Appunti di Algebra Lineare

Roberto Catenacci

Versione del: 7 Marzo 2011

Indice

1 Richiami di algebra.

1.1 Equazioni quadratiche.

Una equazione quadratica in x è una equazione della forma: $ax^2 + bx + c = 0$ dove a, b, c sono numeri reali e $a \neq 0$. Non è difficile risolverla; quando $\Delta = b^2 - 4ac \geq 0$ (la stessa formula vale anche se $\Delta < 0$, ma ne parleremo dopo) si ha:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Ci sono due tipi di equazioni quadratiche; quelle che hanno soluzioni *razionali* (ovvero che si possono scrivere nella forma m/n con m e n numeri interi relativi e $n \neq 0$) e quelle le cui soluzioni sono *irrazionali.* Le soluzioni sono razionali quando Δ è il quadrato di un numero razionale. Esempi di numeri irrazionali sono $\sqrt{2}, \sqrt{3}, \pi$.

Dimostriamo che $\sqrt{2}$ è irrazionale. Supponiamo che sia razionale; allora $\sqrt{2} = m/n$, *con m e n senza fattori comuni*; elevando al quadrato si ha: $m^2 = 2n^2$. Questo significa che m^2 è un numero pari e quindi anche m lo è (questa ultima osservazione deriva dal fatto che il quadrato di un numero dispari è un numero dispari, infatti $(2k + 1)^2 = 4k^2 + 4k + 1$; ora usiamo anche il seguente fatto: se la proposizione A implica la proposizione B, allora la proposizione non-B implica la proposizione non-A. (m dispari implica m^2 dispari allora m^2 non-dispari cioè pari implica m non-dispari cioè pari). Allora $m = 2k$ e quindi $4k^2 = 2n^2$ cioè $n^2 = 2k^2$ così n^2 è pari e anche n lo è. Allora sia m che n sono numeri pari e quindi hanno in comune almeno il fattore 2.

Questa è una contraddizione perchè abbiamo supposto che *m e n siano senza fattori comuni.* La contraddizione si risolve *solo negando l'assunto* che $\sqrt{2}$ sia razionale. Il ragionamento precedente è un esempio di **dimostrazione per assurdo**, il metodo consiste nell'assumere falsa la proposizione che si vuole dimostrare e cercare di ottenere una contraddizione. Questo risultato si generalizza abbastanza facilmente: **le radici n-esime dei numeri interi sono numeri irrazionali a meno che il numero non sia una n-esima potenza.** Dimostreremo questo teorema (che useremo nel corso di Algebra) più sotto, dopo aver discusso alcune proprietà dei polinomi.

La formula risolutiva delle equazioni di secondo grado data sopra implica che una tale equazione ha, al più, due soluzioni. Se $\Delta = b^2 - 4ac = 0$ l'equazione ha una sola soluzione *(meglio dire due soluzioni coincidenti)*, se $\Delta = b^2-4ac > 0$ l'equazione ha due soluzioni distinte. Se $\Delta = b^2 - 4ac < 0$ l'equazione non ha soluzioni reali; introducendo i **numeri complessi** si possono estrarre le radici quadrate dei numeri negativi e si ripristina la simmetria: una equazione di grado *due* ha sempre *due* soluzioni (distinte o coincidenti).

Due equazioni $a_1x^2 + b_1x + c_1 = 0$ e $a_2x^2 + b_2x + c_2 = 0$ hanno le stesse soluzioni se e solo se sono proporzionali, cioè se esiste una costante $k \neq 0$ tale che

$$ka_1 = a_2, kb_1 = b_2, kc_1 = c_2.$$

Questa osservazione è utile perchè ci consente di scrivere subito, senza risolvere l'equazione, la somma e il prodotto delle sue soluzioni: siano α e β le soluzioni dell'equazione $ax^2+bx+c = 0$.

Anche l'equazione:

$$a(x-\alpha)(x-\beta) = ax^2 - a(\alpha+\beta)x + a\alpha\beta = 0$$

ha ovviamente le stesse soluzioni, e quindi, confrontando con l'equazione originaria, si trova:

$$\alpha+\beta = -b/a \text{ e } \alpha\beta = c/a.$$

Esempio: data l'equazione $8x^2-9x+2=0$, trovare una equazione che abbia come soluzioni i quadrati delle soluzioni dell'equazione data. Siano α e β le soluzioni: si ha $\alpha+\beta = 9/8$ e $\alpha\beta = 1/4$. Notiamo anche che $(\alpha^2+\beta^2) = (\alpha+\beta)^2 - 2\alpha\beta = 49/64$ e $\alpha^2\beta^2 = 1/16$. Quindi l'equazione richiesta, *trovata senza risolvere l'equazione di partenza*, è $x^2-(\alpha^2+\beta^2)x+\alpha^2\beta^2 = x^2 - \frac{49}{64}x + \frac{1}{16} = 0$.

Questo metodo funziona per trovare equazioni le cui radici siano **funzioni simmetriche** (una funzione f di α e β è simmetrica se $f(\alpha,\beta) = f(\beta,\alpha)$) delle radici dell'equazione data; infatti **tutte** le funzioni simmetriche si possono scrivere in funzione delle due funzioni simmetriche elementari $f_1(\alpha,\beta) = \alpha+\beta$ e $f_2(\alpha,\beta) = \alpha\beta$.

Esempio: $\alpha^3+\beta^3 = (\alpha+\beta)^3 - 3\alpha\beta(\alpha+\beta)$.

1.2 Numeri complessi.

Per introdurre i numeri complessi si parte di solito dal fatto che l'equazione $x^2+1=0$ non ha soluzioni nel campo dei numeri reali. Si *inventa* allora una soluzione dell'equazione data e la si chiama i. Si tratta poi i come un normale numero a cui si applicano le solite regole di calcolo algebrico e allora si ottiene $i^2 = -1$. Si scrivono poi tutti i simboli del tipo $a+ib$ con a e b numeri reali e si postulano per essi le usuali regole di calcolo algebrico: ecco i **numeri complessi**.

Un modo più *convincente* per introdurre i numeri complessi è quello di farli discendere direttamente dalle coppie di numeri reali.

Indichiamo con $\mathbb{R}$ il campo dei numeri reali con le solite operazioni e con $\mathbb{C}$ il campo dei numeri complessi.

Un **numero complesso** è una coppia ordinata di numeri reali $(a,b) \in \mathbb{R}\times\mathbb{R}$, cioè

$$\mathbb{C} = \{(a,b) : a,b \in \mathbb{R}\}.$$

Definiamo $\forall\,(a,b)\,,(c,d) \in \mathbb{C}$ le seguenti operazioni :

i) **somma**: $(a,b)+(c,d) = (a+b, c+d)$;

ii) **prodotto**: $(a,b)\cdot(c,d) = (ac-bd, ad+bc)$.

Viene naturale **identificare** $a \in \mathbb{R}$ con la coppia $(a,0) \in \mathbb{C}$, cioè ogni numero reale può essere pensato come numero complesso $\Rightarrow$ $\mathbb{C}$ è un "ampliamento" di $\mathbb{R}$, cioè $\mathbb{R} \subset \mathbb{C}$.

L'elemento $(0,1) \in \mathbb{C}$, si denota con i ed è detto **unità immaginaria**. Si ha

$$i^2 = i\cdot i = (0,1)\cdot(0,1) = (-1,0) = -1.$$

Prendiamo $(a,b) \in \mathbb{C}$:

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1) \cdot (b, 0) = a + ib,$$

cioè ogni numero complesso può essere scritto nella forma $a + ib$, detta **forma algebrica**, a è chiamata **parte reale**, ib è la **parte immaginaria**. Le operazioni tra numeri complessi scritti in forma algebrica si eseguono con le regole del calcolo letterale:

$$\begin{aligned}(a + ib) + (c + id) &= a + ib + c + id = (a + c) + i(b + d);\\ (a + ib) \cdot (c + id) &= ac + iad + ibc + i^2bd = ac + iad + ibc - bd\\ &= (ac - bd) + i(ad + bc).\end{aligned}$$

Dato $z = a + ib$, si definisce il numero **complesso coniugato**

$$\overline{z} = a - ib.$$

I numeri complessi $\neq 0$ (ovvero quelli per cui $a^2 + b^2 \neq 0$) hanno un inverso:

$$\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{a - ib}{a^2 + b^2}$$

Infatti

$$z \cdot \frac{1}{z} = (a + ib) \cdot \frac{a - ib}{a^2 + b^2} = 1$$

Per mettere in forma algebrica una frazione con un numero complesso a denominatore, si può moltiplicare numeratore e denominatore per il coniugato del numero, osservando che $z\overline{z} = (a + ib) \cdot (a - ib) = a^2 - iab + iab - i^2b = a^2 + b^2$.

Rappresentazione geometrica dei numeri complessi

Fissiamo nel piano un sistema di riferimento cartesiano Oxy. Un numero complesso (a, b) può essere pensato come punto del piano π (**piano complesso**): a rappresenta l'ascissa e b l'ordinata.

Osservazioni.

1) Se $z \in \mathbb{R} \Rightarrow z = (a, 0) \Rightarrow z$ sta sull'asse x (detto **asse reale**).

2) Se $z = ib$, cioè z è immaginario puro $\Rightarrow z = (0, b)$ sta sull'asse y (detto **asse immaginario**).

3) La moltiplicazione per i equivale quindi a una rotazione di 90° in senso antiorario.

- **Forma trigonometrica di un numero complesso:**

sia $z = a + ib$ un numero complesso e sia P la sua immagine sul piano π. Associamo a z due "coordinate polari" (ρ, θ):

1) $\rho = \sqrt{a^2 + b^2}$ è detto **modulo** di z ed è la lunghezza del segmento che congiunge l'origine con il punto P;

2) θ è detto **anomalia** o **argomento** di z, è l'angolo tra ρ e il semiasse positivo delle x ed **è individuato a meno di multipli di 2π.**

Dalla trigonometria si ha che

$$a = \rho \cos\theta \text{ e } b = \rho \sin\theta$$

Si ottiene quindi:

$$\cos\theta = \frac{a}{\rho} = \frac{a}{\sqrt{a^2+b^2}} \text{ e } \sin\theta = \frac{b}{\rho} = \frac{b}{\sqrt{a^2+b^2}}.$$

Dalle formule precedenti si ottiene:

$$\boxed{a + ib = \rho\,(\cos\theta + i\sin\theta)},$$

detta **forma trigonometrica** del numero complesso $a+ib$.

Due numeri complessi non nulli $z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$ e $z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$ sono uguali se $\rho_1 = \rho_2$ e $\theta_1 = \theta_2 + 2k\pi$, con k intero.

- **Teorema**. Dati $z_1, z_2 \in \mathbb{C}$, il loro prodotto è

$$z_1 \cdot z_2 = \rho_1\rho_2\,[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)].$$

- Tale regola vale anche per un numero qualsiasi di fattori. Se i fattori sono uguali, si ha la **formula di De Moivre**:

$$[\rho\,(\cos\theta + i\sin\theta)]^n = \rho^n\,(\cos n\theta + i\sin n\theta)\,, \quad n \in \mathbb{N}.$$

Esempio:

$$(1+i)^{14} = (\sqrt{2}(\cos\pi/4 + i\sin\pi/4))^{14} = 128(\cos(-\pi/2) + i\sin(-\pi/2)) = -128i$$

infatti $14\pi/4 = 7\pi/2 = 4\pi - \pi/2$ (provate solo a pensare di dover fare il calcolo senza la formula di De Moivre per capirne l'importanza !).

Se si conviene di porre $[\rho\,(\cos\theta + i\sin\theta)]^0 = 1$, e, per $\rho \neq 0$,

$$[\rho\,(\cos\theta + i\sin\theta)]^{-n} = \frac{1}{[\rho\,(\cos\theta + i\sin\theta)]^n} = \rho^{-n}\,[(\cos\theta - i\sin\theta)]^n = \left[\rho^{-n}\,(\cos(-n\theta) + i\sin(-n\theta)\right]$$

quindi la formula di De Moivre vale per $n \in \mathbb{Z}$.

- **Radici n-esime dei numeri complessi.**

Sia $z \in \mathbb{C}$ e $n \in \mathbb{N}$. I numeri $w \in \mathbb{C}$ tali che $w^n = z$, si dicono **radici n-esime** di z.

Prendiamo $z = \rho\,(\cos\theta + i\sin\theta)$ e sia $w = r(\cos\varphi + i\sin\varphi)$ una radice n-esima di z. Si deve avere

$$\begin{gathered} w^n = z \Rightarrow r^n(\cos n\varphi + i\sin n\varphi) = \rho\,(\cos\theta + i\sin\theta) \\ \Rightarrow r^n = \rho \text{ e } n\varphi = \theta + 2k\pi,\ k \in \mathbb{Z} \\ \Rightarrow r = \sqrt[n]{\rho} \text{ (radice aritmetica) e } \varphi = \frac{\theta}{n} + \frac{2k\pi}{n},\ k \in \mathbb{Z}. \end{gathered}$$

L'ultima formula dà infiniti valori per φ, ma solo n sono distinti ($k = 0, 1, ..., n-1$ poi si ripetono a meno di multipli di 2π).

Teorema. $z = \rho\,(\cos\theta + i\sin\theta)$ ammette n radici n-esime w distinte date dalla formula

$$(\sharp) \qquad \boxed{w = \sqrt[n]{z}^{*} = \sqrt[n]{\rho}\left[\cos\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right) + i\sin\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right)\right]},$$

nella quale si pone $k = 0, 1, 2, ..., n-1$.

(L'asterisco serve solo a distinguere la radice complessa da quella aritmetica).

Gli n numeri forniti da ($\sharp$), hanno lo stesso modulo $\sqrt[n]{\rho}$, quindi le loro immagini sul piano complesso stanno su una circonferenza di centro l'origine e raggio $\sqrt[n]{\rho}$.

Scegliamo $z = 1 = 1 + 0i = 1(\cos 0 + i\sin 0)$. Utilizzando ($\sharp$), troviamo le n **radici n-esime dell'unità**:

$$\begin{aligned} \epsilon_k = \sqrt[n]{1}^{*} &= \sqrt[n]{1}\left[\cos\left(0 + \frac{2k\pi}{n}\right) + i\sin\left(0 + \frac{2k\pi}{n}\right)\right] \\ &= \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \text{ con } k = 0, 1, 2, ..., n-1. \end{aligned}$$

Le radici n-esime dell'unità si rappresentano geometricamente come vertici di un poligono regolare di n lati, inscritto in una circonferenza di raggio 1, con un vertice nel punto $(1, 0)$.

Le radici n-esime dell'unità servono anche per calcolare le radici n-esime di qualsiasi numero: basta moltiplicare una qualsiasi radice complessa del numero per le radici n-esime dell'unità. Infatti sia ζ una qualunque radice n-esima di z, ponendo $\zeta_k = \zeta\epsilon_k$ si ottengono tutte le radici di z, infatti:

$$(\zeta_k)^n = (\zeta\epsilon_k)^n = \zeta^n\epsilon_k^n = z$$

Esempi.

- Calcolare le tre radici cubiche complesse di 8.

Scriviamo 8 in forma trigonometrica:

$$\begin{gathered} 8 = 8 + 0i = 8(\cos 0 + i\sin 0). \\ w_k = \sqrt[3]{8}^{*} = \sqrt[3]{8}\left(\cos\frac{2k\pi}{3} + i\sin\frac{2k\pi}{3}\right),\ k = 0, 1, 2. \end{gathered}$$

Quindi

$w_0 = 2(\cos 0 + i\sin 0) = 2$;

$w_1 = 2\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right) = 2\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = -1 + i\sqrt{3}$;

$w_2 = 2\left(\cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}\right) = 2\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = -1 - i\sqrt{3}.$
Le tre radici cubiche di 8 sono: 2, $-1 + i\sqrt{3}$ e $-1 + i\sqrt{3}$.

- Calcolare le quattro radici quarte di $-1 + i\sqrt{3}$.

Si ha:

$$\rho = \sqrt{1+3} = 2,\ \cos\theta = -\frac{1}{2} \text{ e } \sin\theta = \frac{\sqrt{3}}{2} \Rightarrow \theta = \frac{2\pi}{3}$$
$$\Rightarrow -1 + i\sqrt{3} = 2\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right).$$

Applicando la ($\sharp$) con $k = 0, 1, 2, 3$, si ha

$$w_k = \sqrt[4]{-1 + i\sqrt{3}}^* = \sqrt[4]{2}\left[\cos\left(\frac{\pi}{6} + \frac{k\pi}{2}\right) + i\sin\left(\frac{\pi}{6} + \frac{k\pi}{2}\right)\right].$$

$$w_0 = \sqrt[4]{2}\left[\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right] = \sqrt[4]{2}\left(\frac{\sqrt{3}}{2} + i\frac{1}{2}\right) = \frac{\sqrt[4]{2}}{2}\left(\sqrt{3} + i\right);$$
$$w_1 = \sqrt[4]{2}\left[\cos\left(\frac{\pi}{6} + \frac{\pi}{2}\right) + i\sin\left(\frac{\pi}{6} + \frac{\pi}{2}\right)\right] = \sqrt[4]{2}\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right)$$
$$= \sqrt[4]{2}\left(-\frac{1}{2} + i\frac{\sqrt{3};}{2}\right) = \frac{\sqrt[4]{2}}{2}\left(-1 + i\sqrt{3}\right);$$
$$w_2 = \sqrt[4]{2}\left[\cos\left(\frac{\pi}{6} + \pi\right) + i\sin\left(\frac{\pi}{6} + \pi\right)\right] = \sqrt[4]{2}\left(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\right)$$
$$= \sqrt[4]{2}\left(-\frac{\sqrt{3}}{2} - i\frac{1}{2}\right) = \frac{\sqrt[4]{2}}{2}\left(-\sqrt{3} - i\right);$$
$$w_3 = \sqrt[4]{2}\left[\cos\left(\frac{\pi}{6} + \frac{3\pi}{2}\right) + i\sin\left(\frac{\pi}{6} + \frac{3\pi}{2}\right)\right] = \sqrt[4]{2}\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right)$$
$$= \sqrt[4]{2}\left(\frac{1}{2} - i\frac{\sqrt{3};}{2}\right) = \frac{\sqrt[4]{2}}{2}\left(1 - i\sqrt{3}\right).$$

- Calcolare le sei radici seste dell'unità.

Si ha:

$$\epsilon_k = \sqrt[6]{1}^* = \cos\frac{2k\pi}{6} + i\sin\frac{2k\pi}{6},\ \text{con } k = 0, 1, 2, 3, 4, 5.$$

Quindi
$\epsilon_0 = \cos 0 + i \sin 0 = 1;$
$\epsilon_1 = \cos \frac{\pi}{3} + i \sin \frac{\pi}{3} = \frac{1}{2} + i\frac{\sqrt{3}}{2};$
$\epsilon_2 = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2};$
$\epsilon_3 = \cos \pi + i \sin \pi = -1;$
$\epsilon_4 = \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2};$
$\epsilon_5 = \cos \frac{5\pi}{3} + i \sin \frac{5\pi}{3} = \frac{1}{2} - i\frac{\sqrt{3}}{2}.$
Le sei radici trovate sono i vertici di un esagono regolare.

- **Formula di Eulero.**

Consideriamo la funzione esponenziale di variabile complessa $f(z) = e^z$, $z \in \mathbb{C}$, definita formalmente con la stessa serie che si usa per definire l'esponenziale reale.
Valgono le consuete proprietà delle potenze:

$$e^z \cdot e^w = e^{z+w}; \qquad e^z/e^w = e^{z-w}; \qquad (e^z)^n = e^{nz},$$

con $z, w \in \mathbb{C}$, $n \in \mathbb{Z}$.
Se $z = x + iy$, con $x, y \in \mathbb{R}$, si trova, scrivendo esplicitamente le serie coinvolte,

$$e^z = e^{x+iy} = e^x \cdot e^{iy} = e^x (\cos y + i \sin y).$$

Quindi è

$$\boxed{e^{iy} = (\cos y + i \sin y)}.$$

Quest'ultima è la **formula di Eulero**.
Si ha

$$e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1 + 0i = 1$$

e

$$e^{z+2\pi i} = e^z \cdot e^{2\pi i} = e^z,$$

cioè l'esponenziale e^z è periodica di periodo $2\pi i$.
Inoltre

$$e^{-iy} = \cos(-y) + i \sin(-y) = \cos y - i \sin y.$$

Esempio.
$i = 0 + 1i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = e^{\frac{\pi}{2} i};$
$-1 = -1 + 0i = \cos \pi + i \sin \pi = e^{\pi}.$

1.3 Radici n - esime dell'unità.

Abbiamo visto che le radici n - esime dell'unità sono espresse dagli n numeri complessi:

$$\epsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \text{con } k = 0, 1, 2, ..., n-1.$$

E' facile verificare, utilizzando la definizione e la formula del prodotto di numeri complessi in forma trigonometrica:

$$z_1 \cdot z_2 = \rho_1\rho_2\left[\cos\left(\theta_1+\theta_2\right) + i\sin\left(\theta_1+\theta_2\right)\right]$$

che l'insieme delle radici n - esime dell'unità è un **gruppo abeliano**; il prodotto è dato esplicitamente da:

$$\epsilon_k \cdot \epsilon_h = \epsilon_j \text{ dove } j = (k+h) \bmod n$$

E' chiaro che l'elemento neutro del gruppo è $\epsilon_0 = 1$ e che l'inverso di ϵ_k è:

$$(\epsilon_k)^{-1} = \epsilon_{n-k}$$

Un altro risultato utile è:

$$(\epsilon_k)^{-1} = \overline{\epsilon_k}$$

Questo gruppo è **ciclico**, ovvero tutti i suoi elementi sono potenze positive di un qualche suo elemento, detto allora **generatore**. Si verifica subito infatti che, per ogni n :

$$\epsilon_k = (\epsilon_1)^k$$

Vogliamo ora capire, dato n, quali sono i generatori del gruppo delle radici n - esime (oltre ovviamente ϵ_1). Il punto di partenza è la seguente osservazione: sia ϵ_k una radice n - esima, se $MCD(n,k) = d$ allora ϵ_k è anche radice $\frac{n}{d}$ - esima. Infatti basta calcolare:

$$(\epsilon_k)^{\frac{n}{d}} = \left(\cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}\right)^{\frac{n}{d}} = \cos\left(\frac{2k\pi}{n}\frac{n}{d}\right) + i\sin\left(\frac{2k\pi}{n}\frac{n}{d}\right) = \cos\frac{2k\pi}{d} + i\sin\frac{2k\pi}{d} = 1$$

Perchè k è per ipotesi divisibile per d. Diamo ora la seguente definizione: ϵ_k è **radice primitiva** se è radice n - esima ma non radice m - esima con $1 \leq m < n$. Ovvero se $(\epsilon_k)^n = 1$, ma $(\epsilon_k)^m \neq 1$ per $1 \leq m < n$. Ovviamente ϵ_1 **è sempre primitiva**. Si vede subito che ϵ_k è primitiva se e solo se $MCD(n,k) = 1$. La dimostrazione procede per assurdo; se infatti $MCD(n,k) \neq 1$ l'osservazione precedente implica che ϵ_k non è primitiva, infatti $(\epsilon_k)^{\frac{n}{d}} = 1$ e $\frac{n}{d} < n$. Viceversa se ϵ_k non è primitiva, allora $(\epsilon_k)^m = 1$ per qualche m, $1 \leq m < n$. Ovvero, essendo ϵ_1 sempre primitiva, esiste un intero h tale che $km = hn$. Questo implica che $MCD(n,k) \neq 1$, perchè se fosse 1, per un teorema dimostrato in corsi precedenti, esisterebbero due interi x, y tali che $nx + ky = 1$ e quindi, moltiplicando per m si ha:

$$m = nmx + kmy = nmx + nhy = n(mx + hy)$$

e questo è assurdo perchè $1 \leq m < n$ implica che m non possa essere divisibile per n.

