= \dim W + \dim W^\perp; \\ \dim V &= \dim W^\perp + \dim(W^\perp)^\perp \end{aligned}$$

Da queste si deduce $\dim W = \dim(W^\perp)^\perp$, e quindi $W = (W^\perp)^\perp$. □

**Definizione 4.2.15.** Una matrice $P \in \mathbb{C}^{n,n}$ si dice **unitaria** se $P \cdot {}^t\overline{P} = I_n$.

In effetti è una matrice invertibile e la sua inversa coincide con la trasposta della coniugata.

Si può dimostrare, in modo analogo a quanto fatto nel caso reale,

**Definizione 4.2.16.** Se $V$ è uno spazio vettoriale con prodotto scalare su $\mathbb{C}$ e $\mathcal{E}$ ed $\mathcal{F}$ due basi di $V$. Diciamo $P$ la matrice del cambio di base $P^{\mathcal{E},\mathcal{F}}$. Allora se $\mathcal{E}$ è una base ortonormale di $V$ la nuova base $\mathcal{F}$ è ortonormale se e solo se la matrice $P$ è unitaria.

*Proof.* La dimostrazione è analoga a quella fatta per il caso reale; bisogna prestare attenzione a come è fatto il prodotto scalare nel caso complesso e alle proprietà di cui gode. Con le notazioni solite si ha:

$$\begin{aligned} f_i &= p_{1,i}e_1 + p_{2,i}e_2 + \cdots + p_{n,i}e_n \\ f_j &= p_{1,j}e_1 + p_{2,j}e_2 + \cdots + p_{n,j}e_n \\ f_i \bullet f_j &= p_{1,i}\overline{p}_{1,j} + p_{2,i}\overline{p}_{2,j} + \cdots + p_{n,i}\overline{p}_{n,j} \end{aligned}$$

Questo è l'elemento di posto $(i, j)$ nella matrice ${}^tP\overline{P}$. La conclusione segue facilmente. □

Per completare il quadro diamo la

**Definizione 4.2.17.** Una matrice $A \in \mathbb{C}^{n,n}$ si dice **Hermitiana** se ${}^tA = \overline{A}$. Una matrice reale è Hermitiana se e solo se è simmetrica.

### 4.2.2 Endomorfismi autoaggiunti

Cominciamo a considerare spazi vettoriali $V$ su $\mathbb{R}$ con prodotto scalare. Il caso di spazi sui complessi si tratta più o meno allo stesso modo. In un paragrafo a parte metteremo in evidenza le particolarità che si hanno negli spazi complessi.

**Definizione 4.2.18.** Un endomorfismo $f : V \to V$ di uno spazio reale con prodotto scalare si dice **autoaggiunto** se si ha $f(v) \bullet w = v \bullet f(w)$, per ogni scelta dei vettori.

Sussiste la seguente proprietà: per provare che un endomorfismo è autoaggiunto basta verificare la proprietà di autoaggiunzione su una base.

**Proposizione 4.2.19.** *Sia $f : V \to V$ un endomorfismo di un $\mathbb{R}$-spazio vettoriale. Le seguenti condizioni sono equivalenti*

*a) $f$ è autoaggiunto;*

*b) detta $\mathcal{E} = (e_1, e_2, \cdots, e_n)$ una base di $V$ accade che $f(e_i) \bullet e_j = e_i \bullet f(e_j)$, per ogni scelta di vettori in $\mathcal{E}$.*

*Proof.* la $a) \Longrightarrow b)$ è ovvia;

Per la implicazione $b) \Longrightarrow a)$ bisogna provare che per ogni scelta di $v$ e $w$ sia $f(v) \bullet w = v \bullet f(w)$. Essendo $\mathcal{E}$ una base $v = a_1e_1 + a_2e_2 + \cdots + a_ne_n$ e $w = b_1e_1 + b_2e_2 + \cdots + b_ne_n$. Tenendo conto della linearità della $f$ e della proprietà b) si ha

$$\begin{aligned} f(v) \bullet w &= (a_1f(e_1) + a_2f(e_2) + \cdots + a_nf(e_n)) \bullet (b_1e_1 + b_2e_2 + \cdots + b_ne_n) \\ &= a_1b_1(f(e_1) \bullet e_1) + \cdots + a_1b_n(f(e_1) \bullet e_n) + a_nb_1(f(e_n) \bullet e_1) + \cdots \\ &+ a_nb_n(f(e_n) \bullet e_n) \end{aligned}$$

se ora teniamo conto dell'ipotesi b) e sostituiamo, per ogni $(i, j)$, $f(e_i) \bullet e_j$ con $e_i \bullet f(e_j)$ otteniamo $v \bullet f(w)$. □