Esempio: consideriamo le radici seste dell'unità (vedi più sopra)**:** $MCD(6,k) = 1$ per $k = 1,5$ e infatti si vede subito che $\epsilon_3 = -1$ è anche radice quadrata, mentre ϵ_2 e ϵ_4 sono anche radici cubiche: $\left(-\frac{1}{2}+i\frac{\sqrt{3}}{2}\right)^3 = \left(-\frac{1}{2}-i\frac{\sqrt{3}}{2}\right)^3 = 1$.

Possiamo ora concludere facilmente che i generatori sono tutte e sole le radici primitive, infatti solo per le radici primitive le potenze distinte sono proprio n.

Si mostra anche un altro fatto interessante: **ogni radice n - esima è d - esima primitiva per ogni d divisore di n.** Sia infatti $(\epsilon_k)^n = 1$ e sia d il più piccolo intero positivo per cui $(\epsilon_k)^d = 1$ (questo d si dice **periodo** o **ordine** di ϵ_k) allora dividendo n per d si ha $n = qd + r$ (con $0 \leq r < d$) allora si avrebbe

$$1 = (\epsilon_k)^n = (\epsilon_k)^{qd}(\epsilon_k)^r = \left((\epsilon_k)^d\right)^q (\epsilon_k)^r = (\epsilon_k)^r$$

E quindi deve essere $r = 0$ perchè altrimenti d non sarebbe il più piccolo e quindi d deve essere divisore di n. Osserviamo anche che segue che se n è primo e quindi non ha divisori non banali, tutte le radici $\neq 1$ sono primitive. Viceversa è ovvio che se d divide n ogni radice d - esima è anche n - esima.

Concludiamo con altre interessanti proprietà delle radici n -esime dell'unità:

$$\sum_0^{n-1}\epsilon_k = \sum_1^n \epsilon_k = 0$$

Infatti si ha, posto $\sum_0^{n-1}\epsilon_k = S$, $\epsilon_1 S = S$ (osservando che $\epsilon_1\epsilon_{n-1} = \epsilon_n = \epsilon_0$) e quindi $S = 0$.

$$\prod_0^{n-1}\epsilon_k = \prod_1^n \epsilon_k = 1 \text{ se } n \text{ è dispari e } -1 \text{ se } n \text{ è pari}$$

Infatti

$$\prod_1^n \epsilon_k = \prod_1^n (\epsilon_1)^k = (\epsilon_1)(\epsilon_1)^2 ... (\epsilon_1)^n = (\epsilon_1)^{\frac{n(n+1)}{2}}$$

Allora se n è dispari, $n = 2m+1$, e quindi $\frac{n(n+1)}{2} = nm$ e allora

$$(\epsilon_1)^{nm} = ((\epsilon_1)^n)^m = 1,$$

mentre se n è pari, $n = 2m$, e quindi $\frac{n(n+1)}{2} = m(n+1)$ e allora

$$(\epsilon_1)^{m(n+1)} = (\epsilon_1)^{nm}(\epsilon_1)^m = (\epsilon_1)^m = \cos\frac{2m\pi}{n} + i\sin\frac{2m\pi}{n} = \cos\frac{2m\pi}{2m} + i\sin\frac{2m\pi}{2m} = -1$$

1.4 Polinomi.

Un **polinomio** a coefficienti reali (o complessi) nella indeterminata x è un'espressione

$$P(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n$$

con $a_0, a_1, ..., a_n \in \mathbb{R}$ (o $\mathbb{C}$), $a_n \neq 0$, detti **coefficienti** del polinomio e $n \in \mathbb{N}$, detto **grado** del polinomio: $n = \deg P(x)$. Se $a_n = 1$, il polinomio è detto **monico**. L'insieme dei polinomi in x su $\mathbb{R}$ (o $\mathbb{C}$) è denotato con $\mathbb{R}[x]$ (o $\mathbb{C}[x]$). Il **polinomio nullo**, è per definizione, il numero 0 (ad esso, secondo la nostra definizione, non si potrebbe attribuire un grado, perchè non c'è nessun n per cui a_n è diverso da zero; si conviene, per convenzione, attribuirgli qualsiasi grado vogliamo). Un **polinomio costante** è il polinomio nullo oppure un polinomio di grado 0.

Due polinomi in x sono uguali se hanno uguali i coefficienti delle potenze di x corrispondenti. I polinomi si sommano e si moltiplicano secondo le regole note del calcolo letterale. Si usa denotare con $\mathbb{R}_{\leq n}[x]$ (o $\mathbb{C}_{\leq n}[x]$) l'insieme dei polinomi di grado al più n, includendo anche il polinomio nullo.

- **Algoritmo della divisione.**

Dati due polinomi $A(x)$, $B(x) \in \mathbb{R}[x]$, sono determinati in modo unico i polinomi $Q(x)$ (**quoziente**) e $R(x)$ (**resto**) in $\mathbb{R}[x]$, tali che

$$A(x) = B(x)Q(x) + R(x)$$

con $R(x) = 0$ oppure $\deg R(x) < \deg B(x)$.
Se $\deg A(x) = n$, $\deg B(x) = m$:
i) se $m \leqslant n \Rightarrow \deg Q(x) = n - m$ e $\deg R(x) < m$;
ii) se $m > n \Rightarrow Q(x) = 0$ e $R(x) = A(x)$.
Se nel polinomio $A(x) \in \mathbb{R}[x]$, x è sostituita da un $c \in \mathbb{R}$, il risultato è un elemento di $\mathbb{R}$ (e lo stesso vale se invece di $\mathbb{R}$ lavoriamo in $\mathbb{C}$).
Esempio.
$A(x) = x^3 - 2x^2 + 2 \in \mathbb{R}[x]$
$A(3) = 27 - 18 + 2 = 11 \in \mathbb{R}$.

- **Teorema del resto**. Siano $A(x) \in \mathbb{R}[x]$ e $B(x) = x - c$, con $c \in \mathbb{R}$. Il resto della divisione di $A(x)$ per $B(x)$ è $A(c)$.

Infatti $A(x) = (x - c)Q(x) + R(x)$ e ponendo $x = c$ si ottiene $A(c) = R(c) = R(x)$ perchè il grado del resto deve essere, in questo caso, zero (il grado del resto è minore del grado del polinomio divisore) e quindi $R(x)$ è un polinomio costante.
Esempio.
$A(x) = x^3 - 2x^2 + 2$, $B(x) = x - 3$.
$R(x) = A(3) = 11$.
Il risultato si può facilmente verificare eseguendo la divisione tra i due polinomi.
Dati $A(x), B(x) \in \mathbb{R}[x]$ con $B(x) \neq 0$, diciamo che $A(x)$ è **divisibile** per $B(x)$ se $A(x) = B(x)Q(x)$, cioè se il resto della divisione è 0.
$c \in \mathbb{R}$ si dice **radice** (o **zero**) di un polinomio $A(x) \in \mathbb{R}[x]$ se $A(c) = 0$.

- **Teorema di Ruffini**. Se $A(x) \in \mathbb{R}[x]$ e $c \in \mathbb{R}$, allora c è radice di $A(x) \Leftrightarrow A(x)$ è divisibile per $(x - c)$.

- **Teorema fondamentale dell'algebra**. Ogni polinomio $P(x) \in \mathbb{C}[x]$ di grado $n \geqslant 1$, ha sempre n radici in $\mathbb{C}$.

Un polinomio a coefficienti reali non è detto che abbia sempre radici reali, ma chiaramente se z è una radice, anche $\overline{z}$ lo è. Ne consegue che un polinomio a coefficienti **reali** di grado **dispari** deve avere sempre almeno una radice reale. Questo fatto deriva anche da un noto teorema di analisi (Teorema degli zeri di Bolzano): un polinomio a coefficienti reali è una funzione reale continua e, se ha grado dispari, per $x \to \pm\infty$ tende all'infinito con valori di segno discordi; deve quindi passare per zero.

Se il polinomio è **monico** e le radici sono $x_1, ..., x_n$ allora possiamo scomporre il polinomio nella forma

$$P(x) = (x - x_1)(x - x_2)...(x - x_n).$$

Se una radice compare k volte nella scomposizione, diciamo che ha **molteplicità algebrica** k. In questo caso essa è radice non solo del polinomio, ma anche delle sue derivate fino all'ordine $k - 1$.

Se il polinomio non è monico basta riscriverlo così:

$$a_0 + a_1x + a_2x^2 + ... + a_nx^n = a_n(\frac{a_0}{a_n} + \frac{a_1}{a_n}x + a_2x^2 + ... + x^n) = a_n(x - x_1)(x - x_2)...(x - x_n)$$

dove ora le radici sono quelle del polinomio monico $\frac{a_0}{a_n} + \frac{a_1}{a_n}x + a_2x^2 + ... + x^n$.

- L'insieme dei polinomi a coefficienti interi, razionali, reali o complessi è un **anello commutativo con unità** con le operazioni di somma e prodotto usuali.

- Un polinomio $P(x)$ è una **unità** se esiste un altro polinomio $Q(x)$ tale che $P(x)Q(x) = 1$. Se i coefficienti sono in un campo le unità sono tutti e soli i polinomi costanti diversi da zero.

- Un polinomio $P(x) \neq 0$ si dice **irriducibile** se non è una unità, e quando $P(x) = A(x)B(x)$, almeno uno dei due polinomi $A(x)$ e $B(x)$ è una unità.

La riducibilità o meno di un polinomio dipende ovviamente dall'anello in cui si considerano i coefficienti. Ad esempio il polinomio $x^2 - 2$ è riducibile sui reali: $x^2 - 2 = \left(x - \sqrt{2}\right)\left(x + \sqrt{2}\right)$, mentre è irriducibile sui razionali (perchè $\sqrt{2}$ non è un razionale). Ancora $x^2 + 2$ è riducibile sui complessi: $x^2 + 2 = \left(x - \sqrt{2}i\right)\left(x + \sqrt{2}i\right)$, mentre è irriducibile sui reali (perchè $\sqrt{2}i$ non è un reale)

I polinomi irriducibili sui complessi sono solo quelli di grado 1 (segue dal teorema fondamentale dell'algebra) mentre sui reali sono quelli di grado 1 e quelli di grado due **senza radici reali**. Infatti tutti i polinomi di grado maggiore di due sono riducibili sui reali: possiamo ridurci sempre al caso di grado pari (se è di grado dispari abbiamo visto che ha almeno una radice reale α e quindi basta dividere per $x - \alpha$) e, in questo caso, le radici reali o complesse appaiono a coppie, β e $\overline{\beta}$. Il polinomio è quindi divisibile per il polinomio **reale**:

$$P(x) = (x - \beta)\left(x - \overline{\beta}\right) = x^2 - \left(\beta + \overline{\beta}\right)x + \beta\overline{\beta}$$

Ad esempio, calcolando le radici quarte di -1, che sono:

$$\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2},\ \left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2},\ -\left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2},\ -\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2},$$

si trova:

$$x^4+1=\left(x^2-x\sqrt{2}+1\right)\left(x^2+x\sqrt{2}+1\right)$$

Per i polinomi a coefficienti **interi** vale un interessante risultato

- **Teorema delle radici razionali:** sia $P(x)=a_0+a_1x+a_2x^2+...+a_nx^n$ un polinomio a coefficienti interi e supponiamo che $r=\frac{c}{d}$ con c e d primi fra loro, sia una radice razionale di $P(x)$. Allora c divide a_0 e d divide a_n.

  Dimostrazione. Si ha quindi:

  $$a_0+a_1\frac{c}{d}+a_2\left(\frac{c}{d}\right)^2+...+a_n\left(\frac{c}{d}\right)^n=0$$

  Moltiplicando per d^n si ottiene:

  $$-a_nc^n=a_{n-1}c^{n-1}d+....+a_1cd^{n-1}+a_0d^n$$

  ne consegue che a_nc^n è divisibile per d, ma siccome d non divide c allora deve dividere a_n. Analogamente possiamo scrivere invece:

  $$a_nc^n+a_{n-1}c^{n-1}d+....+a_1cd^{n-1}=-a_0d^n$$

  E dedurre come sopra che c deve dividere a_0.

- Questo teorema ha una conseguenza molto interessante: consideriamo l'equazione $x^n-a=0$ con a intero. Se $x=\frac{c}{d}$ fosse una soluzione razionale, il teorema dimostrato sopra implica che $d=1$ perchè $a_n=1$, e quindi $x=r$ deve essere un intero. Allora necessariamente a è una potenza n - esima di un intero. Quindi $\sqrt[n]{a}$ **è razionale se e solo se a è una potenza n - esima di un intero**. Questo risultato sarà utilizzato nel corso di Algebra.

1.5 Polinomi ciclotomici.

I **polinomi ciclotomici** sono polinomi che hanno come radici le radici primitive n - esime dell'unità; sono molto interessanti e hanno importantissime applicazioni in algebra teorica (teoria dei campi finiti) e applicata all'informatica (codici correttori di errori).

- **Definizione.** Sia n un numero intero positivo; la funzione di Eulero di n, indicata con $\varphi(n)$ è il numero dei numeri minori di n e primi con n. E' una funzione molto importante e la utilizzeremo nel corso di algebra. Nella sezione dedicata alle radici dell'unità abbiamo dimostrato che $\varphi(n)$ è anche il numero delle radici n - esime primitive.

- **Definizione.** L'n - esimo polinomio ciclotomico $\Phi_n(x)$ è il polinomio monico di grado $\varphi(n)$ che ha come radici le radici n - esime primitive:

$$\Phi_n(x) = \prod_{k=1}^{\varphi(n)} (x - \alpha_k)$$

dove denotiamo con α_k le radici n - esime **primitive**. Se si considera il polinomio $x^n - 1$, le sue radici sono tutte le radici n - esime, quindi abbiamo la fattorizzazione

$$x^n - 1 = \prod_{k=1}^{n} (x - \epsilon_k)$$

dove ora ϵ_k è una qualsiasi radice n -esima. Quindi $\Phi_n(x)$ divide $x^n - 1$. Sia ora ϵ_k una qualsiasi radice n -esima e sia d il suo periodo (se $d = n$ la radice è primitiva), cioè $(\epsilon_k)^d = 1$, abbiamo già visto che d **deve dividere** n e che, in questo caso, ϵ_k è anche una radice d - esima primitiva.

Abbiamo allora, raggruppando in $\Phi_d(x)$ tutte le radici n - esime di periodo d

$$x^n - 1 = \prod_{d} \Phi_d(x)$$

Dove il prodotto è esteso a tutti i d divisori di n. Per $n = 1$ otteniamo ovviamente

$$\Phi_1(x) = x - 1$$

Se p è un numero primo,

$$x^p - 1 = \Phi_1(x)\Phi_p(x) = (x - 1)\,\Phi_p(x)$$

Per cui

$$\Phi_p(x) = \frac{x^p - 1}{x - 1} = 1 + x + x^2 + ... + x^{p-1}$$

Dalle formule precedenti possiamo allora ottenere facilmente i primi polinomi ciclotomici:

$$\Phi_1(x) = x - 1 \tag{1}$$
$$\Phi_2(x) = 1 + x \tag{2}$$
$$\Phi_3(x) = 1 + x + x^2 \tag{3}$$
$$\Phi_4(x) = 1 + x^2 \tag{4}$$
$$\Phi_5(x) = 1 + x + x^2 + x^3 + x^4 \tag{5}$$
$$\Phi_6(x) = 1 - x + x^2 \tag{6}$$
$$\Phi_7(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 \tag{7}$$
$$\Phi_8(x) = 1 + x^4 \tag{8}$$

Si può dimostrare che i polinomi ciclotomici hanno sempre coefficienti interi e sono irriducibili sui razionali. Contrariamente alle apparenze i coefficienti non sono solo $-1, 0, 1$. Ad esempio, il primo polinomio ciclotomico in cui appare anche un coefficiente 2 è $\Phi_{105}(x)$. Notate che $105 = 3 \times 5 \times 7$ è il più piccolo intero che è il prodotto di tre primi dispari distinti...e ciò non è per niente casuale...

2 Spazi vettoriali.

Uno **spazio vettoriale** è un insieme V di elementi ($\overrightarrow{v}$), detti **vettori**, su cui sono definite le seguenti operazioni:

- una **somma**: per $\overrightarrow{u}, \overrightarrow{v} \in V \Rightarrow \overrightarrow{u} + \overrightarrow{v} \in V$, tale che, per ogni $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} \in V$:

(a) $(\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w} = \overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w})$;
(b) Esiste $\overrightarrow{0} \in V$ tale che: $\overrightarrow{0} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{0} = \overrightarrow{v}$;
(c) Esiste $-\overrightarrow{v} \in V$ tale che: $\overrightarrow{v} + (-\overrightarrow{v}) = (-\overrightarrow{v}) + \overrightarrow{v} = 0$;
(d) $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}$.

- un **prodotto** $\lambda\overrightarrow{v} \in V$ di un vettore $\overrightarrow{v}$ per un *numero* λ reale (o complesso), tale che, per ogni $\lambda, \mu \in \mathbb{R}$ (o $\mathbb{C}$), $\overrightarrow{v}, \overrightarrow{w} \in V$:

(a) $\lambda(\overrightarrow{v} + \overrightarrow{w}) = \lambda\overrightarrow{v} + \lambda\overrightarrow{w}$;
(b) $(\lambda + \mu)\overrightarrow{v} = \lambda\overrightarrow{v} + \mu\overrightarrow{v}$;
(c) $(\lambda\mu)\overrightarrow{v} = \lambda(\mu\overrightarrow{v})$;
(d) $1\overrightarrow{v} = \overrightarrow{v}$
Osservazione.
Si possono definire gli spazi vettoriali su un generico corpo $(K; +; \cdot)$. I vettori godono delle proprietà 1) e 2) ma in questo caso i numeri $\lambda; \mu$ (i cosiddetti *scalari*) stanno in K; $(\lambda + \mu)$ è la somma in K e $(\lambda\mu)$ è il prodotto in K. Comunque, noi ci concentreremo solo sul caso in cui i corpi sono $\mathbb{R}$ o $\mathbb{C}$.
Per completezza, ripetiamo la definizione di spazio vettoriale nel caso più generale.

- Un **corpo** è un insieme $(K; +; \cdot)$ dotato di un'addizione $+$ e di una moltiplicazione $\cdot$ che godono delle proprietà seguenti:

1.) L'addizione è associativa;
2.) La moltiplicazione è associativa;
3.) Se $x; y \in K$ allora $x + y$ e $x \cdot y$ stanno ancora in K;
4.) Esiste un elemento $0 \in K$ tale che $x + 0 = 0 + x = x$ per ogni $x \in K$;
5.) Esiste un elemento $1 \in K$ tale che $x \cdot 1 = 1 \cdot x = x$ per ogni $x \in K$;
6.) Per ogni $x \in K$ esiste un elemento $-x \in K$ tale che $(-x) + x = x + (-x) = 0$;
7.) Per ogni $x \in K$, $x \neq 0$, esiste un elemento $x^{-1} \in K$ tale che $x^{-1} \cdot x = x \cdot x^{-1} = 1$.

- Uno **spazio vettoriale su** K è un insieme V di elementi ($\overrightarrow{v}$), detti **vettori**, su cui sono definite le seguenti operazioni:

1) una **somma**: per $\overrightarrow{u}, \overrightarrow{v} \in V \Rightarrow \overrightarrow{u} + \overrightarrow{v} \in V$, tale che:
(a) $(\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w} = \overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w})$;
(b) Esiste $\overrightarrow{0} \in V$ tale che: $\overrightarrow{0} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{0} = \overrightarrow{v}$;
(c) Esiste $-\overrightarrow{v} \in V$ tale che: $\overrightarrow{v} + (-\overrightarrow{v}) = (-\overrightarrow{v}) + \overrightarrow{v} = 0$;

(d) $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}$,
per ogni $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} \in V$.
2) un **prodotto** $\lambda \overrightarrow{v}$ di un vettore $\overrightarrow{v}$ per un $\lambda \in K$, tale che:
(a) $\lambda (\overrightarrow{v} + \overrightarrow{w}) = \lambda \overrightarrow{v} + \lambda \overrightarrow{w}$;
(b) $(\lambda + \mu) \overrightarrow{v} = \lambda \overrightarrow{v} + \mu \overrightarrow{v}$;
(c) $(\lambda\mu) \overrightarrow{v} = \lambda (\mu \overrightarrow{v})$;
(d) $1 \overrightarrow{v} = \overrightarrow{v}$,
per ogni $\lambda, \mu \in K, \overrightarrow{v}, \overrightarrow{w} \in V$.