**Lemma 4.2.20.** *Sia $f : V \to V$ un endomorfismo autoaggiunto di un $\mathbb{R}$-spazio $V$. Per ogni $\alpha, \beta \in \mathbb{R}$ tali che $\alpha^2 - 4\beta < 0$ accade che*

$$f^2 + \alpha f + \beta I$$

*è invertibile.*

*Proof.* Per ogni $v \neq 0$ consideriamo

$$\begin{aligned} ((f^2 + \alpha f + \beta I)(v) \bullet v) &= (f(v) \bullet f(v)) + \alpha(f(v) \bullet v) + \beta(v \bullet v) \\ &= ||f(v)||^2 + \alpha f(v) \bullet v + \beta||v||^2 \\ &\geq ||f(v)||^2 - |\alpha| \cdot ||f(v)|| \cdot ||v|| + \beta||v||^2 \\ &\quad \text{si è usata la diseguaglianza di Cauchy-Schwartz} \\ &= \left(||f(v)|| - \frac{|\alpha|||v||}{2}\right)^2 + (\beta - \frac{\alpha^2}{4})||v||^2 \\ &= \geq + > 0 > 0 \end{aligned}$$

Per ogni $v \neq 0$ si ha $(f^2+\alpha f+\beta I)(v) \neq 0$, altrimenti saremmo in contraddizione con la precedente. Il che significa che $f^2+\alpha f+\beta I$ è iniettiva, quindi è invertibile. $\square$

**Proposizione 4.2.21.** *Se $f : V \to V$ è un endomorfismo autoaggiunto di un $\mathbb{R}$-sapazio $V$, allora $f$ ha un autovalore.*

*Proof.* Ovviamente il teorema si fa per uno spazio reale perché su uno spazio complesso sappiamo che esistono autovalori. Consideriamo un vettore $v \neq 0$. Allora $(v, f(v), f^2(v), \cdots, f^n(v))$ è un insieme di vettori dipendenti. Quindi esiste una combinazione lineare nulla con coefficienti non tutti nulli. E sia:

$$a_0 v + a_1 f(v) + \cdots + a_n f^n(v) = 0.$$

Il polinomio con gli stessi coefficienti, nella variabile $x$, si decompone

$$0 = a_0+a_1x+\cdots+a_nx^n = c(x^2+\alpha_1x+\beta_1)\cdots(x^2+\alpha_rx+\beta_r)(x-\lambda_1)\cdots(x-\lambda_s)$$

Osserviamo che $c, \alpha_i, \beta_i, \lambda_i$ sono reali, con discriminanti $\Delta_i = \alpha_i^2 - 4\beta_i < 0$. Ovviamente la decomposizione vale per ogni $x$. E quindi abbiamo

$$\begin{aligned} 0 &= a_0v + a_1f(v) = \cdots + a_nf^n(v) \\ &= (a_0I + a_1f + \cdots + a_nf^n)(v) \\ &= c(f^2 + \alpha_1 f + \beta_1 I)\cdots(f^2 + \alpha_r f + \beta_r I)(f - \lambda_1)\cdots(f - \lambda_s)(v) \end{aligned}$$

Ciò per la proprietà di decomposizione dei polinomi; si sa anche che $\forall\ i \Delta_i = \alpha_i^2 - 4\beta_i < 0$. Per il Lemma precedente si ha che $f^2 + \alpha_i f + \beta_i$ è invertibile per ogni $i$. Si ha allora

$$0 = (f - \lambda_1 I)(f - \lambda_2 I)\cdots(f - \lambda_s I)(v)$$

Si conclude che non tutti gli $(f - \lambda_i I)$ possono essere iniettivi; quindi ci sarà necessariamente un autovalore. $\square$

**Proposizione 4.2.22 (Endomorfismi autoaggiunti e matrici simmetriche).** *Sia $V$ uno spazio vettoriale con prodotto scalare, $f : V \to V$ un endomorfismo ed $\mathcal{E} = (e_1, e_2, \cdots, e_n)$ una base ortonormale. Detta $A = M^{\mathcal{E}}(f)$ si ha che le seguenti condizioni sono equivalenti:*

1. *l'endomorfismo è autoaggiunto;*
2. *La matrice $A$ è simmetrica.*

*Proof.* Essendo la base $\mathcal{E}$ ortonormale, per definizione la matrice associata ad $f$ rispetto ad $\mathcal{E}$ si scrive

$$\begin{pmatrix} f(e_1)\bullet e_1 & f(e_2)\bullet e_1 & \cdots & f(e_n)\bullet e_1 \\ f(e_1)\bullet e_2 & f(e_2)\bullet e_2 & \cdots & f(e_n)\bullet e_2 \\ \cdots & \cdots & \cdots & \cdots \\ f(e_1)\bullet e_n & f(e_2)\bullet e_n & \cdots & f(e_n)\bullet e_n \end{pmatrix}$$

Ora è evidente che 1. e 2. sono equivalenti. Basta tenere conto delle definizioni. $\square$

**Theorema 4.2.23** (**Teorema spettrale reale**). *Sia $V$ uno spazio vettoriale reale di dimensione finita con prodotto scalare ed $f : V \to V$ un endomorfismo. Allora le seguenti condizioni sono equivalenti:*

1. *$f$ è autoaggiunto;*

2. *$f$ è semplice e gli autospazi sono a due a due ortogonali;*

3. *Esiste in $V$ una base ortonormale formata da autovettori.*

*Proof.* Proviamo 1. $\Longrightarrow$ 2. Intanto ci sono autovalori e quindi autovettori. Diciamo $W = V_{\lambda_1} \oplus +V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$ il sottospazio contenente tutti gli autovettori. Prendiamo due qualunque autovettori $v_i \in V_{\lambda_i}$ e $v_j \in V_{\lambda_j}$. Si ha $f(v_i) \bullet w_j = v_i \bullet f(v_j) = \lambda_i(v_i \bullet v_j) = \lambda_j(v_i \bullet v_j) = 0$, da cui $(\lambda_i - \lambda_j)(v_i \bullet v_j) = 0$ e quindi $v_i \bullet v_j = 0$.

Adesso se $W = V$ l'implicazione è provata. Se invece $W \subset V$ consideriamo l'ortogonale $W^\perp$ che sarà certamente non nullo. Ora se $w^* \in W$ seguirà che $f(w^*) \in W$. Cioè la restrizione di $f$ a $W$ è un endomorfismo. Analogamente la restrizione di $f$ a $W^\perp$ è un endomorfismo su $W^\perp$. Infatti se $\overline{w} \in W^\perp$ segue $f(\overline{w}) \bullet w = \overline{w} \bullet f(w) = \overline{w} \bullet w' = 0$. E poiché $W^\perp \neq 0$ ci sarà un autovalore e quindi un autovettore $v^* \in W^\perp$. Ma tutti gli autovettori li avevamo presi in $W$ quindi necessariamente $v^* = 0$. Ma ciò è assurdo perché gli autovettori per definizione sono non nulli. Tale assurdo deriva dall'aver supposto $W \subset V$. Quindi $f$ è semplice.

La 2. $\Longrightarrow$ 3. Basta prendere in ciascun autospazio $V_{\lambda_i}$ una base di autovettori e ortonormalizzarla; la base ortonormale di autovettori si otterrà facendo l'unione delle suddette basi.

3. $\Longrightarrow$ 1. Sia $\mathcal{E} = (e_1, e_2, \cdots, e_n)$ la base ortonormale di autovettori. Per $i \neq j$ si ha $f(e_i) \bullet e_j = \lambda_i(e_i \bullet e_j) = 0 = e_i \bullet f(e_j) = \lambda_j(e_i \bullet e_j)$ e quindi $f$ è autoaggiunto. $\square$

Nel caso complesso il teorema spettrale si può enunciare come segue:

**Theorema 4.2.24** (**Teorema spettrale complesso**). *Sia $V$ un $\mathbb{C}$- spazio vettoriale con prodotto scalare complesso e sia $f : V \to V$ un endomorfismo. Allora se $f$ è autoaggiunto esiste in $V$ una base ortonormale di autovettori.*