- Alcune proprietà interessanti che si ricavano dalle definizioni sono le seguenti:

$$0 \overrightarrow{v} = \overrightarrow{0} \text{ e } - \overrightarrow{v} = (-1) \overrightarrow{v}$$

Infatti, per ogni $\overrightarrow{v} \in V$,

$$0 \overrightarrow{v} = 0 \overrightarrow{v} + \overrightarrow{0} = 0 \overrightarrow{v} + \overrightarrow{v} - \overrightarrow{v} = (0+1) \overrightarrow{v} - \overrightarrow{v} = 1 \overrightarrow{v} - \overrightarrow{v} = \overrightarrow{v} - \overrightarrow{v} = \overrightarrow{0}$$

Inoltre:

$$(-\lambda) \overrightarrow{v} + \lambda \overrightarrow{v} = (-\lambda + \lambda) \overrightarrow{v} = 0 \overrightarrow{v} = \overrightarrow{0}$$

cioè $(-\lambda) \overrightarrow{v} = -\lambda \overrightarrow{v}$ (è l'inverso di $\lambda \overrightarrow{v}$).
Esempi importanti.

1. Tutte queste proprietà sono verificate per gli insiemi $\mathbb{R}^n$ o $\mathbb{C}^n$ con le operazioni citate nel paragrafo precedente, in particolare, $(0, 0, ..., 0)$ funge da $\overrightarrow{0}$ e $(-x_1, -x_2, ..., -x_n)$ funge da $-(x_1, x_2, ..., x_n)$.

2. Consideriamo $\mathbb{R}$ (o $\mathbb{C}$) e sia $\mathbb{R}[x]$ (o $\mathbb{C}[x]$) l'insieme dei polinomi in x a coefficienti in $\mathbb{R}$:

$$\mathbb{R}[x] = \{P(x) = a_0 + a_1 x + ... + a_n x^n : a_0, a_1, ..., a_n \in \mathbb{R}\} .$$

$\mathbb{R}[x]$ è uno spazio vettoriale. Infatti, dati
$f(x) = a_0 + a_1 x + ... + a_n x^n$, $g(x) = b_0 + b_1 x + ... + b_m x^m \in \mathbb{R}[x]$
e $\lambda \in \mathbb{R}$, il polinomio somma e il polinomio prodotto

$$\begin{aligned} f(x) + g(x) &= (a_0 + b_0) + (a_1 + b_1)x + ... + (a_n + b_n)x^n + b_{n+1}x^{n+1} + ... + b_m x^m \\ \lambda f(x) &= \lambda a_0 + \lambda a_1 x + ... + \lambda a_n x^n \end{aligned}$$

appartengono a $\mathbb{R}[x]$.

Consideriamo lo spazio $\mathcal{C}(\mathbb{R})$ delle funzioni **continue** da $\mathbb{R}$ in $\mathbb{R}$. Esso è uno spazio vettoriale **reale.** Infatti, se $f; g \in \mathcal{C}(\mathbb{R})$, la loro somma è ancora una funzione continua, e quindi $f + g \in \mathcal{C}(\mathbb{R})$, dove

$$(f + g)(x) = f(x) + g(x).$$

Inoltre, se $f \in \mathcal{C}(\mathbb{R})$ e $\lambda \in \mathbb{R}$, (λf) è ancora una funzione continua, e sta in $\mathcal{C}(\mathbb{R})$.
Il vettore $\mathbf{0}$ è la funzione $\mathbf{0}(x) = 0$, $\forall x \in \mathbb{R}$.

2.1 Sottospazi, generatori e basi.

Un sottoinsieme W di V è detto **sottospazio vettoriale** di V se è esso stesso uno spazio vettoriale (con le stesse operazioni di V). In altre parole, W è un sottospazio vettoriale se
1) $\overrightarrow{0} \in W$;
2) se $\overrightarrow{v} \in W$ anche $-\overrightarrow{v} \in W$;
3) se $\overrightarrow{v}, \overrightarrow{w} \in W$ anche $\overrightarrow{v} + \overrightarrow{w} \in W$;
4) se $\lambda \in \mathbb{R}$ (o $\mathbb{C}$) e $\overrightarrow{v} \in W$ anche $\lambda \overrightarrow{v} \in W$.
Esempio.
Consideriamo i sottoinsiemi di $\mathbb{R}^2$

$$W = \{(a,a) \text{ con } a \in \mathbb{R}\}$$
$$Z = \{(a,1) \text{ con } a \in \mathbb{R}\}.$$

W è un sottospazio di $\mathbb{R}^2$, Z non è un sottospazio, perchè

$$(a,1) + (b,1) = (a+b,2) \notin Z.$$

A partire da U e W sottospazi di V, costruiamo altri sottospazi di V:

- L'intersezione di U e W è costituita dai vettori che stanno sia in U che in W

$$U \cap W = \{\overrightarrow{v} \in V : \overrightarrow{v} \in U \text{ e } \overrightarrow{v} \in W\}.$$

$U \cap W$ è un sottospazio di V, mentre $U \cup W$ in genere non lo è.
Ad esempio, siano U e W due rette in $\mathbb{R}^2$ passanti per l'origine (cioè da $\overrightarrow{0}$). L'unione di U e W non contiene la somma di un elemento di U e uno di W diversi da $\overrightarrow{0}$.
Per fabbricare un sottospazio dall'unione si devono aggiungere tutti gli elementi mancanti (nel caso precedente, tutte le somme).

- Consideriamo lo spazio formato da tutte le somme di elementi di U con elementi di W:

$$U + W = \{\overrightarrow{u} + \overrightarrow{w} \text{ con } \overrightarrow{u} \in U,\ \overrightarrow{w} \in W\}.$$

$U + W$ è un sottospazio vettoriale di V. Naturalmente $U + V$ contiene sia U che V.
Se $U \cap W = \left\{\overrightarrow{0}\right\}$, l'operazione $+$ tra sottospazi si indica con $U \oplus W$ e si dice **somma diretta**; la cosa interessante è che, in questo caso, ogni elemento di $U \oplus W$ si scrive in modo *unico* come somma di un elemento di U e uno di W. (Esercizio!)
Esempio. (*)
Consideriamo in $\mathbb{R}^3$ i seguenti sottospazi:

$$U = \{(x,0,x) : x \in \mathbb{R}\};$$
$$V = \{(x,0,0) : x \in \mathbb{R}\};$$
$$W = \{(x,0,y) : x,y \in \mathbb{R}\};$$
$$Z = \{(0,x,y) : x,y \in \mathbb{R}\}.$$

Si ha

$$\begin{gathered} U \cap V = \{(x,0,x) = (x,0,0)\} = \{(0,0,0)\}; \\ U \cap W = \{(x,0,x) = (x,0,y)\} = \{(x,0,x)\} = U; \\ Z \cap W = \{(0,x,y) = (x,0,y)\} = \{(0,0,y)\}; \\ U + V = \{(x,0,x) + (x',0,0) \text{ con } x, x' \in \mathbb{R}\} \\ = \{(a,0,b) \text{ con } a, b \in \mathbb{R}\} = W. \end{gathered}$$

- Siano $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ vettori di uno spazio vettoriale V e $a_1, ..., a_n \in \mathbb{R}$ (o $\mathbb{C}$). Il vettore

$$\overrightarrow{v} = a_1 \overrightarrow{v}_1 + ... + a_n \overrightarrow{v}_n = \sum_{i=1}^{n} a_i \overrightarrow{v}_i,$$

si dice **combinazione lineare** dei vettori $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$.
Sia $S = \{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ un insieme finito di vettori di V. Lo spazio

$$\text{Span}\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\} = \{\text{tutte le possibili combinazioni lineari di } \overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$$

è detto **span lineare** ed è un sottospazio vettoriale di V, precisamente è il più piccolo sottospazio che contiene i vettori $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$.

Diremo che $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ **generano** lo spazio vettoriale V o che $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ è un **sistema di generatori** di V se $V = \text{Span}\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$. Quindi $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ è un sistema di generatori di V se e solo se $\forall \overrightarrow{v} \in V$ esistono $a_1, ..., a_n \in \mathbb{R}$ (o $\mathbb{C}$) tali che $\overrightarrow{v} = a_1 \overrightarrow{v}_1 + ... + a_n \overrightarrow{v}_n$.

I vettori $\overrightarrow{v}_1, ..., \overrightarrow{v}_n \in V$ si dicono **linearmente dipendenti** se $\exists a_1, ..., a_n \in \mathbb{R}$ (o $\mathbb{C}$) **non tutti nulli**, tali che $a_1 \overrightarrow{v}_1 + ... + a_n \overrightarrow{v}_n = \overrightarrow{0}$. Altrimenti, si dicono **linearmente indipendenti** (cioè $a_1 \overrightarrow{v}_1 + ... + a_n \overrightarrow{v}_n = \overrightarrow{0} \Rightarrow a_i = 0, \forall i$).

Un insieme finito $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ di vettori di V, si dice **base** di V se $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ sono **linearmente indipendenti e se sono un sistema di generatori di** V. Se $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ è una base di V, allora

1) $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ generano $V \Rightarrow \forall \overrightarrow{v} \in V, \overrightarrow{v} = a_1 \overrightarrow{v}_1 + ... + a_n \overrightarrow{v}_n$;

2) gli $a_1, ..., a_n$ sono determinati in modo **unico** da $\overrightarrow{v}$ (ciò è dovuto al fatto che $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$ sono linermente indipendenti).

Gli $a_1, ..., a_n \in \mathbb{R}$ (o $\mathbb{C}$) si dicono **coordinate** di $\overrightarrow{v}$ rispetto alla base $\overrightarrow{v}_1, ..., \overrightarrow{v}_n$.

- **Teorema:** Supponiamo che $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ siano un sistema di generatori per V, e che $\{\overrightarrow{w}_1, ..., \overrightarrow{w}_m\}$ siano linearmente indipendenti. Allora, $m \leq n$.

Dimostrazione. Abbiamo $\overrightarrow{w}_1 = a_1 \overrightarrow{v}_1 + a_2 \overrightarrow{v}_2 + ... + a_n \overrightarrow{v}_n$. Allora, dato che $\overrightarrow{w}_1 \neq 0$, qualcuno degli a_i è non nullo (noi supporremo, per esempio, che sia $a_1 \neq 0$). Dunque, $\{\overrightarrow{w}_1; \overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ è linearmente dipendente, e genera ovviamente V. Allora, possiamo scrivere $\overrightarrow{v}_1$ in funzione di $\overrightarrow{w}_1$ e di $\overrightarrow{v}_2, ..., \overrightarrow{v}_n$, dato che

$$\overrightarrow{v}_1 = \tfrac{1}{a_1} (\overrightarrow{w}_1 - a_2 \overrightarrow{v}_2 - ... - a_n \overrightarrow{v}_n).$$

Dunque, $\{\overrightarrow{w}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_n\}$ genera V.

Sia poi $\overrightarrow{w}_2 = b'\overrightarrow{w}_1 + b_2\overrightarrow{v}_2 + ... + b_n\overrightarrow{v}_n$. Dato che $\overrightarrow{w}_1$ e $\overrightarrow{w}_2$ sono linearmente indipendenti, qualcuno dei $\{b_i\}$ deve essere diverso da zero (altrimenti esisterebbe una combinazione lineare di $\overrightarrow{w}_1$ e $\overrightarrow{w}_2$ che si annullerebbe). Supponiamo che $b_2 \neq 0$. Allora, possiamo scrivere $\overrightarrow{v}_2$ in funzione di $\overrightarrow{w}_1; \overrightarrow{w}_2$ e di $\overrightarrow{v}_3, ..., \overrightarrow{v}_n$, dato che

$$\overrightarrow{v}_1 = \tfrac{1}{b_2}\left(\overrightarrow{w}_2 - b'\overrightarrow{w}_1 - b_3\overrightarrow{v}_3 - ... - b_n\overrightarrow{v}_n\right).$$

Dunque, $\{\overrightarrow{w}_1; \overrightarrow{w}_2; \overrightarrow{v}_3 ...; \overrightarrow{v}_n\}$ genera V.

Di conseguenza, in ognuno dei passaggi, possiamo aggiungere uno dei $\{\overrightarrow{w}_i\}$ al sistema dei generatori ed eliminare contemporaneamente uno dei $\{\overrightarrow{v}_i\}$. Non è possibile, allora, che $m > n$, perchè, altrimenti, procedendo nello stesso modo, si potrebbe scrivere $\overrightarrow{w}_{n+1}$ come combinazione lineare di $\{\overrightarrow{w}_1; ...; \overrightarrow{w}_n\}$. Siamo però sicuri che questo non può mai accadere, perchè abbiamo scelto l'ipotesi che $\{\overrightarrow{w}_1, ..., \overrightarrow{w}_m\}$ siano linearmente indipendenti.

- **Teorema:** Tutte le basi di uno spazio vettoriale sono costituite dallo stesso numero di elementi.

Dimostrazione. Supponiamo che $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$ e $\{\overrightarrow{w}_1, ..., \overrightarrow{w}_m\}$ siano due basi di V. Allora, esse sono entrambe sistemi di generatori per V, ed entrambe sono linearmente indipendenti. Così, per il teorema precedente, ricaviamo sia $m \leq n$, sia $n \leq m$, cioè che $m = n$.

- Se uno spazio vettoriale V possiede una base $\{\overrightarrow{v}_1, ..., \overrightarrow{v}_n\}$, il numero n si dice **dimensione** di V e si indica con $\dim V$. Se $V = \left\{\overrightarrow{0}\right\}$, si pone $\dim V = 0$. Se $V = \left\{\overrightarrow{0}\right\}$, oppure V possiede una base, diciamo che V ha **dimensione finita**. Con questa definizione $\mathbb{R}^n$ ha dimensione n.

Esempio.

1) In $\mathbb{R}^2$, consideriamo i vettori

$$\overrightarrow{e}_1 = (1,0), \overrightarrow{e}_2 = (0,1).$$

$\overrightarrow{e}_1, \overrightarrow{e}_2$ generano $\mathbb{R}^2$. Infatti, sia $\overrightarrow{v} = (v_1, v_2) \in \mathbb{R}^2$. Si ha

$$\overrightarrow{v} = (v_1, v_2) = (v_1, 0) + (0, v_2) = v_1(1,0) + v_2(0,1) = v_1\overrightarrow{e}_1 + v_2\overrightarrow{e}_2.$$

Inoltre, $\overrightarrow{v} = (v_1, v_2) = \overrightarrow{0} \Leftrightarrow v_1 = v_2 = 0 \Rightarrow \overrightarrow{e}_1, \overrightarrow{e}_2$ linearmente indipendenti $\Rightarrow \{\overrightarrow{e}_1, \overrightarrow{e}_2\}$ è una base di $\mathbb{R}^2$ e $\mathbb{R}^2$ ha dimensione 2. $\{\overrightarrow{e}_1, \overrightarrow{e}_2\}$ si chiama **base canonica** di $\mathbb{R}^2$.

2) In $\mathbb{R}^3$, la base canonica è costituita dai tre vettori

$$\overrightarrow{e}_1 = (1,0,0), \overrightarrow{e}_2 = (0,1,0), \overrightarrow{e}_3 = (0,0,1).$$

3) Il sottospazio $\mathbb{R}[x]_{\leqslant 3}$ dei polinomi di grado $\leqslant 3$, possiede una base finita: $\{1, x, x^2, x^3\}$. Infatti, il polinomio:

$$P(x) = a_0 + a_1x + a_2x^2 + a_3x^3 = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

è combinazione lineare di $\{1, x, x^2, x^3\} \Rightarrow \{1, x, x^2, x^3\}$ è un sistema di generatori. Inoltre, se $a_0 + a_1x + a_2x^2 + a_3x^3 = 0 \Rightarrow a_0 = a_1 = a_2 = a_3 = 0$,

quindi $1, x, x^2, x^3$ sono linearmente indipendenti. $\mathbb{R}[x]_{\leqslant 3}$ ha dimensione 4 (numero di elementi della base).

- Se U ha base $\overrightarrow{v}_1, ..., \overrightarrow{v}_l$ e V ha base $\overrightarrow{e}_1, ..., \overrightarrow{e}_m$, allora

$$U + V = \text{Span}\{\overrightarrow{v}_1, ..., \overrightarrow{v}_l, \overrightarrow{e}_1, ..., \overrightarrow{e}_m\}.$$

- **Teorema delle dimensioni**. Se U e V sono spazi vettoriali di dimensione finita, allora

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

In particolare (ricordando che per la somma diretta $U \oplus V$ si ha $U \cap V = \left\{\overrightarrow{0}\right\}$),

$$\dim(U \oplus V) = \dim U + \dim V.$$

Dimostrazione. Prima di tutto, verifichiamo che, se A è un sottospazio vettoriale di B, e $\{\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha\}$ è una base di A, allora è sempre possibile estenderla ad una base di B della forma $\left\{\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha; \overrightarrow{b}_1; \overrightarrow{b}_2; ...; \overrightarrow{b}_\beta\right\}$ di B. Sia, infatti, $\{\overrightarrow{c}_1; \overrightarrow{c}_2; ...; \overrightarrow{c}_k\}$ una base di B. Allora $\{\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha; \overrightarrow{c}_1; \overrightarrow{c}_2; ...; \overrightarrow{c}_k\}$ è sicuramente un sistema di generatori di B. Ora, però, essi non sono linerarmente indipendenti. Per ottenere una base di B, dobbiamo togliere da $\{\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha; \overrightarrow{c}_1; \overrightarrow{c}_2; ...; \overrightarrow{c}_k\}$ alcuni vettori. Procediamo così: scorriamo da sinistra a destra i vettori di $\{\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha; \overrightarrow{c}_1; \overrightarrow{c}_2; ...; \overrightarrow{c}_k\}$. Un elemento verrà eliminato (e quindi non comparirà nella base estesa di V), se esso si può scrivere come combinazione lineare dei vettori alla sua sinistra. In questo modo, $\overrightarrow{a}_1; ... \overrightarrow{a}_\alpha$ non sono toccati, perchè già sappiamo che sono linearmente indipendenti, e, quando si arriva all'ultimo elemento a destra, si ottiene una base di B.

Ora dimostriamo il teorema delle dimensioni.

Sia $W = U \cap V$. Supponiamo che $\dim(W) = r$; $\dim(U) = m$ e $\dim(V) = n$. Sia $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r\}$ una base di W. Poichè W è un sottospazio vettoriale sia di U che di V, possiamo estendere la base data ad una base $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{u}_1; \overrightarrow{u}_2; ...; \overrightarrow{u}_{m-r}\}$ di U e ad una base $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_{n-r}\}$ di V. Considero ora l'insieme $\mathcal{B} = \{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{u}_1; \overrightarrow{u}_2; ...; \overrightarrow{u}_{m-r}; \overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_{n-r}\}$.

Ogni vettore di $U+V$ si può scrivere come $\overrightarrow{u} + \overrightarrow{v}$, per qualche $\overrightarrow{u} \in U$; $\overrightarrow{v} \in V$. Poichè $\overrightarrow{u}$ si può scrivere come combinazione lineare di $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{u}_1; \overrightarrow{u}_2; ...; \overrightarrow{u}_{m-r}\}$ e $\overrightarrow{v}$ si può scrivere come combinazione lineare di $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_{n-r}\}$, $\overrightarrow{u} + \overrightarrow{v}$ è combinazione lineare di elementi di $\mathcal{B}$. Dunque, $\mathcal{B}$ è un sistema di generatori per $U + V$.

Dimostro ora che gli elementi di $\mathcal{B}$ sono linearmente indipendenti. Supponiamo che sia

$$a_1\overrightarrow{w}_1 + ... + a_r\overrightarrow{w}_r + b_1\overrightarrow{u}_1 + ... + b_{m-r}\overrightarrow{u}_{m-r} + c_1\overrightarrow{v}_1 + ... + c_{n-r}\overrightarrow{v}_{n-r} = 0,$$

dove $\{a_i; b_j; c_k\}$ sono scalari. Sia

$$\overrightarrow{v} = a_1\overrightarrow{w}_1 + ... + a_r\overrightarrow{w}_r + b_1\overrightarrow{u}_1 + ... + b_{m-r}\overrightarrow{u}_{m-r};$$

si ha anche

$$\overrightarrow{v} = -c_1\overrightarrow{v}_1 - ... - c_{n-r}\overrightarrow{v}_{n-r}.$$

Dalle ultime due uguaglianze si vede subito che $\overrightarrow{v}$ appartiene sia ad U che a V, e quindi sta in $U \cap V$. Pertanto, $\overrightarrow{v}$ è combinazione lineare di $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r\}$, ovvero $\overrightarrow{v} = d_1\overrightarrow{w}_1 + ... + d_r\overrightarrow{w}_r$. Sostituendo nella relazione precedente, si trova

$$d_1\overrightarrow{w}_1 + ... + d_r\overrightarrow{w}_r + c_1\overrightarrow{v}_1 + ... + c_{n-r}\overrightarrow{v}_{n-r} = 0.$$

Ma $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_{n-r}\}$ è una base di V, e quindi la relazione precedente, per l'indipendenza lineare, forza $c_1 = c_2 = ... = c_{n-r} = 0$ e $d_1 = d_2 = ... = d_r = 0$. Così, $\overrightarrow{v} = 0$.

Sostituendo ancora, $a_1\overrightarrow{w}_1 + ... + a_r\overrightarrow{w}_r + b_1\overrightarrow{u}_1 + ... + b_{m-r}\overrightarrow{u}_{m-r} = 0$. Pertanto, essendo $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r; \overrightarrow{u}_1; \overrightarrow{u}_2; ...; \overrightarrow{u}_{m-r}\}$ una base di U, anche $a_1 = a_2 = ... = a_r = 0$ e $b_1 = b_2 = ... = b_{m-r} = 0$.

Esempio.