*Proof.* Se $f$ è autoaggiunto esiste una base ortonormale di autovettori. Nel caso complesso abbiamo certamente autovalori. E siano $\lambda_1, \lambda_2 \cdots, \lambda_k$ quelli distinti e $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_k}$ i corrispondenti autospazi. Sia $W = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$. Per provare che esiste una base di autovettori basta provare $W = V$. Supponiamo che $W \subset V$, cioè sia contenuto propriamente in $V$. Allora l' ortogonale $W^\perp \neq (0_V)$. È ovvio che la restrizione di $f$ a $W$ è un endomorfismo su $W$. Ed è facile provare che $f(W^\perp) \subseteq W^\perp$. Infatti se $w' \in W^\perp$ e $w \in W$ allora $f(w') \bullet w = w' \bullet f(w) = w' \bullet \overline{w} = 0$. E quindi $f : W^\perp \to W^\perp$ è un endomorfismo. Deve ammettere un autovalore e un corrispondente autovettore $v^* \in W^\perp$. Allora $v^* \in W \cap W^\perp = (0_V)$. Ma gli autovettori sono non nulli e quindi abbiamo una contraddizione. Questa deriva dall' avere ammesso che $W^\perp \subset V$. Quindi possiamo "coprire" $V$ con autovettori che costituiscono una base per $V$. Alla fine tale base può essere ortonormalizzata col procedimento di Gram-Schmidt. $\square$

**Osservazione** 4.2.25. Se $V$ è uno spazio vettoriale complesso con prodotto scalare ed $f : V \to V$ un endomorfismo. Allora esiste in $V$ una base ortonormale di autovettori se e solo se $f$ è **normale**. Ricordiamo che $f$ è normale se commuta con l'endomorfismo aggiunto, cioè $ff^* = f^*f$. Approfondiremo questo concetto in seguito.

**Theorema 4.2.26 (Diagonalizzazione delle matrici simmetriche reali).** *Sia $A \in \mathbb{R}^{n,n}$ una matrice simmetrica reale. Allora esiste una matrice ortogonale $P$ che diagonalizza $A$, cioè tale che $P^{-1}AP = D$.*

*Proof.* Sia $f : \mathbb{R}^n \to \mathbb{R}^n$ l'endomorfismo associato ad $A$ mediante la base canonica $E$. Visto che $E$ è ortonormale rispetto al prodotto scalare euclideo ed $f$ è associato ad una matrice simmetrica reale, $f$ è autoaggiunto. In base al teorema spettrale esiste una base ortonormale $F$ di autovettori. Detta $P$ la matrice del cambio di base dalla base canonica $E$ a questa base $F$ la matrice $P$ deve essere ortogonale perché trasforma una base ortonormale in una base ortonormale. Detta $D$ la matrice associata ad $f$ rispetto a $F$ per la similitudine si ha $P^{-1}AP = D$. $\square$

Nel caso complesso il precedente teorema si può enunciare come segue:

**Theorema 4.2.27.** *Ogni matrice Hermitiana $A \in \mathbb{C}^{n,n}$ si può diagonalizzare mediante una matrice unitaria.*

## 4.3 Forme quadratiche

**Definizione 4.3.1.** Si dice **forma quadratica** una funzione $q : \mathbb{R}^n \to \mathbb{R}$ definita da un polinomio omogeneo di 2° grado in $n$ variabili. Detto $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ in generale si ha

$$q(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \cdots + 2a_{n-1,n}x_{n-1}x_n + a_{nn}x_n^2 \tag{4.15}$$

Si definisce anche la **matrice della forma quadratica**

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

come la matrice simmetrica avente i coefficienti $a_{ii}$ di $x_i^2$ come elementi della diagonale principale, mentre nei posti $(i, j)$ si mettono la metà dei coefficienti di $x_{ij}$. Ovviamente $a_{ij} = a_{ji}$.

Ponendo $\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ la forma quadratica (4.15) si può anche scrivere nella forma matriciale: $q(\underline{X}) = {}^t\underline{X}A\underline{X}$.

Un **cambiamento lineare di coordinate** è una trasformazione del tipo $\underline{X} = P\underline{Y}$, dove $P$ è una matrice invertibile. Se $q(\underline{X}) = {}^t\underline{X}A\underline{X}$, cambiando le variabili si ha $q'(\underline{Y}) = q(P\underline{Y}) = {}^t(P\underline{Y})A(P\underline{Y}) = {}^t\underline{Y}({}^tPAP)\underline{Y} = {}^t\underline{Y}B\underline{Y}$. Abbiamo posto $B = {}^tPAP$. Si vede subito che ${}^tB = {}^tP\ {}^tAP = B$. Cioè $B$ è una matrice simmetrica. Quindi $q'(\underline{Y})$ è una forma quadratica. Inoltre due forme quadratiche ottenute l'una per un cambiamento di coordinate dall'altra si dicono **equivalenti**.

**Proposizione 4.3.2 (Segno di una forma quadratica).** *Una forma quadratica $q(\underline{X})$ è detta*

- *definita positiva se* $\quad q(\underline{X}) > 0 \quad \forall \underline{X} \neq 0$
- *semidefinita positiva se* $\quad q(\underline{X}) \geq 0 \quad \forall \underline{X} \in \mathbb{R}^n$
- *definita negativa se* $\quad q(\underline{X}) < 0 \quad \forall \underline{X} \neq 0$
- *semidefinita negativa se* $\quad q(\underline{X}) \leq 0 \quad \forall \underline{X} \in \mathbb{R}^n$
- *non definita in ogni altro caso.*

**Proposizione 4.3.3.** *Sia $q(\underline{X}) = {}^t\underline{X}A\underline{X}$ una forma quadratica e $q'(\underline{Y})$ una forma quadratica equivalente. Sia $A = (a_{ij})$ la matrice associata. Sussistono le seguenti proprietà:*

*i)* *La forma* $q'(\underline{Y})$ *ha lo stesso segno di* $q(\underline{X})$*;*

*ii)* *se accade che* $a_{ii} < 0$ *e* $a_{jj} > 0$ *allora la forma* $q(\underline{X})$ *è non definita;*

*iii)* *se* $q(\underline{X})$ *è semidefinita positiva allora:* $a_{ii} \geq 0 \quad \forall i$*;*

*iv)* *se* $q(\underline{X})$ *è semidefinita negativa allora:* $a_{ii} \leq 0 \quad \forall i$*;*

*v)* *se* $q(\underline{X})$ *è definita positiva allora:* $a_{ii} > 0 \quad \forall i$*;*

*vi)* *se* $q(\underline{X})$ *è definita negativa allora:* $a_{ii} < 0 \quad \forall i$*;*

*Proof.* Per rispondere alla *i)* detta $\underline{X} = P\underline{Y}$ la trasformazione lineare invertibile che lega le due forme quadratiche equivalenti $q(\underline{X})$ e $q'(\underline{Y})$, dall'essere: $q'(\underline{Y}) = q(\underline{X}) = q(P\underline{Y})$ e $q(\underline{X}) = q'(P^{-1}\underline{Y})$ si deduce che al variare di $\underline{X}$ e $\underline{Y}$, $q(\underline{X})$ e $q'(\underline{Y})$ assumono gli stessi valori. Quindi hanno lo stesso segno.
Per rispondere alle altre questioni basta osservare che detti $\underline{X}_1 = (1, 0, \cdots, 0)$; $\underline{X}_2 = (0, 1, 0, \cdots, 0); \cdots; \underline{X}_n = (0, 0, \cdots, 1)$ si ha $q(\underline{X}_i) = a_{ii}$. Dopodiché basta applicare le definizioni. □

**Definizione 4.3.4.** Una forma quadratica si dice **in forma canonica** se la matrice associata è diagonale.

**Proposizione 4.3.5.** *Ogni forma quadratica* $q(\underline{X}) = {}^t\underline{X}A\underline{X}$ *ammette una forma equivalente del tipo* $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$.