Vediamo che per gli spazi dell'esempio (*) è soddisfatto il teorema delle dimensioni. Si ha: $U \cap V = \{(0,0,0)\} \Rightarrow \dim(U \cap V) = 0$.

U è formato da vettori del tipo $(x,0,x) = x(1,0,1)$, cioè da multipli di $(1,0,1)$, quindi è generato da $(1,0,1) \Rightarrow \dim U = 1$.

V è formato da vettori del tipo $(x,0,0) = x(1,0,0)$, quindi è generato dal vettore $(1,0,0) \Rightarrow \dim V = 1$.

Un vettore di $W = U + V$ è, invece, dato da $a(1,0,0) + b(0,0,1)$, cioè è combinazione lineare dei vettori $(1,0,0)$ e $(0,0,1) \Rightarrow \dim W = 2$. Allora vale $\dim W = \dim U + \dim V - \dim(U \cap V)$.

Esempio.

Siano $U = \{(x,y,0) : x,y \in \mathbb{R}\}$ e $V = \{(0,a,b) : a,b \in \mathbb{R}\}$. Si ha $\mathbb{R}^3 = U + V$, e quindi $\dim(U+V) = 3$. D'altra parte, $\dim(U \cap V) = 1$, perchè $U \cap V$ è l'asse y, così la formula è verificata.

$\dim(U+V) = \dim U + \dim V - \dim(U \cap V)$;

$3 = 2 + 2 - 1$.

3 Prodotti Scalari e Hermitiani.

Consideriamo $\mathbb{R}^n$ con la base canonica $\overrightarrow{e}_1, ..., \overrightarrow{e}_n$ ($\overrightarrow{e}_1 = (1,0,...,0)$, $\overrightarrow{e}_2 = (0,1,0,...,0)$, ..., $\overrightarrow{e}_n = (0,...,0,1)$) e siano $\overrightarrow{v}, \overrightarrow{w} \in \mathbb{R}^n$ due vettori: $\overrightarrow{v} = (v_1, v_2, ..., v_n)$, $\overrightarrow{w} = (w_1, w_2, ..., w_n)$. Definiamo il p**rodotto scalare standard** di $\overrightarrow{v}$ e $\overrightarrow{w}$:

$$\overrightarrow{v} \cdot \overrightarrow{w} = \sum_{i=1}^{n} v_i w_i.$$

Proprietà: $\forall \overrightarrow{v}, \overrightarrow{u}, \overrightarrow{w} \in \mathbb{R}^n$, $a, b \in \mathbb{R}$:
1) $(a\overrightarrow{v} + b\overrightarrow{u}) \cdot \overrightarrow{w} = a\overrightarrow{v} \cdot \overrightarrow{w} + b\overrightarrow{u} \cdot \overrightarrow{w}$;
2) $\overrightarrow{v} \cdot (a\overrightarrow{u} + b\overrightarrow{w}) = a\overrightarrow{v} \cdot \overrightarrow{u} + b\overrightarrow{v} \cdot \overrightarrow{w}$;
3) $\overrightarrow{v} \cdot \overrightarrow{v} \geqslant 0$ e $\overrightarrow{v} \cdot \overrightarrow{v} = 0 \Leftrightarrow \overrightarrow{v} = \overrightarrow{0}$;
$\sqrt{\overrightarrow{v} \cdot \overrightarrow{v}}$ si indica con$\|\overrightarrow{v}\|$ e si dice **norma** o **lunghezza** di $\overrightarrow{v}$.
4) $\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \|\overrightarrow{v}\| \cos\theta$, dove θ è l'angolo compreso tra i due vettori $\overrightarrow{u}$ e $\overrightarrow{v}$.
Due vettori $\overrightarrow{u}$ e $\overrightarrow{v}$ si dicono **ortogonali** se $\overrightarrow{u} \cdot \overrightarrow{v} = 0$.

3.1 Prodotti scalari reali.

Sia V uno spazio vettoriale **reale**. Un **prodotto scalare** è un'applicazione che ad ogni coppia di vettori $\overrightarrow{u}, \overrightarrow{v} \in V$ associa un numero reale $\langle \overrightarrow{u}; \overrightarrow{v} \rangle$, e che soddisfa i seguenti assiomi:

1. (Proprietà lineare) $\langle a_1\overrightarrow{u}_1 + a_2\overrightarrow{u}_2; \overrightarrow{v} \rangle = a_1 \langle \overrightarrow{u}_1; \overrightarrow{v} \rangle + a_2 \langle \overrightarrow{u}_2; \overrightarrow{v} \rangle$, $\forall \overrightarrow{u}_1; \overrightarrow{u}_2; \overrightarrow{v} \in V$; $\forall a_1; a_2 \in \mathbb{R}$.
2. (Proprietà simmetrica) $\langle \overrightarrow{u}; \overrightarrow{v} \rangle = \langle \overrightarrow{v}; \overrightarrow{u} \rangle$, $\forall \overrightarrow{u}; \overrightarrow{v} \in V$.
3. (Proprietà definita positiva) $\langle \overrightarrow{u}; \overrightarrow{u} \rangle \geq 0$, e $\langle \overrightarrow{u}; \overrightarrow{u} \rangle = 0$ se e solo se $\overrightarrow{u} = \overrightarrow{0}$.

Esempio.
Il *prodotto scalare standard* descritto in precedenza è un prodotto scalare su $\mathbb{R}^n$.
Esempio.
Sia V lo spazio vettoriale $\mathcal{C}([a; b])$ delle funzioni **continue** sull'intervallo chiuso $[a; b]$. Su V, definiamo

$$\langle f; g \rangle = \int_a^b f(x)\, g(x)\, dx.$$

In questo modo, si definisce un prodotto scalare su $\mathcal{C}([a; b])$.
Esempio.
Considero il seguente prodotto in $\mathbb{R}^2$:
$\langle \overrightarrow{u}; \overrightarrow{v} \rangle = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2$, per $\overrightarrow{u} = (x_1; x_2)$ e $\overrightarrow{v} = (y_1; y_2)$. Verifico che esso è un prodotto scalare su $\mathbb{R}^2$.
1.) Se $\overrightarrow{u} = (x_1; x_2)$; $\overrightarrow{v} = (y_1; y_2)$ e $\overrightarrow{w} = (z_1; z_2)$, dati $a; b \in \mathbb{R}$, si ha
$\langle a\overrightarrow{u} + b\overrightarrow{w}; \overrightarrow{v} \rangle = \langle (ax_1 + bz_1; ax_2 + bz_2); (y_1; y_2) \rangle =$
$= (ax_1 + bz_1) y_1 - (ax_1 + bz_1) y_2 - (ax_2 + bz_2) y_1 + 3 (ax_2 + bz_2) y_2 =$
$= a (x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2) + b (z_1y_1 - z_1y_2 - z_2y_1 + 3z_2y_2) =$
$= a \langle \overrightarrow{u}; \overrightarrow{v} \rangle + b \langle \overrightarrow{w}; \overrightarrow{v} \rangle$. La proprietà lineare è verificata.
2.) $\langle \overrightarrow{u}; \overrightarrow{v} \rangle = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2 = y_1x_1 - y_2x_1 - y_1x_2 + 3y_2x_2 = \langle \overrightarrow{v}; \overrightarrow{u} \rangle$.
La proprietà simmetrica è verificata.
3.) $\langle \overrightarrow{u}; \overrightarrow{u} \rangle = x_1^2 - 2x_1x_2 + 3x_2^2 = (x_1 - x_2)^2 + 2x_2^2$.
Dunque, $\langle \overrightarrow{u}; \overrightarrow{u} \rangle \geq 0$ e $\langle \overrightarrow{u}; \overrightarrow{u} \rangle = 0$ se e solo se $\overrightarrow{u} = 0$. Anche la proprietà definita positiva è verificata.
Alla fine, ho verificato che $\langle ; \rangle$ è un prodotto scalare su $\mathbb{R}^2$.

- DISUGUAGLIANZA DI CAUCHY-SCHWARTZ **(caso reale).**

Se V è uno spazio vettoriale reale, per ogni $\overrightarrow{u}; \overrightarrow{v} \in V$ si ha:

$$\langle \overrightarrow{u}; \overrightarrow{v} \rangle^2 \leq \langle \overrightarrow{u}; \overrightarrow{u} \rangle \cdot \langle \overrightarrow{v}; \overrightarrow{v} \rangle.$$

Dimostrazione. Per ogni numero reale t,
$\langle t\overrightarrow{u} + \overrightarrow{v}; t\overrightarrow{u} + \overrightarrow{v} \rangle = t^2 \langle \overrightarrow{u}; \overrightarrow{u} \rangle + 2t \langle \overrightarrow{u}; \overrightarrow{v} \rangle + \langle \overrightarrow{v}; \overrightarrow{v} \rangle$. Questo polinomio in t è sempre ≥ 0, quindi non può avere due radici reali distinte. Così, il discriminante Δ deve essere minore di zero.

$\Delta = (4 \langle \overrightarrow{u}; \overrightarrow{v} \rangle)^2 - 4 \langle \overrightarrow{u}; \overrightarrow{u} \rangle \cdot \langle \overrightarrow{v}; \overrightarrow{v} \rangle \leq 0$, da cui si ottiene la disuguaglianza di Cauchy-Schwartz.

- Sia V uno spazio vettoriale dotato di un prodotto scalare $\langle ; \rangle$. Due vettori $\overrightarrow{u}; \overrightarrow{v} \in V$ si dicono **ortogonali** se $\langle \overrightarrow{u}; \overrightarrow{v} \rangle = 0$.

Osservazione.
$\overrightarrow{0}$ è ortogonale ad ogni altro vettore, perchè $\left\langle \overrightarrow{0}; \overrightarrow{v} \right\rangle = \langle 0\overrightarrow{v}; \overrightarrow{v} \rangle = 0 \langle \overrightarrow{v}; \overrightarrow{v} \rangle = 0$, per ogni $\overrightarrow{v} \in V$.

Esempio.
I vettori $\overrightarrow{e}_1 = (1; 0; 1)$ e $\overrightarrow{e}_2 = (-2; 1; 2)$ sono ortogonali rispetto al prodotto scalare standard di $\mathbb{R}^3$.

- Sia V uno spazio vettoriale con un sottospazio vettoriale $W \subseteq V$. Si definisce il **complemento ortogonale** di W come

$$W^\perp = \{\overrightarrow{v} \in V : \langle \overrightarrow{v}; \overrightarrow{w} \rangle = 0, \forall \overrightarrow{w} \in W\}.$$

Verifichiamo che $W^\perp$ è un sottospazio vettoriale di V.
1.) Prima di tutto, se $\overrightarrow{w}_1^\perp$ e $\overrightarrow{w}_2^\perp$ stanno in $W^\perp$, si ha, per ogni $\overrightarrow{w} \in W$,
$\left\langle \overrightarrow{w}_1^\perp + \overrightarrow{w}_2^\perp; \overrightarrow{w} \right\rangle = \left\langle \overrightarrow{w}_1^\perp; \overrightarrow{w} \right\rangle + \left\langle \overrightarrow{w}_2^\perp; \overrightarrow{w} \right\rangle = 0 + 0 = 0$.
Dunque, $\overrightarrow{w}_1^\perp + \overrightarrow{w}_2^\perp \in W^\perp$.
2.) Se $\lambda \in \mathbb{R}$, allora $\left\langle \lambda\overrightarrow{w}_1^\perp; \overrightarrow{w} \right\rangle = \lambda \left\langle \overrightarrow{w}_1^\perp; \overrightarrow{w} \right\rangle = \lambda 0 = 0$. Anche $\lambda\overrightarrow{w}_1^\perp \in W^\perp$.
3.) $\overrightarrow{0} \in W^\perp$, per l'osservazione precedente.
4.) Infine, se $\overrightarrow{w}^\perp \in W^\perp$, allora, per ogni $\overrightarrow{w} \in W$,
$\left\langle -\overrightarrow{w}^\perp; \overrightarrow{w} \right\rangle = -\left\langle \overrightarrow{w}^\perp; \overrightarrow{w} \right\rangle = 0$. Si vede così che $-\overrightarrow{w}^\perp \in W^\perp$.

Esempio.
Consideriamo in $\mathbb{R}^3$ i vettori

$$\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \overrightarrow{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \overrightarrow{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \overrightarrow{v}_4 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix},$$
$$\overrightarrow{v}_5 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \overrightarrow{v}_6 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix},$$

e i sottospazi $U = \text{Span}\{\overrightarrow{v}_1, \overrightarrow{v}_2\}$ e $V = \text{Span}\{\overrightarrow{v}_3, \overrightarrow{v}_4, \overrightarrow{v}_5, \overrightarrow{v}_6\}$.

Calcoliamo la dimensione e una base per i sottospazi $U \cap V$ e $U + V$.

Si nota subito che $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_4$ sono indipendenti e che $\overrightarrow{v}_3, \overrightarrow{v}_5, \overrightarrow{v}_6$ sono loro combinazioni lineari: $\overrightarrow{v}_6 = 2\overrightarrow{v}_1$, $\overrightarrow{v}_3 = \overrightarrow{v}_1 + \overrightarrow{v}_2$ e $\overrightarrow{v}_5 = \overrightarrow{v}_1 + \overrightarrow{v}_2 - \overrightarrow{v}_4$. Quindi, $V = \text{Span}\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_4\}$. Ne segue che $U \cap V = \text{Span}\{\overrightarrow{v}_1, \overrightarrow{v}_2\}$ ha dimensione due e una base è $\overrightarrow{v}_1, \overrightarrow{v}_2$. $U + V = \text{Span}\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_4\}$ ha dimensione tre e una base è $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_4$.

Per scrivere una base del complemento ortogonale a $U \cap V$, che ha dimensione 1, basta scrivere un vettore (a, b, c) ortogonale a $\overrightarrow{v}_1, \overrightarrow{v}_2$. Deve essere:

$$\begin{cases} a + b + c = 0 \\ b - c = 0 \end{cases} \Rightarrow \begin{cases} a = -2c \\ b = c \\ c = t \end{cases} \Rightarrow (a, b, c) = t(-2, 1, 1).$$

Quindi, per $t = 1$, una base di $(U \cap V)^{\perp}$ è data dal vettore $(-2, 1, 1)$.

3.2 Prodotti Hermitiani.

Sia V uno spazio vettoriale **complesso**. Un **prodotto hermitiano** è un'applicazione che ad ogni coppia di vettori $\overrightarrow{u}; \overrightarrow{v} \in V$ associa un numero complesso $\langle \overrightarrow{u}; \overrightarrow{v} \rangle$, e che soddisfa i seguenti assiomi:

1. (Proprietà lineare) $\langle a_1 \overrightarrow{u}_1 + a_2 \overrightarrow{u}_2; \overrightarrow{v} \rangle = a_1 \langle \overrightarrow{u}_1; \overrightarrow{v} \rangle + a_2 \langle \overrightarrow{u}_2; \overrightarrow{v} \rangle$, $\forall \overrightarrow{u}_1; \overrightarrow{u}_2; \overrightarrow{v} \in V$; $\forall a_1; a_2 \in \mathbb{C}$.

2. (Proprietà sesquilineare) $\langle \overrightarrow{u}; b_1 \overrightarrow{v}_1 + b_2 \overrightarrow{v}_2 \rangle = \overline{b_1} \langle \overrightarrow{u}; \overrightarrow{v}_1 \rangle + \overline{b_2} \langle \overrightarrow{u}; \overrightarrow{v}_2 \rangle$, $\forall \overrightarrow{u}; \overrightarrow{v}_1; \overrightarrow{v}_2 \in V$; $\forall b_1; b_2 \in \mathbb{C}$.

3. (Proprietà simmetrica coniugata) $\langle \overrightarrow{u}; \overrightarrow{v} \rangle = \overline{\langle \overrightarrow{v}; \overrightarrow{u} \rangle}$, $\forall \overrightarrow{u}; \overrightarrow{v} \in V$.

4. (Proprietà definita positiva) $\langle \overrightarrow{u}; \overrightarrow{u} \rangle$ è reale, $\langle \overrightarrow{u}; \overrightarrow{u} \rangle \geq 0$, e $\langle \overrightarrow{u}; \overrightarrow{u} \rangle = 0$ se e solo se $\overrightarrow{u} = \overrightarrow{0}$.

- IL PRODOTTO HERMITIANO STANDARD.

Consideriamo $\mathbb{C}^n$ con la base canonica $\overrightarrow{e}_1, ..., \overrightarrow{e}_n$ ($\overrightarrow{e}_1 = (1, 0, ..., 0)$, $\overrightarrow{e}_2 = (0, 1, 0, ..., 0)$, ..., $\overrightarrow{e}_n = (0, ..., 0, 1)$) e siano $\overrightarrow{v}, \overrightarrow{w} \in \mathbb{C}^n$ due vettori: $\overrightarrow{v} = (v_1, v_2, ..., v_n)$, $\overrightarrow{w} = (w_1, w_2, ..., w_n)$. Definiamo il p**rodotto hermitiano standard** di $\overrightarrow{v}$ e $\overrightarrow{w}$:

$$\overrightarrow{v} \cdot \overrightarrow{w} = \sum_{i=1}^{n} v_i \overline{w}_i.$$

Le definizioni di ortogonalità e di complemento ortogonale si estendono senza alcuna differenza anche al caso complesso. Anche la definizione di norma si estende senza differenze agli spazi vettoriali complessi dotati di un prodotto hermitiano $\langle ; \rangle$:

$$\| \overrightarrow{v} \| = \sqrt{\langle \overrightarrow{v}; \overrightarrow{v} \rangle}.$$

• DISUGUAGLIANZA DI CAUCHY-SCHWARTZ **(caso complesso).**

Se V è uno spazio vettoriale complesso, per ogni $\overrightarrow{u}; \overrightarrow{v} \in V$ si ha:

$$|\langle \overrightarrow{u}; \overrightarrow{v} \rangle|^2 \leq \langle \overrightarrow{u}; \overrightarrow{u} \rangle \langle \overrightarrow{v}; \overrightarrow{v} \rangle.$$

Dim.) Per ogni numero reale t,
$0 \leq \langle \overrightarrow{u} - \langle \overrightarrow{u}; \overrightarrow{v} \rangle t \overrightarrow{v}; \overrightarrow{u} - \langle \overrightarrow{u}; \overrightarrow{v} \rangle t \overrightarrow{v} \rangle = \langle \overrightarrow{u}; \overrightarrow{u} \rangle - t\overline{\langle \overrightarrow{u}; \overrightarrow{v} \rangle} \cdot \langle \overrightarrow{u}; \overrightarrow{v} \rangle - t \langle \overrightarrow{u}; \overrightarrow{v} \rangle \langle \overrightarrow{v}; \overrightarrow{u} \rangle +$ $t^2 \overline{\langle \overrightarrow{u}; \overrightarrow{v} \rangle} \cdot \langle \overrightarrow{u}; \overrightarrow{v} \rangle \cdot \langle \overrightarrow{v}; \overrightarrow{v} \rangle =$
$= \langle \overrightarrow{u}; \overrightarrow{u} \rangle - 2t |\langle \overrightarrow{u}; \overrightarrow{v} \rangle|^2 + |\langle \overrightarrow{u}; \overrightarrow{v} \rangle|^2 t^2 \cdot \langle \overrightarrow{v}; \overrightarrow{v} \rangle.$
La disuguaglianza si ricava prendendo $t = \frac{1}{\langle \overrightarrow{v}; \overrightarrow{v} \rangle}$: $\langle \overrightarrow{u}; \overrightarrow{u} \rangle - \frac{\|\langle \overrightarrow{u}; \overrightarrow{v} \rangle\|^2}{\langle \overrightarrow{v}; \overrightarrow{v} \rangle} \geq 0.$
La disuguaglianza si può scrivere anche così: (abbiamo a che fare con numeri reali non negativi!)

$$|\langle \overrightarrow{u}; \overrightarrow{v} \rangle| \leq \sqrt{\langle \overrightarrow{u}; \overrightarrow{u} \rangle \langle \overrightarrow{v}; \overrightarrow{v} \rangle} = \|\overrightarrow{u}\| \|\overrightarrow{v}\|$$

• DISUGUAGLIANZA TRIANGOLARE.

Se V è uno spazio vettoriale reale o complesso, per ogni $\overrightarrow{u}, \overrightarrow{v} \in V$ si ha:

$$\|\overrightarrow{u} + \overrightarrow{v}\| \leq \|\overrightarrow{u}\| + \|\overrightarrow{v}\|.$$

Dalla precedente disuguaglianza si ottiene:

$$|\langle \overrightarrow{u}; \overrightarrow{v} \rangle|^2 = (\mathrm{Re}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle)^2 + (\mathrm{Im}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle)^2 \leq \|\overrightarrow{u}\|^2 \|\overrightarrow{v}\|^2$$

per cui:

$$|\mathrm{Re}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle| \leq \sqrt{(\mathrm{Re}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle)^2 + (\mathrm{Im}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle)^2} \leq \|\overrightarrow{u}\| \|\overrightarrow{v}\|$$

Inoltre:

$$\|\overrightarrow{u} + \overrightarrow{v}\|^2 = \langle \overrightarrow{u} + \overrightarrow{v}; \overrightarrow{u} + \overrightarrow{v} \rangle = \|\overrightarrow{u}\|^2 + \|\overrightarrow{v}\|^2 + \langle \overrightarrow{u}; \overrightarrow{v} \rangle + \langle \overrightarrow{v}; \overrightarrow{u} \rangle =$$
$$\|\overrightarrow{u}\|^2 + \|\overrightarrow{v}\|^2 + 2\, \mathrm{Re}\, \langle \overrightarrow{u}; \overrightarrow{v} \rangle \leq \|\overrightarrow{u}\|^2 + \|\overrightarrow{v}\|^2 + 2 \|\overrightarrow{u}\| \|\overrightarrow{v}\| = (\|\overrightarrow{u}\| + \|\overrightarrow{v}\|)^2$$

Avendo a che fare con numeri non negativi, prendendo la radice quadrata segue subito l'enunciato.