*Proof.* Come sappiamo la matrice $A$ è simmetrica reale e quindi si può diagonalizzare mediante mediante una matrice $P$ ortogonale. Si ha quindi $A' = P^{-1}AP$, dove $A'$ è diagonale e sulla diagonale ha gli autovalori, contati con la loro molteplicità. Mediante la trasformazione $\underline{X} = P\underline{Y}$ la forma quadratica $q$ viene trasformata in una forma equivalente $q'(\underline{Y}) = {}^t\underline{Y}B\underline{Y}$, associata alla matrice $B = {}^tPAP$. Ma poiché la matrice $P$ è ortogonale si ha che ${}^tP = P^{-1}$. Quindi $A' = {}^tPAP$ e la forma quadratica è $q'(\underline{Y}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$. □

Dalla Proposizione precedente si deduce che se si deve studiare il segno di una forma quadratica $q(\underline{X})$ si deve trovare il polinomio caratteristico della matrice associata il quale avrà radici reali perché la matrice è simmetrica. Detto $P$ il numero delle radici positive, $N$ il nomero delle radici negative, e $Z$ il numero delle radici nulle si ha che

$$q(\underline{X}) \begin{cases} \text{definita positiva}, & \text{se } P = n \\ \text{definita negativa}, & \text{se } N = n \\ \text{semidefinita positiva}, & \text{se } N = 0 \\ \text{semidefinita negativa}, & \text{se } P = 0 \\ \text{non definita}, & \text{se } P > 0 \ \text{ e } N < 0 \end{cases}$$

È importante osservare che non necessariamente si devono calcolare le radici del polinomio caratteristico ma bisogna conoscere il segno di dette radici. Ci viene in aiuto la cosiddetta Regola di Cartesio, che dice:

**Proposizione 4.3.6.** *Se* $f(\underline{X}) = a_0 + a_1\underline{X} + a_2\underline{X}^2 + \cdots + a_n\underline{X}^n$ *è un polinomio a coefficienti reali avente tutte le radici reali, quelle positive sono tante quante le variazioni di segno consecutive nella sequenza* $(a_0, a_1, a_2, \cdots, a_n)$*.*

*Esempio* 4.3.7. Sia $q(x, y, z) = 4x^2 + 4xy + 4xz + 4y^2 + 4yz + 4z^2$ una forma quadratica reale.

1. Stabilire il segno della forma quadratica.

2. Ridurre $q$ a forma canonica.

3. Diagonalizzare la matrice associata mediante una matrice ortogonale $P$.

*Proof.* Innanzitutto bisogna considerare la matrice simmetrica associata alla forma. Essa è la matrice

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

l'equazione caratteristica è data dal determinante $|A - TI|$ della matrice caratteristica eguagliato a zero

$$\begin{vmatrix} 4-T & 2 & 2 \\ 2 & 4-T & 2 \\ 2 & 2 & 4-T \end{vmatrix} = 0$$

Usando le proprietà dei determinti è facile ottenere $(T-2)^2(T-8) = 0$. Quindi gli autovalori sono $\lambda_1 = \lambda_2 = 2$ e $\lambda_3 = 8$. Una forma canonica equivalente è $q(x', y', z') = 2x'^2 + 2y'^2 + 8z'^2$. Essa è definita positiva e la matrice diagonale ad essa associata è $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$. Adesso bisogna trovare gli autospazi associati $V_2$ e $V_8$. Poniamo 2 al posto di $T$ nella matrice caratteristica. Si ottiene $x + y + z = 0$, da cui $x = -y - z$ e quindi si hanno i due autovettori $u_1 = (1, -1, 0)$ e $u_2 = (1, 0, -1)$. Questi due autovettori non sono ortogonali fra loro. Possiamo ortonormalizzare queste due autovettori col procedimento di Gram-Schmidt.

Innanzitutto normalizziamo $u_1$ dividendo per la sua norma $\|u_1\|$. Si ottiene il vettore $e_1 = \frac{u_1}{\|u_1\|} = \frac{(1,-1,0)}{\sqrt{2}}$. Adesso dobbiamo costruire il versore $e_2 = \frac{u_2 - (u_2 \bullet e_1)e_1}{\text{norma del numeratore}}$. Si ottiene dopo semplici, ma attenti calcoli, $e_2 = \frac{(1,0,-1) - \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)}{\text{norma del numeratore}}$; sviluppando i calcoli a numeratore si ha $(\frac{1}{2}, \frac{1}{2}, -1)$. La norma di questo vettore è $\sqrt{\frac{1}{4} + \frac{1}{4} + 1}$. In ultimo $e_2 = \frac{(1,1,-2)}{\sqrt{6}}$.
Adesso dobbiamo calcolare l'autospazio $V_8$. Sostituiamo 8 al posto di $T$ nella matrice caratteristica. Si ha $\begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$. Risolvendo il sisstema lineare omogeneo costruito su questa matrice si ha che le soluzioni del sistema lineare omogeneo costruito su questa matrice fornisce la soluzione $(x, x, x)$. Quindi

possiamo prendere come autovettore $u_3 = (1,1,1)$. Normalizzandolo si ha $e_3 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. È facile verificare che il versore $e_3$ è ortogonale a $e_1$ ed $e_2$.