3.3 Procedimento di ortogonalizzazione di Gram-Schmidt.

Dato un prodotto scalare o hermitiano, si dice **proiezione ortogonale** del vettore $\vec{v}$ sul vettore $\vec{u}$, il vettore:

$$P_{\vec{u}}(\vec{v}) = \frac{\langle \overrightarrow{v}; \overrightarrow{u} \rangle}{\langle \overrightarrow{u}; \overrightarrow{u} \rangle} \vec{u}$$

$P_{\vec{u}}(\vec{v})$ geometricamente rappresenta il vettore che si ottiene proiettando $\vec{v}$ lungo la direzione rappresentata dal vettore $\vec{u}$; è facile verificare che $\vec{v} - P_{\vec{u}}(\vec{v})$ è ortogonale al vettore $\vec{u}$.

Sia $\{\overrightarrow{v}_1; ...; \overrightarrow{v}_n\}$ una base di uno spazio vettoriale V. Troviamo per V una base **ortogonale (rispetto a un dato prodotto scalare o hermitiano)** $\{\overrightarrow{w}_1; ...; \overrightarrow{w}_n\}$, togliendo da ogni vettore $\overrightarrow{v}_i$ le sue proiezioni ortogonali rispetto ai vettori $\overrightarrow{w}_j$ con $j < i$.

$\overrightarrow{w}_1 = \overrightarrow{v}_1$;

$\overrightarrow{w}_2 = \overrightarrow{v}_2 - \frac{\langle \overrightarrow{v}_2; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1$;

$\overrightarrow{w}_3 = \overrightarrow{v}_3 - \frac{\langle \overrightarrow{v}_3; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1 - \frac{\langle \overrightarrow{v}_3; \overrightarrow{w}_2 \rangle}{\langle \overrightarrow{w}_2; \overrightarrow{w}_2 \rangle} \overrightarrow{w}_2$;

...

$\overrightarrow{w}_n = \overrightarrow{v}_n - \frac{\langle \overrightarrow{v}_n; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1 - \frac{\langle \overrightarrow{v}_n; \overrightarrow{w}_2 \rangle}{\langle \overrightarrow{w}_2; \overrightarrow{w}_2 \rangle} \overrightarrow{w}_2 - ... - \frac{\langle \overrightarrow{v}_n; \overrightarrow{w}_{n-1} \rangle}{\langle \overrightarrow{w}_{n-1}; \overrightarrow{w}_{n-1} \rangle} \overrightarrow{w}_{n-1}$.

E' semplice, ma noioso, verificare che i vettori ottenuti sono fra loro ortogonali, ad esempio:

$$\langle \overrightarrow{w}_1; \overrightarrow{w}_2 \rangle = \left\langle \overrightarrow{w}_1; \overrightarrow{v}_2 - \frac{\langle \overrightarrow{v}_2; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1 \right\rangle = \langle \overrightarrow{w}_1; \overrightarrow{v}_2 \rangle - \overline{\langle \overrightarrow{v}_2; \overrightarrow{w}_1 \rangle} = \langle \overrightarrow{w}_1; \overrightarrow{v}_2 \rangle - \langle \overrightarrow{w}_1; \overrightarrow{v}_2 \rangle = 0$$

- Dividendo poi ciascuno dei vettori $\{\overrightarrow{w}_1; ...; \overrightarrow{w}_n\}$ per la sua norma, si ottiene una base **ortonormale** di V, ovvero una base costituita da vettori di norma 1 e ortogonali a due a due: $\langle w_i; w_j \rangle = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}$. In forma compatta, si scrive $\langle w_i; w_j \rangle = \delta_{i;j}$, dove $\delta_{i;j}$ è il **simbolo di Kronecker** e vale

$$\delta_{i;j} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}.$$

Esempio.

Considero il sottospazio U di $\mathbb{R}^4$ generato da $\overrightarrow{v}_1 = (1; 1; 1; 1)$; $\overrightarrow{v}_2 = (1; 2; 4; 5)$; $\overrightarrow{v}_3 = (1; -3; -4; -2)$.

Trovo una base $\{\overrightarrow{w}_1; \overrightarrow{w}_2; \overrightarrow{w}_3\}$ ortogonale di U.

$\overrightarrow{w}_1 = \overrightarrow{v}_1 = (1; 1; 1; 1)$;

$\overrightarrow{w}_2 = \overrightarrow{v}_2 - \frac{\langle \overrightarrow{v}_2; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1 = (1; 2; 4; 5) - \frac{12}{4}(1; 1; 1; 1) = (-2; -1; 1; 2)$.

$\overrightarrow{w}_3 = \overrightarrow{v}_3 - \frac{\langle \overrightarrow{v}_3; \overrightarrow{w}_1 \rangle}{\langle \overrightarrow{w}_1; \overrightarrow{w}_1 \rangle} \overrightarrow{w}_1 - \frac{\langle \overrightarrow{v}_3; \overrightarrow{w}_2 \rangle}{\langle \overrightarrow{w}_2; \overrightarrow{w}_2 \rangle} \overrightarrow{w}_2 =$

$= (1; -3; -4; -2) - \frac{-8}{4}(1; 1; 1; 1) - \frac{-7}{10}(-2; -1; 1; 2) = \left(\frac{8}{5}; -\frac{17}{5}; -\frac{13}{10}; \frac{7}{5}\right)$.

Normalizziamo questa base dividendo ciasuno dei vettori per la rispettiva norma: alla fine, si trova

$\overrightarrow{u}_1 = \frac{1}{2}(1; 1; 1; 1)$; $\overrightarrow{u}_2 = \frac{1}{\sqrt{10}}(-2; -1; 1; 2)$; $\overrightarrow{u}_3 = \frac{1}{\sqrt{910}}(16; -17; -13; 14)$.

- **Teorema:**

Se V è uno spazio vettoriale con un dato prodotto scalare o hermitiano, e W è un suo sottospazio, allora $V = W \oplus W^{\perp}$.

Dimostrazione. Sia $\{\overrightarrow{w}_1; \overrightarrow{w}_2; ...; \overrightarrow{w}_r\}$ una base di W. Allora, è possibile, applicando il metodo di ortogonalizzazione di Gram-Schmidt, trasformarla in una base ortonormale di W, $\{\overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_r\}$. Ora, estendiamo quest'ultima ad una base di V,

$\{\overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_r; \overrightarrow{v}_{r+1}; \overrightarrow{v}_{r+2}; ...; \overrightarrow{v}_n\}$. Applicando di nuovo il procedimento di Gram-Schmidt, si ottiene una base ortogonale di tutto V, $\{\overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_r; \overrightarrow{u}_{r+1}; \overrightarrow{u}_{r+2}; ...; \overrightarrow{u}_n\}$. I primi vettori non sono modificati dall'ortogonalizzazione, perchè già erano ortogonali. I successivi $(n-r)$ vettori $\{\overrightarrow{u}_{r+1}; \overrightarrow{u}_{r+2}; ...; \overrightarrow{u}_n\}$ sono ortogonali a ciascuno dei $\{\overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v}_r\}$, quindi stanno in $W^\perp$. In questo modo, si è verificato che $V = W + W^\perp$.

Resta ancora da verificare che $W \cap W^\perp = \{0\}$. Sia $\overrightarrow{w} \in W \cap W^\perp$. Allora, $\langle \overrightarrow{w}; \overrightarrow{v} \rangle = 0$, $\forall \vec{v} \in W$. In particolare, $\langle \overrightarrow{w}; \overrightarrow{w} \rangle = 0$. Questo è possibile se e solo se $\overrightarrow{w} = 0$. Ne consegue che $W \cap W^\perp = \{0\}$.

4 Rette e Piani in $\mathbb{R}^n$.

Sia $\overrightarrow{v} \in \mathbb{R}^n$ e sia $\overrightarrow{x}_0$ un punto di $\mathbb{R}^n$.

- Si dice **retta di direzione $\overrightarrow{v}$ passante per $\overrightarrow{x}_0$** il sottoinsieme (è un sottospazio solo se passa per l'origine! cioè se $\overrightarrow{x}_0$ e $\overrightarrow{v}$ sono proporzionali):

$$r = \{\overrightarrow{x}_0 + t\overrightarrow{v} \text{ con } -\infty < t < \infty\}.$$

In $\mathbb{R}^2$, $\overrightarrow{v} = (v_x, v_y)$, $\overrightarrow{x}_0 = (x_0, y_0)$ e $r = (x, y)$. Le **equazioni parametriche** di r sono:

$$\begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \end{cases}$$

Eliminando il parametro t si ottengono le solite **equazioni cartesiane**.
Analogamente, le equazioni parametriche di una retta in $\mathbb{R}^3$ sono:

$$\begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z \end{cases}$$

- Si dice **piano in $\mathbb{R}^3$ di direzioni $\overrightarrow{v}$ e $\overrightarrow{w}$ passante per $\overrightarrow{x}_0$** il sottoinsieme

$$P = \{\overrightarrow{x}_0 + t\overrightarrow{v} + s\overrightarrow{w} \text{ con } -\infty < t, s < \infty\},$$

con $\overrightarrow{x}_0, \overrightarrow{v}, \overrightarrow{w} \in \mathbb{R}^3$ e $\overrightarrow{v}, \overrightarrow{w}$ **indipendenti.** Le equazioni parametriche di un piano sono date da:

$$\begin{cases} x = x_0 + tv_x + sw_x \\ y = y_0 + tv_y + sw_y \\ z = z_0 + tv_z + sw_z \end{cases}$$

- Si dice **m-piano di $\mathbb{R}^n$ di direzioni** $\overrightarrow{v}_1, ..., \overrightarrow{v}_m$ ($m < n$ e indipendenti) **passante per** $\overrightarrow{x}_0$ il sottoinsieme

$$P = \left\{\overrightarrow{x}_0 + \sum_{i=1}^{m} t_i \overrightarrow{v}_i \text{ con } -\infty < t < \infty\right\}.$$

- Consideriamo ora $\mathbb{R}^n$ con il **prodotto scalare standard**;

Diciamo che le rette $r = \overrightarrow{p} + t\overrightarrow{v}$ e $r' = \overrightarrow{q} + t\overrightarrow{v}'$ sono:
parallele se $\overrightarrow{v} = a\overrightarrow{v}'$ con $a \neq 0$ (e, naturalmente, le due rette non coincidono).
ortogonali se $\overrightarrow{v} \cdot \overrightarrow{v}' = 0$;
incidenti se $r \cap r' \neq \emptyset$;
complanari se $\exists$ un 2-piano π tale che $r \subset \pi$ e $r' \subset \pi$.
Esempio.
Consideriamo le rette in $\mathbb{R}^3$

$$\alpha = (0,0,1) + t\,(1,-1,0) = (0,0,1) + t\vec{v}$$
$$\beta = (1,0,0) + t\,(1,-1,0) = (1,0,0) + t\vec{u}$$
$$\gamma = (0,1,0) + t\,(0,0,1) = (0,1,0) + t\vec{w}$$

Poichè $\vec{v} = \vec{u}$, α e β sono parallele. Inoltre $\vec{u} \cdot \vec{v} = (0,0,1) \cdot (1,-1,0) = 0$, quindi γ è ortogonale a α.

Vediamo se α e γ sono incidenti. Ci chiediamo: quando un generico $P \in \alpha$ è uguale a $P' \in \gamma$? Poniamo:

$$(0,0,1) + t\,(1,-1,0) = (0,1,0) + s\,(0,0,1)$$
$$\Rightarrow \quad (t,-t,1) = (0,1,s)$$
$$\Rightarrow \quad t = 0 \text{ e } t = 1 \Rightarrow \ \textbf{impossibile!}$$

Le rette α e γ non sono incidenti. Troviamo una retta δ parallela a γ e incidente con α. Poichè deve essere δ parallela a γ, si avrà

$$\delta = (a,b,c) + t(0,0,1).$$

Per trovare (a,b,c), imponiamo che δ sia incidente con α:

$$(t,-t,1) = (a,b,c+t).$$

Prendiamo $t = 0$. Abbiamo

$$(0,0,1) = (a,b,c) \Rightarrow \delta = (0,0,1) + t\,(0,0,1).$$

Vogliamo cercare ora un piano che contenga le due rette parallele α e β. Sia

$$\pi = (a,b,c) + tv_1 + sv_2.$$

π deve contenere $\alpha \Rightarrow \pi = (0,0,1) + t\,(1,-1,0) + s\,(a,b,c)$ infatti per $s = 0$, π si riduce alla retta α.

Per trovare (a,b,c), imponiamo che contenga β.

$$(1,0,0) + t'\,(1,-1,0) = (0,0,1) + t\,(1,-1,0) + s\,(a,b,c)$$
$$\Rightarrow (1,0,0) = (0,0,1) + s\,(a,b,c) \text{ e } t = t'$$

Fissiamo un s possibile, per esempio $s = 1$:

$$(1,0,0) = (a, b, c+1) \Rightarrow a = 1, b = 0, c = -1$$
$$\Rightarrow \pi = (0,0,1) + t\,(1,-1,0) + s\,(1,0,-1).$$

Infatti il piano π per $s = 0$ si riduce alla retta α, mentre per $s = 1$ si riduce alla retta β.

Osservazioni.

Due rette incidenti e distinte sono complanari.

Due rette parallele sono complanari.

Due rette non incidenti sono complanari solo se parallele.

Esempio.

Le rette α e δ dell'esempio precedente sono incidenti, quindi complanari. Troviamo il piano che le contiene.

$$\alpha = (0,0,1) + t\,(1,-1,0)$$
$$\delta = (0,0,1) + t\,(0,0,1).$$

Il piano che le contiene è il piano generato dai vettori direzione delle due rette:

$$\pi = (0,0,1) + t\,(1,-1,0) + s(0,0,1).$$

- Due $2-$ piani π e $\acute{\pi}$ sono **paralleli** se $\acute{\pi} = \pi + q$, $q \neq (0,0,...,0)$. Se $\pi = p + tv_1 + sv_2$ e $\acute{\pi} = \acute{p} + t\acute{v}_1 + s\acute{v}_2$, allora π e $\acute{\pi}$ sono paralleli se:

$$\text{Span}\{v_1, v_2\} = \text{Span}\{\acute{v}_1, \acute{v}_2\} \text{ e } p - \acute{p} \notin \text{Span}\{v_1, v_2\}.$$

La stessa definizione vale per gli $m-$ piani.

Esempio.

Consideriamo i piani

$$\pi = (0,0,0) + t(1,0,0) + s(1,0,1)$$
$$\acute{\pi} = (0,1,0) + t'(1,0,0) + s'(0,0,1).$$

Si ha

$$\text{Span}\{v_1, v_2\} = \{v = av_1 + bv_2 = a(1,0,0) + b(1,0,1) = (a+b, 0, b)\}$$
$$\text{Span}\{\acute{v}_1, \acute{v}_2\} = \{v = \acute{a}\acute{v}_1 + \acute{b}\acute{v}_2 = \acute{a}(1,0,0) + \acute{b}(0,0,1) = (\acute{a}, 0, \acute{b})\}$$

I due spazi coincidono: basta prendere $\acute{b} = b$, $\acute{a} = a + b$, quindi i due piani sono paralleli (non coincidono perchè uno passa per l'origine e l'altro no).

Invece,

$$\pi = (0,0,0) + t(1,0,0) + s(1,0,1)$$
$$\acute{\pi} = (0,0,1) + \acute{t}(1,0,0) + \acute{s}(0,1,0)$$

non sono paralleli e $\pi \cap \acute{\pi} = r$. Infatti, i punti in comune si trovano imponendo $\pi = \acute{\pi}$:

$$(t+s,0,s) = (t',s',1),$$

cioè $s'=0$, $s=1$, $t'=t+1$. Allora i punti del tipo $(t+1,0,1) \in \pi \cap \pi'$, cioè la retta

$$r = (1,0,1) + t(1,0,0)$$

è l'intersezione dei due piani.

- Due piani di $\mathbb{R}^3$, $\pi = p+tv_1+sv_2$ e $\pi' = q+t'v'_1+s'v'_2$ sono **ortogonali** se le loro **direzioni normali** (cioè le direzioni delle rette ortogonali ai piani) sono fra loro ortogonali.

Esempio.
Dato

$$\pi = (0,0,0) + t(1,0,0) + s(1,0,1),$$

con $\vec{v}_1 = (1,0,0)$ e $\vec{v}_2 = (1,0,1)$, calcoliamo la direzione $\overrightarrow{n}$ normale a π. Posto $\overrightarrow{n} = (a,b,c)$, deve essere

$$(a,b,c)\cdot(1,0,0) = a = 0$$
$$(a,b,c)\cdot(1,0,1) = a+c = 0 \Rightarrow c = 0,$$

quindi $\overrightarrow{n} = t(0,1,0)$. Consideriamo, inoltre, il piano

$$\pi' = (0,0,0) + t'(0,1,0) + s'(2,0,1).$$

Calcoliamo la sua normale $\overrightarrow{n'}$:

$$(a',b',c')\cdot(0,1,0) = b' = 0$$
$$(a',b',c')\cdot(2,0,1) = 2a'+c' = 0 \Rightarrow c' = -2a'$$
$$\Rightarrow \overrightarrow{n'} = t'(-1,0,2).$$

Ora $\overrightarrow{n}\cdot\overrightarrow{n'} = (0,1,0)\cdot(-1,0,2) = 0 \Rightarrow \overrightarrow{n} \perp \overrightarrow{n'} \Rightarrow \pi,\pi'$ ortogonali.
Esempio.
Dato il piano

$$\pi = (0,0,0) + t(1,0,0) + s(1,0,1),$$

trovare la retta passante per $(0,0,0)$ e ortogonale al piano π.
La retta cercata ha equazione

$$r = (0,0,0) + t(a,b,c)$$

e la direzione (a,b,c) deve essere ortogonale a $(1,0,0)$ e a $(1,0,1)$. Si deve quindi avere

$$(1,0,0)\cdot(a,b,c) = a = 0,$$
$$(1,0,1)\cdot(a,b,c) = a+c = 0 \Rightarrow c = 0.$$

Allora $r = (0,0,0) + t(0,1,0)$.
Determiniamo, visto che in questo caso la retta e il piano sono anche sottospazi, $\pi \oplus r$. Osserviamo che $(1,0,0)$, $(1,0,1)$ e $(0,1,0)$ sono tre vettori indipendenti, per cui: $\pi \oplus r = \mathbb{R}^3$.

5 Matrici e calcolo matriciale.

Consideriamo l'insieme:

$$M(n \times m) = \left\{ \begin{array}{c} \text{tabelle di } nm \text{ numeri (reali o complessi)} \\ \text{ordinati in } n \text{ righe e in } m \text{ colonne} \end{array} \right\}$$

$$= \{\textbf{matrici a } n \textbf{ righe ed } m \textbf{ colonne}\}$$

Rappresentiamo una matrice come una tabella della forma:

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix},$$

la indichiamo con A e scriviamo $A = (a_{ij})$, dove a_{ij} è il generico elemento di A contrassegnato da un indice di riga i e un indice di colonna j. Se $m = n$, la matrice si dice **quadrata di ordine** n. Se $A, B \in M(n \times m)$ e $\lambda \in \mathbb{R}$ o $\mathbb{C}$, si possono definire:

1) $A + B = C \in M(n \times m)$, sommando gli elementi nello stesso posto:

$$c_{ij} = a_{ij} + b_{ij}$$

2) $\lambda A \in M(n \times m)$, con la formula:

$$(\lambda A)_{ij} = \lambda a_{ij}.$$

Se $A \in M(n \times m)$ e $B \in M(m \times q)$, si definisce il prodotto tra matrici

$$AB = C \in M(n \times q)$$

con il **prodotto righe per colonne**: $(AB)_{ik}$, cioè l'elemento di posto i, k della matrice AB, se gli elementi della matrice sono numeri reali, è il **prodotto scalare standard** della i-esima riga di A, $(a_{i1}a_{i2}...a_{im})$, per la k-esima colonna di B, $\begin{pmatrix} b_{1k} \\ \vdots \\ b_{mk} \end{pmatrix}$:

$$(AB)_{ik} = (C)_{ik} = (a_{i1}a_{i2}...a_{im}) \begin{pmatrix} b_{1k} \\ \vdots \\ b_{mk} \end{pmatrix} = a_{i1}b_{1k} + a_{i2}b_{2k} + ... + a_{im}b_{mk}.$$

Il generico elemento di C è, quindi:

$$(AB)_{ij} = c_{ij} = \sum_m a_{im} b_{mj}.$$

La stessa formula si usa nel caso di matrici complesse.

Notiamo che A e B possono essere moltiplicate tra loro solo se

numero colonne di A = numero righe di B.

Due matrici quadrate dello stesso ordine si possono sempre moltiplicare.

ESEMPIO.

(2×2)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}.$$

OSSERVAZIONE **(importante!)**.

Se si possono fare sia AB che BA, in generale si ha che $AB \neq BA$. Inoltre, $AB = O$, dove $O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, **non** implica che A o B siano zero:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

- La **matrice trasposta** A^t di A è la matrice ottenuta da A scambiando le righe con le colonne. Se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, allora $A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. In generale, $(A^t)_{ik} = a_{ki}$.

Proprietà immediate della matrice trasposta:

1) $(A+B)^t = A^t + B^t$;

2) $(AB)^t = B^t A^t$.

3) $(A^t)^t = A$.

Se $A = (a_{ij})$ è una matrice quadrata, gli elementi $a_{11}, a_{22}, ..., a_{nn}$ costituiscono la **diagonale principale** di A. Una matrice si dice **diagonale** se gli elementi fuori della diagonale principale sono nulli.

Consideriamo matrici **quadrate** (per ora 2×2). La matrice diagonale con tutti 1 sulla diagonale principale è la **matrice identità** $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ed ha le stesse proprietà del numero 1, cioè:

$$AI = IA = A.$$

La matrice identità si definisce analogamente per ogni ordine.

- Sappiamo che se $a \in \mathbb{R}$ e $a \neq 0$, possiamo definire un inverso a^{-1} tale che $aa^{-1} = 1$. Nel caso delle matrici, non basta che $A \neq O$ per poter definire una A^{-1} tale che

$$AA^{-1} = A^{-1}A = I.$$

Serve la nozione di **determinante** della matrice A ($\det A$). Nel caso 2×2, si ha

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Si vede subito che $\det(AB) = \det A \det B$. Se $\det A \neq 0$, si può definire A^{-1}:

$$\text{se } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ si ha } A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Si verifica immediatamente che $AA^{-1} = A^{-1}A = I$.

- **Proprietà dell'inversa**:

$a)$ $\det A^{-1} = \dfrac{1}{\det A}$;
$b)$ $(A^{-1})^{-1} = A$;
$c)$ $(AB)^{-1} = B^{-1}A^{-1}$;
$d)$ $(A^t)^{-1} = (A^{-1})^t$.
Si noti anche che $\det A^t = \det A$.

- Per le matrici $n \times n$, il **determinante** si definisce ricorsivamente nel seguente modo: se

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

allora

$$\det A = \sum_{p=1}^{n} (-1)^{p+1} a_{1p} \det(A_{1p})$$

dove A_{1p} è la sottomatrice di A ottenuta da A eliminando la prima riga e la p-esima colonna. Anche qui $\det A^t = \det A$ e $\det AB = \det A \det B$.

ESEMPI.

1) Consideriamo la matrice 2×2

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

Calcoliamo $\det A = \sum_{p=1}^{2} (-1)^{p+1} a_{1p} \det A_{1p}$. Si ha: $A_{11} = a_{22}$, $A_{12} = a_{21}$, per cui

$$\det A = (-1)^2 a_{11} \det A_{11} + (-1)^3 a_{12} \det A_{12} = a_{11}a_{22} - a_{12}a_{21}$$

(analoga alla definizione data precedentemente).
2) Consideriamo la matrice 3×3

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

Si ha:

$$A_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix},\ A_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix},\ A_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$
$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}.$$

- Formula generale per l'inversa: sia A_{pq} la sottomatrice ottenuta eliminando da A la p-esima riga e la q-esima colonna. Si definisce **complemento** $\tilde{a}_{ij}$ dell'elemento a_{ij} di A:

$$\tilde{a}_{ij} = (-1)^{i+j} \det A_{ij}.$$

Allora

$$(A^{-1})_{ij} = \frac{1}{\det A} \tilde{a}_{ji} \text{ (notare lo scambio: } ji \textbf{ e non } ij).$$

Applicazione al caso 2×2. Se $\tilde{A} = (\tilde{a}_{ij})$ è la **matrice dei complementi**, la formula precedente si riscrive $A^{-1} = \frac{1}{\det A} \tilde{A}^t$. Allora se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, si ha

$$\tilde{a}_{11} = d,\ \ \tilde{a}_{22} = a,\ \ \tilde{a}_{12} = -c,\ \ \tilde{a}_{21} = -b,$$

per cui

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix},\ \ \tilde{A}^t = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

quindi: $A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.
NOTA IMPORTANTE.
$\tilde{a}_{ij}$ **serve anche a calcolare il determinante usando la riga che è più comoda (cioè quella con più zeri):**

$$\det A = \sum_j a_{ij}\tilde{a}_{ij} \qquad \textbf{(dove } i \textbf{ è scelto da noi).}$$

ESEMPIO.

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

Calcoliamo il determinante scegliendo la 3^a riga, $i = 3$.

$$\tilde{a}_{31} = \det\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = -2, \; \tilde{a}_{32} = -\det\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = 2,$$
$$\tilde{a}_{33} = \det\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0.$$
$$\det A = \sum_j a_{3j}\tilde{a}_{3j} = 0 \cdot (-2) + 0 \cdot 2 + 2 \cdot 0 = 0.$$

- **Osservazioni utili per il calcolo del determinante.**

1.) Se una matrice ha una riga (oppre una colonna) interamente costituita da zeri, allora il determinante è 0.

ESEMPIO.

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

Si vede subito che la seconda riga è nulla, quindi il determinante è 0, come si vede immediatamente usando la seconda riga per calcolarlo.

2.) Se una matrice ha due righe (oppure due colonne) identiche, allora il determinante è 0.

ESEMPIO

$$A = \begin{pmatrix} 5 & 6 & 7 & 6 \\ 1 & -3 & 5 & -3 \\ 4 & 9 & -3 & 9 \\ 2 & 7 & 8 & 7 \end{pmatrix}$$

La seconda e la quarta colonna sono uguali, dunque il determinante è 0.

3.) Se la matrice è triangolare, il determinante è semplicemente il prodotto degli elementi sulla diagonale principale.

ESEMPIO

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

$det\,(A) = (2) \cdot (-3) \cdot (5) \cdot 4 = -120.$

4.) Se ad una riga (oppure ad una colonna) aggiungiamo una combinazione lineare delle altre $(n-1)$ righe (oppure colonne), il determinante non viene modificato. Questo può essere utile per calcolare "a mano" il determinante di matrici molto grosse.

ESEMPIO

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

Osserviamo che la seconda riga ha già due zeri: se ci fosse un terzo zero, il calcolo del determinante sarebbe più semplice. Possiamo però scrivere una matrice A' avente lo stesso determinante di A ma con tre zeri sulla seconda riga. Infatti, siano $\{C_1; C_2; C_3; C_4\}$ le colonne. Se a C_3 aggiungiamo $2C_1$, il determinante non cambia, e otteniamo una nuova matrice A' avente lo stesso determinante di A.

$$A' = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 1 & 0 & 0 & 0 \\ 3 & -1 & 7 & -2 \\ 4 & -3 & 8 & 2 \end{pmatrix}.$$

$$det\,(A') = (-1)^3 \cdot 1 \cdot det \begin{pmatrix} 2 & 4 & 3 \\ -1 & 7 & -2 \\ -3 & 8 & 2 \end{pmatrix} = (-1) \cdot \left[1 \cdot 2 \cdot det \begin{pmatrix} 7 & -2 \\ 8 & 2 \end{pmatrix} \right] +$$
$$+ (-1) \cdot \left[(-1) \cdot 4 \cdot det \begin{pmatrix} -1 & -2 \\ -3 & 2 \end{pmatrix} \right] + (-1) \cdot \left[1 \cdot 3 \cdot det \begin{pmatrix} -1 & 7 \\ -3 & 8 \end{pmatrix} \right] =$$
$$= (-1) \cdot [60] + (-1) \cdot [32] + (-1) \cdot [39] = -131.$$

5.) Se una singola riga (oppure una singola colonna) di una matrice viene moltiplicata per un numero λ, allora il determinante della matrice risulta esso pure moltiplicato per λ. In particolare, data una matrice $n \times n$, se essa viene moltiplicata per λ, ogni sua riga è moltiplicata per tale numero. Alla fine, il determinante risulterà essere moltiplicato per λ^n.

- ALGORITMO DI GAUSS (PER CALCOLARE L'INVERSA.)

Sia A una matrice quadrata $n \times n$. Eseguiamo i seguenti passi:

- Formare la matrice $\begin{pmatrix} A & | & I \end{pmatrix}$, dove I è la matrice $n \times n$ identità.
- Procedere con operazioni elementari tra le righe finchè non si ottiene $\begin{pmatrix} I & | & B \end{pmatrix}$.
- Porre $B = A^{-1}$.
- ATTENZIONE: se nel procedimento si genera nella metà di sinistra una riga tutta nulla, allora A non è invertibile.

ESEMPIO

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

Abbiamo:

$M = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 2 & -1 & 3 & | & 0 & 1 & 0 \\ 4 & 1 & 8 & | & 0 & 0 & 1 \end{pmatrix}$. Siano R_1; R_2; R_3 le righe di M. Eseguiamo le seguenti operazioni: $R_2 \longmapsto R_2 - 2R_1$; $R_3 \longmapsto R_3 - 4R_1$.

Troviamo:

$M' = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & -1 & -1 & | & -2 & 1 & 0 \\ 0 & 1 & 0 & | & -4 & 0 & 1 \end{pmatrix}$. Siano R_1'; R_2'; R_3' le righe di M'. Eseguiamo la seguente operazione: $R_3 \longmapsto R_3 + R_2$.

Troviamo:

$M'' = \begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & -1 & -1 & | & -2 & 1 & 0 \\ 0 & 0 & -1 & | & -6 & 1 & 1 \end{pmatrix}$. Cambiamo segno alla seconda riga.

$\begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 2 & -1 & 0 \\ 0 & 0 & -1 & | & -6 & 1 & 1 \end{pmatrix}$. Siano R_1''; R_2''; R_3'' le righe. Eseguiamo le seguenti operazioni: $R_1'' \longmapsto R_1'' + 2R_3''$; $R_2'' \longmapsto R_2'' + R_3''$.Troviamo:

$\begin{pmatrix} 1 & 0 & 0 & | & -11 & 2 & 2 \\ 0 & 1 & 0 & | & -4 & 0 & 1 \\ 0 & 0 & -1 & | & -6 & 1 & 1 \end{pmatrix}$. Cambiamo segno alla terza riga. Alla fine si ottiene

$\begin{pmatrix} 1 & 0 & 0 & | & -11 & 2 & 2 \\ 0 & 1 & 0 & | & -4 & 0 & 1 \\ 0 & 0 & 1 & | & 6 & -1 & -1 \end{pmatrix}$.

E quindi: $A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$.

- Rango di una matrice. Sia A una matrice $n \times m$. Si chiama **minore** di A una sottomatrice **quadrata** contenuta in A. L'**ordine** del minore è l'ordine della sottomatrice quadrata.

ESEMPIO.
Data la matrice

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

i minori sono:
ordine 1: $1, 0, 2, 3$;
ordine 2: $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$;
ordine 3: $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Si dice **rango** di A e si indica con $r(A)$, l'ordine (cioè il numero di righe e di colonne) **massimo** del minore di A con determinante $\neq 0$.

ESEMPIO.
$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ ha rango 1; $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ ha rango 1;
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ha rango 2; $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ha rango 0;
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ ha rango 2; $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ ha rango 2.

- Ovviamente $r(A) \leqslant \min\{n, m\}$. Una matrice si dice di rango massimo se è raggiunta l'uguaglianza; se la matrice è quadrata, è di rango massimo se ha determinante diverso da zero.

Proprietà essenziale. Il rango di A è il numero massimo di colonne linearmente indipendenti (o di righe linearmente indipendenti, perchè $r(A) = r(A^t)$).

ESEMPIO.
$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ha rango 2, perchè

$$\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Infatti il determinante della matrice è uguale a zero (verificare) e il minore $\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, che ha ordine 2 ha determinante diverso da zero.

Teorema. Sia $AB = C$. Allora se A *è invertibile* $r(AB) = r(B)$; se A *e* B *sono invertibili* $r(AB) = r(A) = r(B)$.

- Notazioni matriciali dei vettori e delle basi. (**Convenzione di Einstein**) Una matrice può anche essere costituita da oggetti più generali dei numeri reali, basta che sugli oggetti siano definite operazioni che danno senso a prodotti e somme; ad esempio, considereremo matrici di numeri complessi e anche di vettori.

Consideriamo la base $e = (\overrightarrow{e}_1, \overrightarrow{e}_2, ..., \overrightarrow{e}_n)$, il vettore $\overrightarrow{v}$ e il vettore colonna delle componenti di $\overrightarrow{v}$: $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$. Allora $\overrightarrow{v} = \sum_{i=1}^{n} v_i \overrightarrow{e}_i$. Si può scrivere, utilizzando la regola formale del prodotto righe $\times$ colonne:

$$\overrightarrow{v} = \sum_{i=1}^{n} v_i \overrightarrow{e} = (\overrightarrow{e}_1, \overrightarrow{e}_2, ..., \overrightarrow{e}_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = ev$$

ed anche, per il prodotto scalare usuale:

$$\overrightarrow{v} \cdot \overrightarrow{w} = v^t w = (v_1, ..., v_n) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v_1 w_1 + ... + v_n w_n.$$

6 Applicazioni lineari.

Un'applicazione $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ si dice **lineare** se $\forall \overrightarrow{v}, \overrightarrow{w} \in \mathbb{R}^n$, $\forall a, b \in \mathbb{R}$, si ha

$$\varphi (a\overrightarrow{v} + b\overrightarrow{w}) = a\varphi (\overrightarrow{v}) + b\varphi (\overrightarrow{w})$$

(con le ovvie modifiche si tratta il caso complesso).

La stessa definizione si usa per definire più in generale le applicazioni lineari $\varphi : V \to W$, **con** V **e** W **spazi vettoriali su un campo** $\mathbb{K}$: $\forall \overrightarrow{v}, \overrightarrow{u} \in V$, $\forall a, b \in \mathbb{K}$, deve essere:

$$\varphi (a\overrightarrow{v} + b\overrightarrow{u}) = a\varphi (\overrightarrow{v}) + b\varphi (\overrightarrow{u})$$

Consideriamo, ad esempio, il caso $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$. La proprietà di linearità ci dice che **per conoscere** φ **basta sapere come opera sui vettori della base**. Infatti, supponiamo di usare la stessa base $e = (\overrightarrow{e}_1, \overrightarrow{e}_2)$ sia nell'$\mathbb{R}^2$ di partenza, sia in quello di arrivo. Se $\overrightarrow{v} = v_1\overrightarrow{e}_1 + v_2\overrightarrow{e}_2 = (\overrightarrow{e}_1, \overrightarrow{e}_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, allora

$$\varphi(\overrightarrow{v}) = v_1\varphi(\overrightarrow{e}_1) + v_2\varphi(\overrightarrow{e}_2)$$

quindi basta dare

$$\varphi(\overrightarrow{e}_1) = a\overrightarrow{e}_1 + c\overrightarrow{e}_2, \quad \varphi(\overrightarrow{e}_2) = b\overrightarrow{e}_1 + d\overrightarrow{e}_2.$$

Allora φ ha una rappresentazione matriciale:

$$\varphi(\overrightarrow{v}) = (\overrightarrow{e}_1, \overrightarrow{e}_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (a\overrightarrow{e}_1 + c\overrightarrow{e}_2, b\overrightarrow{e}_1 + d\overrightarrow{e}_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

cioè

$$\begin{aligned} \varphi(\overrightarrow{v}) &= (a\overrightarrow{e}_1 + c\overrightarrow{e}_2)\, v_1 + (b\overrightarrow{e}_1 + d\overrightarrow{e}_2)\, v_2 = \\ &= (av_1 + bv_2)\,\overrightarrow{e}_1 + (cv_1 + dv_2)\,\overrightarrow{e}_2 = \begin{pmatrix} av_1 + bv_2 \\ cv_1 + dv_2 \end{pmatrix}, \end{aligned}$$

ovvero, in una base qualsiasi $e = (\overrightarrow{e}_1, \overrightarrow{e}_2)$, $\varphi(\overrightarrow{v}) = eAv$, dove A **è la matrice che rappresenta la** φ e si costruisce:

1) mettendo in **colonna** i trasformati dei vettori della base

$$\varphi(\overrightarrow{e}_1) = (\overrightarrow{e}_1, \overrightarrow{e}_2)\begin{pmatrix} a \\ c \end{pmatrix}, \varphi(\overrightarrow{e}_2) = (\overrightarrow{e}_1, \overrightarrow{e}_2)\begin{pmatrix} b \\ d \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

2) oppure mettendo in **riga** i trasformati delle coordinate del vettore $\overrightarrow{v}$

$$\varphi\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} av_1 + bv_2 \\ cv_1 + dv_2 \end{pmatrix} \Rightarrow A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Nel caso genearale, sia

$$f : V \to W$$

una applicazione lineare da V, spazio vettoriale di dimensione n a W spazio vettoriale di dimensione m. Indichiamo con $e = (\overrightarrow{e}_1, \overrightarrow{e}_2, ..., \overrightarrow{e}_n)$ una base di V e con $u = (\overrightarrow{u}_1, \overrightarrow{u}_2, ..., \overrightarrow{u}_m)$ una base di W. Un vettore $\overrightarrow{v} = \sum_{i=1}^{n} v_i\overrightarrow{e}_i = ev$, dove $v = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ v_n \end{pmatrix}$, è trasformato da f nel vettore $\vec{w}$:

$$\vec{w} = f(\vec{v}) = f\left(\sum_{i=1}^{n} v_i\overrightarrow{e}_i\right) = \sum_{i=1}^{n} v_i f(\overrightarrow{e}_i)$$

Posto:

$$f\left(\overrightarrow{e}_i\right) = \sum_{k=1}^{m} a_{ki}\overrightarrow{u}_k$$

Ovvero:

$$\begin{aligned}
f\left(\overrightarrow{e}_1\right) &= a_{11}\overrightarrow{u}_1 + a_{21}\overrightarrow{u}_2 + ... + a_{m1}\overrightarrow{u}_m \\
f\left(\overrightarrow{e}_2\right) &= a_{12}\overrightarrow{u}_1 + a_{22}\overrightarrow{u}_2 + ... + a_{m2}\overrightarrow{u}_m \\
&\dots\dots\dots\dots\dots \\
f\left(\overrightarrow{e}_n\right) &= a_{1n}\overrightarrow{u}_1 + a_{2n}\overrightarrow{u}_k + ... + a_{mn}\overrightarrow{u}_m
\end{aligned}$$

Otteniamo:

$$f(\vec{v}) = \sum_{k=1,i=1}^{m,n} a_{ki}v_i\overrightarrow{u}_k$$

D'altra parte, se consideriamo la matrice a m righe e n colonne A, costituita con i coefficienti a_{ki}, è immediato verificare, prendendo la regola formale del prodotto righe $\times$ colonne:

$$u\,(Av) = \left(\overrightarrow{u}_1, \overrightarrow{u}_2, ..., \overrightarrow{u}_m\right)\begin{pmatrix} a_{11} & a_{12} & . & a_{1n} \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & . & a_{mn} \end{pmatrix}\begin{pmatrix} v_1 \\ . \\ . \\ v_n \end{pmatrix} = f(\vec{v})$$

Abbiamo quindi generalizzato l'associazione di una matrice ad una applicazione lineare, *data una base del dominio V e una base del codominio W.*

Viceversa, ogni matrice, scelta una base nel dominio e una nel codominio, dà un'applicazione lineare (per vedere che ciò ha senso, si deve parlare di trasformazione di basi, come faremo in seguito).

- Se $f : V \to V$ è iniettiva e suriettiva si dice che f è un **isomorfismo** e se A è la matrice associata a f in una certa base, allora A è **invertibile** e A^{-1} è la matrice associata a f^{-1} nella stessa base. $(w = Av \Rightarrow v = A^{-1}w)$.

- E' chiaro che **due spazi vettoriali V e W della stessa dimensione finita sono isomorfi,** perchè se $\left(\overrightarrow{e}_1, \overrightarrow{e}_2, ..., \overrightarrow{e}_n\right)$ è una base di V e $\left(\overrightarrow{u}_1, \overrightarrow{u}_2, ..., \overrightarrow{u}_n\right)$ è una base di W, costruiamo l'applicazione lineare $\varphi : V \to W$ definita sui vettori delle basi da: $\varphi\left(\overrightarrow{e}_1\right) = \vec{u}_1, \varphi\left(\overrightarrow{e}_2\right) = \vec{u}_2, ..., \varphi\left(\overrightarrow{e}_n\right) = \vec{u}_n$; è immediato verificare che è iniettiva e suriettiva (e quindi è un isomorfismo) e che, nelle basi indicate, la matrice associata è I.

- Se $\varphi : V \to W$ è lineare, consideriamo due sottoinsiemi interessanti (verificheremo poi che sono sottospazi):

1) il **nucleo** di φ:

$$\ker\varphi = \left\{\overrightarrow{v} \in V \text{ tali che } \varphi\left(\overrightarrow{v}\right) = \overrightarrow{0}\right\}$$

2) l'**immagine** di φ:

$$\operatorname{Im}\varphi = \{\overrightarrow{w} \in W \text{ tali che esiste } \overrightarrow{v} \in V \text{ per cui } \varphi(\overrightarrow{v}) = \overrightarrow{w}\}.$$

Se $\ker\varphi = \left\{\overrightarrow{0}\right\}$ (cioè nel nucleo c'è solo il vettore nullo) , allora φ è iniettiva, cioè vettori distinti vanno in vettori distinti. Infatti, se due vettori $\overrightarrow{v}_1$ e $\overrightarrow{v}_2$ hanno la stessa immagine, $\varphi(\overrightarrow{v}_1) = \varphi(\overrightarrow{v}_2) \Rightarrow \varphi(\overrightarrow{v}_1 - \overrightarrow{v}_2) = \overrightarrow{0} \Rightarrow \overrightarrow{v}_1 - \overrightarrow{v}_2 \in \ker\varphi = \left\{\overrightarrow{0}\right\} \Rightarrow \overrightarrow{v}_1 - \overrightarrow{v}_2 = \overrightarrow{0} \Rightarrow \overrightarrow{v}_1 = \overrightarrow{v}_2$.

Inoltre, se $\varphi : V \to V$ (quindi $m = n$) è iniettiva, presa una base di V, $(\overrightarrow{e}_1, \overrightarrow{e}_2, ..., \overrightarrow{e}_n)$ possiamo allora prendere $(\varphi(\overrightarrow{e}_1), \varphi(\overrightarrow{e}_2), ..., \varphi(\overrightarrow{e}_n))$ come base di $\operatorname{Im}\varphi$ che ha quindi dimensione n; allora $\operatorname{Im}\varphi$ e V sono isomorfi perchè hanno la stessa dimensione: una applicazione lineare $\varphi : V \to V$ se è iniettiva è anche suriettiva.

Come sopra si dice che φ è un **isomorfismo** e si ha che la matrice A associata a φ (prendendo ora la stessa base nel dominio e nell'immagine) è **invertibile**: $w = Av \Rightarrow v = A^{-1}w$.