In conclusione la matrice ortogonale che diagonalizza $A$ è la matrice

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

□

**Theorema 4.3.8 (Sylvester).** *Una forma quadratica associata ad una matrice simmetrica reale $A$ è definita positiva se e solo se tutti i minori principali sono positivi.*

Ricordiamo che si chiama minore principale, di ordine $r$, di una matrice il determinante della sottomatrice formata dagli elementi di posto $(i,j)$, con $1 \le i \le r$ e $1 \le j \le r$.

È interessante il seguente **metodo alternativo di diagonalizzazione di una matrice simmetrica**

**Osservazione** 4.3.9. È facile convincersi che quando in una matrice quadrata $A \in \mathbb{R}^{n,n}$ si fa una riduzione per righe, si ottiene lo stesso risultato se si moltiplica a sinistra $A$ per la matrice identica $I_n$ a cui è stata fatta la riduzione che si vuole effettuare in $A$. È evidente che questo discorso si può approfondire. Analogamente se si vogliono effettuare riduzioni per colonne basta moltiplicare a destra la matrice identica a cui sono state fatte le riduzioni volute.

Facciamo vedere come si può **diagonalizzare una matrice simmetrica reale** senza bisogno di calcolare autovalori e autovettori. Consideriamo un esempio molto semplice, ma molto istruttivo. Sia $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ una matrice simmetrica reale di ordine 2, con $a \neq 0$.

$$\frac{\begin{vmatrix} a & b \\ b & c \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}} \xrightarrow{R_2 \leftrightarrow R_2 - \frac{b}{a} R_1} \frac{\begin{vmatrix} a & b \\ 0 & c - \frac{b^2}{a} \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}} \xrightarrow{C_2 \leftrightarrow C_2 - \frac{b}{a} C_1} \frac{\begin{vmatrix} a & 0 \\ 0 & c - \frac{b^2}{a} \end{vmatrix}}{\begin{vmatrix} 1 & -\frac{b}{a} \\ 0 & 1 \end{vmatrix}}$$

Si sarà notato che dopo aver effettuato la prima riduzione per righe, si è poi effettuata la stessa riduzione sulle colonne, continuando fino alla completa diagonalizzazione della matrice. Quando si conclude l'algoritmo l'ultima matrice a denominatore ci fornisce la matrice $P$ del cambio di base. Nel nostro caso la matrice $P = \begin{pmatrix} 1 & -\frac{b}{a} \\ 0 & 1 \end{pmatrix}$ è la matrice del cambio di base dalla base canonica alla nuova base $\mathcal{F} = [(1,0); (-\frac{b}{a})]$. La trasposta $P^T$moltiplicata a sinistra per

$A$ permette la riduzione per righe voluta. In definitiva si ottiene

$$P^T \cdot A \cdot P = D = \begin{pmatrix} a & 0 \\ 0 & c - \frac{b^2}{a} \end{pmatrix}$$

L'algoritmo che abbiamo usato permette di diagonalizzare una matrice simmetrica reale e ci fornisce anche la matrice $P$ del cambio di coordinate.

**Bisogna fare però molta attenzione**. Il tipo di diagonalizzazione ottenuto è diverso da quello ottenuto nei numeri precedenti. Infatti la matrice ottenuta non è simile a quella di partenza.

C'è però da dire che le due matrici hanno lo stesso rango e lo stesso determinante. Nel nostro caso $a \neq 0$, $\det A = \det D = ac - b^2$. Quindi possiamo dire che se $a > 0$ e $\det A > 0$ la forma quadratica associata ad $A$ è definita positiva.

L' algoritmo usato permette di dimostrare in generale il criterio di Sylvester.