- **Gli isomorfismi danno i cambiamenti di base**, cioè trasformano vettori indipendenti in vettori indipendenti in modo biunivoco. Se abbiamo $\overrightarrow{e}_1, ..., \overrightarrow{e}_n$ indipendenti e in numero "giusto", cioè n, gli $\overrightarrow{e'}_i = \varphi(\overrightarrow{e}_i)$ sono ancora indipendenti (verificare !) ed in numero "giusto". Quindi, se e ed e' sono due basi di V, dato un vettore $\overrightarrow{v}$, si può scrivere $\overrightarrow{v} = ev$, ma anche $\overrightarrow{v} = e'v'$. Allora, se $e' = eB$ (attenzione, si esegue sempre il prodotto righe $\times$ colonne e quindi si deve scrivere in questo ordine) con B associata a un isomorfismo, e quindi invertibile, segue:

$$ev = \overrightarrow{v} = e'v' = eBv' \Rightarrow v = Bv' \Rightarrow v' = B^{-1}v.$$

Inoltre se $\varphi : V \to V$, è una applicazione lineare, si ha:

$$\varphi(\overrightarrow{v}) = eAv = e'A'v' \Rightarrow eAv = eBA'B^{-1}v \Rightarrow A = BA'B^{-1} \Rightarrow \boxed{A' = B^{-1}AB}$$

cioè se B è la matrice associata al cambiamento di base, le componenti dei vettori trasformano con B^{-1}, mentre le matrici delle applicazioni lineari trasformano come $B^{-1}AB$.

NOTA.

$\det A' = \det A$, perchè $\det A' = \det B^{-1} \det A \det B = \det A$ e rango(A') = rango(A), perchè $r(B^{-1}AB) = r(AB) = r(A)$, essendo B e B^{-1} invertibili. Allora due matrici che in due diverse basi rappresentano la **stessa** applicazione lineare hanno uguale determinante e uguale rango. **Non è vero il viceversa.**

ESEMPIO. Consideriamo $\mathbb{R}^2$ con base $\overrightarrow{e}_1, \overrightarrow{e}_2$ e il vettore $\overrightarrow{v} = ev = (\overrightarrow{e}_1, \overrightarrow{e}_2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ vogliamo vedere cosa diventano le sue componenti nella base $(\overrightarrow{e'}_1, \overrightarrow{e'}_2) = (\overrightarrow{e}_1 + \overrightarrow{e}_2, \overrightarrow{e}_1 - \overrightarrow{e}_2)$:
Abbiamo $B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ per cui $v' = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix}$

- Altre proprietà delle applicazioni lineari. Sia $\varphi : V \to W$ un'applicazione lineare.

1) $\ker \varphi$ è un sottospazio di V. Infatti, se $\overrightarrow{v}_1, \overrightarrow{v}_2 \in \ker \varphi \Rightarrow \varphi(\overrightarrow{v}_1) = \overrightarrow{0}$ e $\varphi(\overrightarrow{v}_2) = \overrightarrow{0} \Rightarrow \varphi(\overrightarrow{v}_1 + \overrightarrow{v}_2) = \varphi(\overrightarrow{v}_1) + \varphi(\overrightarrow{v}_2) = \overrightarrow{0} \Rightarrow \overrightarrow{v}_1 + \overrightarrow{v}_2 \in \ker \varphi$.

Inoltre, se $\overrightarrow{v}_1 \in \ker \varphi$, $\varphi(\lambda \overrightarrow{v}_1) = \lambda \varphi(\overrightarrow{v}_1) = \lambda \cdot \overrightarrow{0} = \overrightarrow{0} \Rightarrow \lambda \overrightarrow{v}_1 \in \ker \varphi$.

2) Osserviamo, in particolare, che $\ker \varphi = \left\{ \overrightarrow{0} \right\}$ **se e solo se φ è iniettiva**. Infatti, se φ è iniettiva, allora l'unico elemento che viene mandato in $\overrightarrow{0} \in W$ può essere solo $\overrightarrow{0} \in V$. Viceversa, se $\ker \varphi = \left\{ \overrightarrow{0} \right\}$, allora $\varphi(\overrightarrow{v_1}) = \varphi(\overrightarrow{v_2})$ implica che $\varphi(\overrightarrow{v_1} - \overrightarrow{v_2}) = \overrightarrow{0}$, ovvero $\overrightarrow{v_1} - \overrightarrow{v_2} = \overrightarrow{0} \in V$, cioè $\overrightarrow{v_1} = \overrightarrow{v_2}$.

3) $\mathrm{Im}\, \varphi$ è un sottospazio di W. Infatti, se $\overrightarrow{v}_1, \overrightarrow{v}_2 \in \mathrm{Im}\, \varphi$, allora $\exists \overrightarrow{w}_1, \overrightarrow{w}_2 \in V$ tali che $\overrightarrow{v}_1 = \varphi(\overrightarrow{w}_1)$ e $\overrightarrow{v}_2 = \varphi(\overrightarrow{w}_2)$. $\overrightarrow{v}_1 + \overrightarrow{v}_2 = \varphi(\overrightarrow{w}_1) + \varphi(\overrightarrow{w}_2) = \varphi(\overrightarrow{w}_1 + \overrightarrow{w}_2)$, cioè $\exists \overrightarrow{w} = \overrightarrow{w}_1 + \overrightarrow{w}_2 \in V$ tale che $\varphi(\overrightarrow{w}) = \overrightarrow{v}_1 + \overrightarrow{v}_2$, cioè $\overrightarrow{v}_1 + \overrightarrow{v}_2 \in \mathrm{Im}\, \varphi$. Inoltre, se $\overrightarrow{v}_1 \in \mathrm{Im}\, \varphi$, $\lambda \overrightarrow{v}_1 = \lambda \varphi(\overrightarrow{w}_1) = \varphi(\lambda \overrightarrow{w}_1) \Rightarrow \lambda \overrightarrow{v}_1 \in \mathrm{Im}\, \varphi$.

4) Se $\{\overrightarrow{v}_1; \overrightarrow{v}_2; ...; \overrightarrow{v_n}\}$ è una base di V, allora $\{\varphi(\overrightarrow{v}_1); \varphi(\overrightarrow{v}_2); ...; \varphi(\overrightarrow{v_n})\}$ è un sistema di generatori per $\mathrm{Im}\, \varphi$. $\{\varphi(\overrightarrow{v}_1); \varphi(\overrightarrow{v}_2); ...; \varphi(\overrightarrow{v_n})\}$ è una base per $\mathrm{Im}\, \varphi$ solo nel caso in cui φ sia iniettiva.

Osserviamo, in particolare, che sussiste il fondamentale:

- **Teorema delle dimensioni**. Se $\varphi : V \to W$ è un'applicazione lineare,

$$\boxed{\dim\ker \varphi + \dim\mathrm{Im}\, \varphi = n = \dim V}.$$

Si vede dalla formula che **se $n < m$ l'applicazione non può essere suriettiva e che se $n > m$, l'applicazione non può essere iniettiva**.

- DIMOSTRAZIONE DEL TEOREMA DELLA DIMENSIONE.

Supponiamo che $F : V \longmapsto W$ sia un'applicazione lineare, che $\dim(\ker(F)) = r$ e che $\dim(\mathrm{Im}(F)) = s$. Sia $\{\overrightarrow{u_1}; \overrightarrow{u_2}; ...; \overrightarrow{u_s}\}$ una base di $\mathrm{Im}(F)$, e sia $\{\overrightarrow{w_1}; \overrightarrow{w_2}; ...; \overrightarrow{w_r}\}$ una base di $\ker(F)$. $\{\overrightarrow{u_1}; \overrightarrow{u_2}; ...; \overrightarrow{u_s}\}$ stanno in $\mathrm{Im}(F)$, dunque esistono $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_s}\}$ in V tali che $F(\overrightarrow{v_1}) = \overrightarrow{u_1}$; $F(\overrightarrow{v_2}) = \overrightarrow{u_2}$;...; $F(\overrightarrow{v_s}) = \overrightarrow{u_s}$. Ora affermo che:

$\mathcal{B} = \{\overrightarrow{w_1}; \overrightarrow{w_2}; ...; \overrightarrow{w_r}; \overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_s}\}$ è una base di V.

1) $\mathcal{B}$ genera V.

Sia $\overrightarrow{v} \in V$. Allora, $F(\overrightarrow{v}) \in \mathrm{Im}(F)$. $\{\overrightarrow{u_1}; \overrightarrow{u_2}; ...; \overrightarrow{u_s}\}$ generano $\mathrm{Im}(F)$, dunque esistono dei numeri reali $a_1; a_2; ...; a_s$ tali che $F(\overrightarrow{v}) = a_1\overrightarrow{u_1} + a_2\overrightarrow{u_2} + ... + a_s\overrightarrow{u_s}$. Sia $\overrightarrow{\widehat{v}} = a_1\overrightarrow{v_1} + a_2\overrightarrow{v_2} + ... + a_s\overrightarrow{v_s} - \overrightarrow{v}$. Allora, per la linearità di F ed il fatto che $F(\overrightarrow{v_j}) = \overrightarrow{u_j}$, si vede che $F\left(\overrightarrow{\widehat{v}}\right) = 0$. Ma allora, visto che $\overrightarrow{\widehat{v}} \in \ker(F)$, si ha $\overrightarrow{\widehat{v}} = b_1\overrightarrow{w_1} + b_2\overrightarrow{w_2} + ... + b_r\overrightarrow{w_r}$, per una scelta opportuna di numeri reali o complessi $b_1; b_2; ...; b_r$. Alla fine,

$$\overrightarrow{v} = a_1\overrightarrow{v_1} + a_2\overrightarrow{v_2} + ... + a_s\overrightarrow{v_s} - (b_1\overrightarrow{w_1} + b_2\overrightarrow{w_2} + ... + b_r\overrightarrow{w_r}).$$

2) $\{\overrightarrow{w_1}; \overrightarrow{w_2}; ...; \overrightarrow{w_r}; \overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_s}\}$ sono linearmente indipendenti.
Supponiamo che
$\alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + ... + \alpha_r\overrightarrow{w_r} + \beta_1\overrightarrow{v_1} + \beta_2\overrightarrow{v_2} + ... + \beta_s\overrightarrow{v_s} = 0$, per un'opportuna scelta di $\{\alpha_1; \alpha_2; ...; \alpha_r; \beta_1; \beta_2; ...; \beta_s\}$. Allora,

$$\begin{aligned}\overrightarrow{0} = F\left(\overrightarrow{0}\right) &= F\left(\alpha_1\overrightarrow{w_1} + \alpha_2\overrightarrow{w_2} + ... + \alpha_r\overrightarrow{w_r} + \beta_1\overrightarrow{v_1} + \beta_2\overrightarrow{v_2} + ... + \beta_s\overrightarrow{v_s}\right) = \\ &= \beta_1 F\left(\overrightarrow{v_1}\right) + \beta_2 F\left(\overrightarrow{v_2}\right) + ... + \beta_s F\left(\overrightarrow{v_s}\right) = \\ &= \beta_1\overrightarrow{u_1} + \beta_2\overrightarrow{u_2} + ... + \beta_s\overrightarrow{u_s}.\end{aligned}$$

Ma $\{\overrightarrow{u_1}; \overrightarrow{u_2}; ...; \overrightarrow{u_s}\}$ sono linearmente indipendenti, dunque $\beta_1 = \beta_2 = ... = \beta_s = 0$. Poichè anche $\{\overrightarrow{w_1}; \overrightarrow{w_2}; ...; \overrightarrow{w_r}\}$ sono linearmente indipendenti, si ottiene, di conseguenza, anche che $\alpha_1 = \alpha_2 = ... = \alpha_r = 0$.

Poichè $\mathcal{B}$ è una base di V, deve essere $n = r + s = \dim(\ker(F)) + \dim(\mathrm{Im}(F))$.
OSSERVAZIONE: Se $\varphi : V \to V$, se A è una matrice associata a φ in una qualsiasi base di V:

$$\boxed{\dim\mathrm{Im}\,\varphi = r(A)}$$

La formula ha senso, perchè abbiamo visto che non dipende dalla base scelta per scrivere A. Inoltre, siccome $r(A)$ è il numero di colonne indipendenti e le colonne di A sono i trasformati dei vettori della base, allora $\mathrm{Im}\,\varphi$ è generata da questi vettori.
COROLLARIO IMPORTANTE: Se consideriamo **endomorfismi**, ovvero applicazioni $F : V \longmapsto V$ si ha l'isomorfismo:

$$V = \ker F \oplus \mathrm{Im}\, F$$

- DEFINIZIONE

Si definisce $Hom(\mathbb{R}^n; \mathbb{R}^m)$ come lo spazio vettoriale di tutte le applicazioni lineari $\varphi : \mathbb{R}^n \longmapsto \mathbb{R}^m$. Esso è dotato delle seguenti operazioni:
(**addizione**) $(\varphi_1 + \varphi_2)(\overrightarrow{v}) = \varphi_1(\overrightarrow{v}) + \varphi_2(\overrightarrow{v})$,
$\forall\varphi_1; \varphi_2 \in Hom(\mathbb{R}^n; \mathbb{R}^m)$; $\forall\overrightarrow{v} \in \mathbb{R}^n$.
(**moltiplicazione per scalari**) $(\lambda\varphi_1)(\overrightarrow{v}) = \lambda \cdot \varphi_1(\overrightarrow{v})$,
$\forall\lambda \in \mathbb{R}$, $\forall\varphi_1 \in Hom(\mathbb{R}^n; \mathbb{R}^m)$, $\forall\overrightarrow{v} \in \mathbb{R}^n$.
$Hom(\mathbb{R}^n; \mathbb{R}^m)$ può anche essere identificato con lo spazio delle matrici reali $m \times n$. Infatti, fissate le basi standard su $\mathbb{R}^n$ e su $\mathbb{R}^m$, si ha che ad ogni applicazione lineare $\varphi : \mathbb{R}^n \longmapsto \mathbb{R}^m$ corrisponde un'unica matrice A_φ reale $m \times n$ e, viceversa, ad ogni matrice reale A di dimensione $m \times n$ corrisponde l'applicazione lineare $\varphi_A(\overrightarrow{v}) = eAv$, $\varphi : \mathbb{R}^n \longmapsto \mathbb{R}^m$. Dunque, $Hom(\mathbb{R}^n; \mathbb{R}^m)$ è uno spazio vettoriale di dimensione mn.
Analogamente si definisce $Hom(V; W)$, ricordando che ogni spazio vettoriale reale (complesso) di dimensione n è isomorfo a $\mathbb{R}^n$ ($\mathbb{C}^n$).

ESEMPIO.
In $\mathbb{R}^3$ con la base canonica $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, consideriamo l'applicazione φ:

$$\begin{aligned} \varphi(\overrightarrow{e}_1) &= \overrightarrow{e}_1 + \overrightarrow{e}_2 \\ \varphi(\overrightarrow{e}_2) &= \overrightarrow{e}_1 + \overrightarrow{e}_3 \\ \varphi(\overrightarrow{e}_3) &= \overrightarrow{e}_2 - \overrightarrow{e}_3. \end{aligned}$$

La matrice associata è $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Troviamo

$$\ker\varphi = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x + y = 0 \\ x + z = 0 \\ y - z = 0 \end{cases}$$
$$\Rightarrow \begin{cases} y = z \\ z = -x \end{cases}$$

cioè $\overrightarrow{v} \in \ker\varphi \Leftrightarrow$ è del tipo $\begin{pmatrix} x \\ -x \\ -x \end{pmatrix} = x \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

Una base di $\ker\varphi$ è $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \Rightarrow \dim\ker\varphi = 1$.

Troviamo

$$\operatorname{Im}\varphi = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : A \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\}.$$

Intanto, si vede che $\det A = 0$, quindi $r(A) = 2$. Cerchiamo ora una base per $\operatorname{Im}\varphi$. Si ha

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Leftrightarrow \begin{cases} x' + y' = x \\ x' + z' = y \\ y' - z' = z \end{cases}$$

$$\Rightarrow x - y = z,$$

cioè $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \operatorname{Im}\varphi$ se e solo se è del tipo

$$\begin{pmatrix} x \\ y \\ x-y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Una base per $\operatorname{Im}\varphi$ è allora $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ e

$$\operatorname{dimIm}\varphi = 2 = r(A).$$

Si poteva osservare che, dalla definizione di φ,

$$\varphi(\overrightarrow{e}_3) = \varphi(\overrightarrow{e}_1) - \varphi(\overrightarrow{e}_2),$$

quindi i vettori che stanno in $\operatorname{Im}\varphi$ sono del tipo

$$\varphi(\overrightarrow{w}) = a\varphi(\overrightarrow{e}_1) + b\varphi(\overrightarrow{e}_2).$$

Quindi

$$\varphi(\overrightarrow{e}_1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$
$$\varphi(\overrightarrow{e}_2) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

sono una base di $\operatorname{Im}\varphi$.

6.1 Lo spazio duale.

Sia V uno spazio vettoriale su $\mathbb{R}$ (o su $\mathbb{C}$). Indichiamo con V^* l'insieme di tutte le applicazioni lineari di V in $\mathbb{R}$ (o $\mathbb{C}$).

- TEOREMA

V^* è uno spazio vettoriale su $\mathbb{R}$ (o $\mathbb{C}$).

Dim.) Si verifica subito che V^* è uno spazio vettoriale su $\mathbb{R}$ (o $\mathbb{C}$), con le seguenti operazioni:

$$(\varphi_1 + \varphi_2)(\overrightarrow{v}) = (\varphi_1)(\overrightarrow{v}) + \varphi_2(\overrightarrow{v}),$$
$$\forall \overrightarrow{v} \in V,\ \forall \varphi_1; \varphi_2 \in V^*,$$

e

$$(c \cdot \varphi)(\overrightarrow{v}) = c(\varphi(\overrightarrow{v})),$$
$$\forall \overrightarrow{v} \in V,\ \forall \varphi \in V^*,\ \forall c \in \mathbb{R} \text{ (o } \mathbb{C}).$$

Gli elementi di V^* si dicono FUNZIONALI su V, e V^* si dice lo SPAZIO DUALE di V.

- TEOREMA

Sia V uno spazio vettoriale di dimensione finita su $\mathbb{R}$ (o $\mathbb{C}$). Allora, lo spazio duale V^* ha dimensione finita, e si ha $dim(V) = dim(V^*)$.

Dim.) Sia $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_n}\}$ una base di V. Voglio trovare una base di V^*. Sia v_i^* il funzionale lineare tale che

$$v_i^*(\overrightarrow{v_j}) \begin{cases} 1 \text{ per } i = j \\ \\ 0 \text{ per } i \neq j \end{cases}.$$

Si vuole dimostrare che $\{v_1^*; v_2^*; ...; v_n^*\}$ è una base di V^*.

1) $\{v_1^*; v_2^*; ...; v_n^*\}$ **generano** V. Sia F un generico funzionale su V. Allora, F è completamente determinato conoscendo i suoi valori sulla base $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_n}\}$. Siano $a_1 = F(\overrightarrow{v_1})$; $a_2 = F(\overrightarrow{v_2})$;...; $a_n = F(\overrightarrow{v_n})$. Allora, $F = a_1 v_1^* + a_2 v_2^* + ... + a_n v_n^*$.

2) $\{v_1^*; v_2^*; ...; v_n^*\}$ sono **linearmente indipendenti.** Siano $\{c_1; c_2; ...; c_n\}$ in $\mathbb{R}$ (o $\mathbb{C}$) tali che

$$c_1 v_1^* + c v_2^* + ... + c_n v_n^* = 0.$$

Allora, $(c_1 v_1^* + c v_2^* + ... + c_n v_n^*)(\overrightarrow{v_1}) = 0$ implica $(c_1 v_1^*)(\overrightarrow{v_1}) = 0$; cioè $c_1 = 0$. Procedendo allo stesso modo per ogni altro indice $i \in \{1; 2; ...; n\}$, si ottiene la tesi.

La base $\{v_1^*; v_2^*; ...; v_n^*\}$ si dice **base duale** di V.

Ora, lavoriamo sullo spazio vettoriale reale V.

• TEOREMA DI RIESZ.

Sia V uno spazio vettoriale reale di dimensione finita munito di un prodotto interno (bilineare, simmetrico e definito positivo). Se $T : V \longrightarrow \mathbb{R}$ è un'applicazione lineare, allora esiste un unico vettore $\overrightarrow{u} \in V$ tale che $T(\overrightarrow{v}) = \langle \overrightarrow{u}; \overrightarrow{v} \rangle$, $\forall \overrightarrow{v} \in V$.

Dim.) Sia $\left\{ \overrightarrow{v_1'}; \overrightarrow{v_2'}; ...; \overrightarrow{v_n'} \right\}$ una base di V. Utilizzando il procedimento di ortogonalizzazione di Gram-Schmidt, ci possiamo ricondurre ad una base ortonormale di V, $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_n}\}$. Allora, si pone

$$\overrightarrow{u} = a_1 \overrightarrow{v_1} + a_2 \overrightarrow{v_2} + ... + a_n \overrightarrow{v_n},$$

dove

$$a_1 = T(\overrightarrow{v_1});\ a_2 = T(\overrightarrow{v_2});\ ...;\ a_n = T(\overrightarrow{v_n}).$$

$\forall\ \overrightarrow{v} \in V$, sia

$$\overrightarrow{v} = b_1 \overrightarrow{v_1} + b_2 \overrightarrow{v_2} + ... + b_n \overrightarrow{v_n}.$$

Allora,

$$\begin{gathered} \langle \overrightarrow{u}; \overrightarrow{v} \rangle = a_1 b_1 + a_2 b_2 + ... + a_n b_n = \\ = T(\overrightarrow{v_1})\, b_1 + T(\overrightarrow{v_2})\, b_2 + ... + T(\overrightarrow{v_n})\, b_n = \\ = T(\overrightarrow{v}). \end{gathered}$$

Il vettore $\overrightarrow{u}$ è unico. Infatti, se esistesse un altro vettore $\overrightarrow{u'}$ che soddisfa la stessa proprietà, allora si avrebbe

$$\langle \overrightarrow{u}; \overrightarrow{v} \rangle = \left\langle \overrightarrow{u'}; \overrightarrow{v} \right\rangle, \forall \overrightarrow{v} \in V,$$

ovvero

$$\left\langle \overrightarrow{u} - \overrightarrow{u'}; \overrightarrow{v} \right\rangle = \overrightarrow{0}, \forall \overrightarrow{v} \in V.$$

Poichè il prodotto scalare è non degenere, deve essere $\overrightarrow{u} - \overrightarrow{u'} = \overrightarrow{0}$, cioè $\overrightarrow{u} = \overrightarrow{u'}$.

• TEOREMA

Sia V uno spazio vettoriale reale di dimensione finita munito di un prodotto scalare. Se $T : V \longrightarrow \mathbb{R}$ è un'applicazione lineare, e $\overrightarrow{u_T} \in V$ è l'unico vettore tale che $T(\overrightarrow{v}) = \langle \overrightarrow{u_T}; \overrightarrow{v} \rangle$, $\forall \overrightarrow{v} \in V$, allora l'applicazione $* : V^* \longmapsto V$ che ad ogni $T \in V^*$ associa $\overrightarrow{u_T}$ è un isomorfismo.

Dim.) $*$ è lineare:

$$(T_1 + T_2)(\overrightarrow{v}) = \langle \overrightarrow{u_{T_1+T_2}}; \overrightarrow{v} \rangle = \langle \overrightarrow{u_{T_1}}; \overrightarrow{v} \rangle + \langle \overrightarrow{u_{T_2}}; \overrightarrow{v} \rangle;$$
$$(cT)(\overrightarrow{v}) = \langle \overrightarrow{u_{cT}}; \overrightarrow{v} \rangle = c \langle \overrightarrow{u_T}; \overrightarrow{v} \rangle.$$

$Ker\,(*) = \left\{ \overrightarrow{0} \right\}$, perchè $\overrightarrow{u_T} = \overrightarrow{0}$ se e solo se $T = 0$. Quindi, $*$ è iniettiva. $*$ è anche suriettiva, perchè V e V^* hanno la stessa dimensione.

Il **doppio duale** è $V^{**} = (V^*)^*$, ovvero il duale di V^*. V^{**} è facilmente identificabile con V. Infatti, se $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_n}\}$ è una base ortonormale di V, $\{v_1^*; v_2^*; ...; v_n^*\}$ è la sua base duale e $\{\widehat{v_1}; \widehat{v_2}; ...; \widehat{v_n}\}$ è la base duale di $\{v_1^*; v_2^*; ...; v_n^*\}$, si ha che l'applicazione

$$\Psi : V \longmapsto V^{**},$$
$$\Psi(\overrightarrow{v_i}) = \widehat{v_i}$$

è un isomorfismo (essendo la composizione di due isomorfismi). Inoltre, vale la seguente proprietà: se $\widehat{v} = \Psi(\overrightarrow{v})$, allora

$$\widehat{v}(F) = F(\overrightarrow{v_i}), \forall F \in V^*.$$

Ψ si dice **isomorfismo naturale** tra V e V^{**}.

7 Sistemi lineari.

Un **sistema lineare** di m equazioni ed n incognite è un insieme di equazioni della forma:

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

dove $x_1, ..., x_n \in \mathbb{R}$ (o $\mathbb{C}$) sono le **incognite**, $b_1, ..., b_m \in \mathbb{R}$ (o $\mathbb{C}$) sono i **termini noti** e $a_{ij} \in \mathbb{R}$ (o $\mathbb{C}$) sono i **coefficienti**.

Si può anche scrivere:

$$Ax = b,$$

dove A è la matrice

$$A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix}$$

e x e b sono i vettori:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad e \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

- Ci chiediamo tre cose in successione:
- a) Il sistema ammette soluzioni? (si dice anche: il sistema è **compatibile**?)
- b) Quante soluzioni ci sono?
- c) Quali sono?

Per rispondere ad a), è utile considerare la matrice A', detta *matrice completa*, ottenuta aggiungendo ad A la colonna dei termini noti:

$$A' = \begin{pmatrix} a_{11} & ... & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & ... & a_{mn} & b_m \end{pmatrix}.$$

Allora, se il sistema ammette x come soluzione, vuol dire che la colonna

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

è combinazione lineare (con coefficienti x_i) delle colonne della matrice A. Infatti, il sistema si può scrivere nel modo seguente:

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

Quindi necessariamente A e A' devono avere lo stesso rango r (ricordare che il rango è il numero di colonne linearmente indipendenti). Viceversa, se A e A' hanno lo stesso rango, vuol dire che la colonna dei b_i è combinazione lineare delle altre e quindi il sistema ammette soluzione. In altre parole, $b \in \operatorname{Im} A$ (=immagine di A), quindi:

- **Un sistema è risolubile se e solo se rango**(A) = **rango**$(A') = r$.

Per rispondere a b), consideriamo prima il caso omogeneo, cioè il caso in cui b è il vettore nullo:

$$b = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

In questo caso, ovviamente, rango(A) = rango$(A') = r$, per cui:

Un sistema omogeneo è sempre risolubile ($x = 0$ è sempre soluzione).

Indichiamo ancora con $\ker A$ l'insieme degli x di $\mathbb{R}^n$ tali che $Ax = 0$ e con $\operatorname{Im} A$ l'insieme degli y di $\mathbb{R}^m$ tali che esiste un x di $\mathbb{R}^n$ per cui $Ax = y$.

Il sistema omogeneo si scrive $Ax = 0$, cioè $x \in \ker A$, e il *teorema delle dimensioni* ci dice che

$$\dim \ker A + \dim \operatorname{Im} A = n.$$

Ma allora: $\dim \ker A =$ (dimensione dello spazio delle soluzioni del sistema omogeneo) $= n - r$ (perchè rango$(A) = \dim \operatorname{Im} A$).

- Si usa dire che **il sistema ha ∞^{n-r} soluzioni**.

Nel caso non omogeneo, basta osservare che, se x è una soluzione qualsiasi e x' è una soluzione del sistema omogeneo, allora

$$x' = x + x'$$

è ancora soluzione. Infatti:$Ax' = Ax + Ax' = Ax = b$, perchè $Ax' = 0$. Viceversa, se x e x' sono due soluzioni distinte del sistema non omogeneo, allora $Ax = Ax' = b$, cioè $A(x - x') = 0$, cioè $x - x' \in \ker A = \{$soluzioni del sistema omogeneo$\}$.

- Quindi, **se un sistema di m equazioni ed n incognite è compatibile e di rango r, allora ammette ∞^{n-r} soluzioni (**che si ottengono sommando a una soluzione particolare del sistema dato **tutte** le soluzioni del corrispondente sistema omogeneo**).**

Per rispondere a c), se il sistema è compatibile e di rango r, si possono eliminare $m - r$ e equazioni e considerare come termini noti $n - r$ delle incognite, per arrivare ad un sistema quadrato di rango massimo, cioè ad un **sistema normale**, cioè un sistema di r equazioni in r incognite, del tipo

$$\boxed{Cy = d}$$

dove C è una sottomatrice quadrata $r \times r$ di A, con $\det C \neq 0$, y è una colonna contenente r delle incognite x_i e d è una colonna contenente r dei termini noti b_i sommati a combinazioni lineari delle altre $n - r$ incognite, la cui unica soluzione è $y = C^{-1}d$ (unica, perchè $r - r = 0$). Tutte le soluzioni del sistema originario si ottengono quindi trattando le $n - r$ incognite che compaiono in d come parametri variabili.

ESEMPI.

1) Consideriamo il sistema di 3 equazioni e 3 incognite:

$$\begin{cases} x + y = 1 \\ y + z = 0 \\ x - z = 1 \end{cases}$$

Abbiamo:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad A' = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix}.$$

Inoltre, rango(A) $= 2 =$ rango(A') (perchè la 4^a colonna è uguale alla 1^a). Il sistema è compatibile con $\infty^{3-2} = \infty$ soluzioni.

Passiamo al sistema normale: guardando la matrice A, si vede che il minore

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

ha det $\neq 0$, per cui il sistema

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

è normale di rango massimo:

$$C \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -z \end{pmatrix},$$

dove

$$C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad d = \begin{pmatrix} 1 \\ -z \end{pmatrix}.$$

Allora

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

La soluzione del sistema è quindi

$$(1 + z, -z, z).$$

2) Sistema di 3 equazioni in 4 incognite:

$$\begin{cases} z+2t=3 \\ 2x+4y-2z=4 \\ 2x+4y-z+2t=7 \end{cases}$$

Si ha

$$A=\begin{pmatrix} 0 & 0 & 1 & 2 \\ 2 & 4 & -2 & 0 \\ 2 & 4 & -1 & 2 \end{pmatrix},\ b=\begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix},\ A' = \begin{pmatrix} 0 & 0 & 1 & 2 & 3 \\ 2 & 4 & -2 & 0 & 4 \\ 2 & 4 & -1 & 2 & 7 \end{pmatrix}.$$

Si verifica che rango$(A)=2$ e poichè b è somma delle colonne di A, anche rango$(A')=2$. Il sistema è quindi compatibile ed essendo $n=4, m=3, r=2$, si hanno $\infty^{4-2}=\infty^2$ soluzioni. Eliminiamo una equazione e consideriamo 2 incognite come termini noti. Osserviamo che $\det\begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix}=-2\neq 0$, per cui il sistema:

$$\begin{cases} z=3-2t \\ 2x-2z=4-4y \end{cases}$$

è normale di rango massimo.

$$C\begin{pmatrix} x \\ z \end{pmatrix}=d$$

con

$$C=\begin{pmatrix} 0 & 1 \\ 2 & -2 \end{pmatrix},\ d=\begin{pmatrix} 3-2t \\ 4-4y \end{pmatrix},\ C^{-1}=-\frac{1}{2}\begin{pmatrix} -2 & -1 \\ -2 & 0 \end{pmatrix}.$$

Allora

$$\begin{pmatrix} x \\ z \end{pmatrix}=C^{-1}d=\begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3-2t \\ 4-4y \end{pmatrix}=\begin{pmatrix} 3-2y+2-2y \\ 3-2t \end{pmatrix}.$$

La soluzione generale del sistema è $(5-2y-2t, y, 3-2t, t)$.

8 Autovalori e Autovettori.

Sia A una matrice **quadrata** di ordine n reale (o complessa) e $\vec{v}$ un vettore colonna.

Se $A\vec{v} = \lambda\vec{v}$ con $\lambda \in \mathbb{R}$ (o $\mathbb{C}$) e $\vec{v} \neq 0$, allora λ è detto **autovalore** di A e $\vec{v}$ è l'**autovettore** corrispondente a λ. Si indica con V_λ l'**autospazio** di λ, cioè

$$V_\lambda = \{\vec{v} : A\vec{v} = \lambda\vec{v}\} \cup \left\{\vec{0}\right\}$$

V_λ **include tutti gli autovettori ed anche il vettore nullo ed è un sottospazio di** $\mathbb{R}^n$ (o di $\mathbb{C}^n$).

Analogamente, se $\varphi : \mathbb{R}^n \longmapsto \mathbb{R}^n$ (oppure $\varphi : \mathbb{C}^n \longmapsto \mathbb{C}^n$) è un'applicazione lineare, e se $\varphi\left(\overrightarrow{v}\right) = \lambda\overrightarrow{v}$ con $\lambda \in \mathbb{R}$ (o $\mathbb{C}$) e $\overrightarrow{v} \neq 0$, allora λ è detto **autovalore** di φ e $\overrightarrow{v}$ è l'**autovettore** corrispondente a λ.

- OSSERVAZIONI **(importanti!)**.

1) Se λ è un autovalore di A con autovettore $\vec{v}$, allora λ^m è un autovalore di A^m **con gli stessi autovettori**, per ogni m intero positivo. Infatti, per $m = 2$,

$$A^2\vec{v} = AA\vec{v} = A\lambda\vec{v} = \lambda A\vec{v} = \lambda\lambda\vec{v} = \lambda^2\vec{v}.$$

Allo stesso modo si procede per m generico. Se A è invertibile, lo stesso vale per ogni m intero negativo.

2) Autospazi di autovalori distinti hanno in comune solo il vettore nullo, cioè **autovettori di autovalori distinti sono indipendenti**. Infatti, se

$$\vec{v} \in V_\lambda \text{ e } \vec{v} \in V_\mu \Rightarrow A\vec{v} = \lambda\vec{v} \text{ e } A\vec{v} = \mu\vec{v} \Rightarrow (\lambda - \mu)\,\vec{v} = 0$$

ma $\lambda \neq \mu \Rightarrow \vec{v} = 0$.

- <u>Come si trovano autovalori e autovettori</u>**.**

$A\vec{v} = \lambda\vec{v}$ con $\vec{v} \neq \vec{0}$ si può riscrivere come $(A - \lambda I)\,\vec{v} = 0$, che è un sistema lineare omogeneo con soluzioni non banali $\vec{v} \neq \vec{0}$ se e solo se

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

Quest'ultima è l'**equazione degli autovalori**. (Può accadere che una o più soluzioni siano complesse; in questo caso **non** sono autovalori se si considerano matrici reali che operano su $\mathbb{R}^n$. Invece, tutte le matrici complesse che operano su $\mathbb{C}^n$ hanno autovalori).

Indichiamo con $P(\lambda)$ la seguente espressione:

$$P(\lambda) = \det(A - \lambda I).$$

$P(\lambda)$ è un polinomio in λ, detto **polinomio caratteristico** di A, ed ha grado n. Si può verificare subito che il coefficiente di λ^n è $(-1)^n$, mentre il termine noto è $\det A$ (infatti, il termine noto di un polinomio è $P(0)$ e, nel caso del polinomio caratteristico, $P(0) = \det(A - 0I) = \det A$).

8.1 Diagonalizzazione delle matrici.

Una matrice si dice **diagonalizzabile** se esiste una base fatta da autovettori. In **questa** base A è chiaramente scritta in forma diagonale **con gli autovalori sulla diagonale**.

C'è un importante teorema che dice che le matrici A **reali e simmetriche**, cioè tali che $A = A^t$, **sono diagonalizzabili** (ovviamente non è obbligatorio che una matrice sia reale e simmetrica per essere diagonalizzabile). Nel caso complesso, la condizione di simmetria è $A = \overline{A}^t$, dove $\overline{A}^t$ è formata dai complessi coniugati degli elementi di A^t.

Per una matrice reale 2×2 il risultato è ovvio:

$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è simmetrica se $b = c$. Troviamo gli autovalori: $\det(A - \lambda I) = 0$ diventa

$$\det \begin{pmatrix} a-\lambda & b \\ b & d-\lambda \end{pmatrix} = 0 \Leftrightarrow (a-\lambda)(d-\lambda) - b^2 = 0$$
$$\Leftrightarrow ad + \lambda^2 - (a+d)\lambda - b^2 = 0 \Leftrightarrow \lambda^2 - (a+d)\lambda - b^2 + ad = 0.$$

$$\begin{aligned} \Delta &= (a+d)^2 + 4b^2 - 4ad = a^2 + d^2 + 4b^2 - 4ad \\ &= (a-d)^2 + 4b^2 \geqslant 0. \end{aligned}$$

Se $\Delta = 0$, allora $b = 0$ e la matrice è gia diagonale. Se invece gli autovalori sono distinti fra loro ($\Delta > 0$), gli autovettori formano una base, perchè sono 2 e sono indipendenti perchè appartengono ad autospazi di autovalori distinti.

Se A è diagonalizzabile, vuol dire che cambiando base, cioè passando ad una base fatta di autovettori, diventa $A' = M^{-1}AM$, con A' diagonale. La formula non è unica, nel senso che **qualsiasi matrice M che ha come colonne le componenti di una base di autovettori, trasforma A in una matrice diagonale.**

- DEFINIZIONE

Due matrici (reali o complesse) A e A' di dimensione $n \times n$ si dicono **simili** se esiste una matrice invertibile S, anch'essa di dimensione $n \times n$, tale che $A' = S^{-1}AS$.

Due matrici simili hanno lo **stesso determinante** e la **stessa traccia.** Per matrici diagonalizzabili, cioè simili ad una matrice diagonale, il determinante è il **prodotto** di tutti gli autovalori (contati con le rispettive molteplicità), mentre la traccia è la **somma** degli stessi.

In particolare, si verifica (esercizio facile) che la similitudine è una relazione di equivalenza.

- OSSERVAZIONE

Si parla spesso anche di "diagonalizzazione di applicazioni lineari", invece che di "diagonalizzazione di matrici". In effetti, i due concetti sono equivalenti, dato che, fissate le basi per dominio e codominio, le applicazioni lineari $\varphi : \mathbb{R}^n \longmapsto \mathbb{R}^n$ sono in corrispondenza biunivoca con le matrici $n \times n$ reali (e lo stesso è vero anche nel caso complesso). "Diagonalizzare una matrice" A significa trovare una matrice diagonale simile ad A. "Diagonalizzare un'applicazione lineare" φ significa trovare una base di autovettori per l'applicazione lineare φ. Le due operazioni sono in corrispondenza, perchè il cambiamento di base per un'applicazione lineare corrisponde a trasformare la matrice associata in una nuova matrice ad essa simile.

Base standard	$\Longrightarrow$	Base di autovettori
Applicazione lineare φ	$\Longrightarrow$	Applicazione lineare φ
Matrice associata: A	$\Longrightarrow$	Matrice associata: $B^{-1}AB$ (B matrice cambiamento base)
A è una matrice diagonalizzabile	$\Longrightarrow$	$A' = B^{-1}AB$ è diagonale.

- CONDIZIONI DI DIAGONALIZZABILITÀ

Sia $d(\lambda)$ la **molteplicità algebrica** dell'autovalore λ (cioè la molteplicità di λ come radice del polinomio caratteristico) e $m(\lambda)$ la **molteplicità geometrica** di λ (cioè la dimensione dell'autospazio V_λ). Si ha sempre (vedi più sotto) $1 \leqslant m(\lambda) \leqslant d(\lambda)$. Per cui, nel **caso complesso** (cioè nel caso di matrici operanti su vettori di $\mathbb{C}^n$), siccome un polinomio di grado n ha sempre n radici, si ha sempre

$$\sum d(\lambda) = n,$$

dove $\sum$ è la somma su tutti gli autovalori. Allora, se per ogni λ si ha $m(\lambda) = d(\lambda)$, cioè gli autovalori sono tutti **regolari**, è chiaro che **la matrice è diagonalizzabile**, perchè ci sono abbastanza autovettori indipendenti per formare una base $\sum m(\lambda) = n$.

Nel caso **reale**, questo non basta, perchè non è detto che ci siano n radici reali e quindi non è garantito che $\sum d(\lambda) = n$, quindi nemmeno che $\sum m(\lambda) = n$. Allora, una matrice reale è diagonalizzabile se $d(\lambda) = m(\lambda)$ per ogni λ **e, inoltre,** $\sum d(\lambda) = n$.

- TEOREMA:

Sia λ un autovalore della matrice $A \in M(\mathbb{R}; n \times n)$. Allora, $m(\lambda) \leqslant d(\lambda)$.

Dim. Sia $r = m(\lambda)$. Sia poi $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_r}; \overrightarrow{w_1}; \overrightarrow{w_2}; ...; \overrightarrow{w_{n-r}}\}$ una base di $\mathbb{R}^n$ tale che $\{\overrightarrow{v_1}; \overrightarrow{v_2}; ...; \overrightarrow{v_r}\}$ siano autovettori di A con autovalore λ. Rispetto a questa base, la matrice A si trasforma in

$$\begin{pmatrix} \lambda & 0 & 0 & 0 & | & a_{1,1} & \cdots & a_{1,n-r} \\ 0 & \lambda & 0 & 0 & | & a_{2,1} & & a_{2,n-r} \\ 0 & 0 & \ddots & 0 & | & \vdots & & \vdots \\ 0 & 0 & 0 & \lambda & | & a_{r,1} & \cdots & a_{r,n-r} \\ - & - & - & - & + & - & - & - \\ 0 & 0 & \cdots & 0 & | & b_{1,1} & \cdots & b_{1,n-r} \\ \vdots & \vdots & & \vdots & | & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & | & b_{n-r,1} & \cdots & b_{n-r,n-r} \end{pmatrix}.$$

Il suo polinomio caratteristico, nella indeterminata t, è della forma $(t-\lambda)^r P(t)$, dunque la molteplicità algebrica dell'autovalore λ è almeno $r = m(\lambda)$.

- CASI PARTICOLARI.

Se una matrice **complessa** ha autovalori **tutti distinti** è diagonalizzabile. Infatti, in questo caso, $d(\lambda) = 1$ per ogni λ e $\sum d(\lambda) = n$ e quindi deve essere anche $m(\lambda) = 1$, quindi siamo nel caso di autovalori regolari. Una matrice **reale** che ha autovalori tutti distinti è diagonalizzabile se, **inoltre**, $\sum d(\lambda) = n$.

Una matrice della forma

$$A = \begin{pmatrix} M_1 & 0 & ... & 0 \\ 0 & M_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & M_s \end{pmatrix},$$

dove $M_1, M_2, ..., M_s$ sono matrici **quadrate**, è detta **matrice a blocchi**. Si può dimostrare che il suo determinante è uguale al prodotto dei determinanti dei vari blocchi $M_1, M_2, ..., M_s$. Inoltre, per lo studio di autovalori, autospazi e diagonalizzabilità si possono **considerare i vari blocchi separatamente.**

OSSERVAZIONE.

Se una matrice A è diagonalizzabile tutte le sue potenze intere A^m lo sono. Se A è invertibile e diagonalizzabile, anche le sue potenze intere negative lo sono (basta ricordare che A e A^m hanno gli stessi autovettori).

ESEMPI.

1) Trovare una matrice M che diagonalizza

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

Determiniamo gli autovalori:

$$\det(A - \lambda I) = \left(\frac{3}{2} - \lambda\right)\left(\frac{3}{2} - \lambda\right) - \frac{1}{4};$$
$$\det(A - \lambda I) = 0 \Leftrightarrow \lambda^2 - 3\lambda + \frac{9}{4} - \frac{1}{4} = 0 \Leftrightarrow \lambda^2 - 3\lambda + 2 = 0$$

Le soluzioni sono $\lambda_1 = 1$ e $\lambda_2 = 2$.

Per $\lambda_1 = 1$, si ha

$$\begin{pmatrix} \frac{3}{2} - 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} - 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} \frac{1}{2}x - \frac{1}{2}y = 0 \\ -\frac{1}{2}x + \frac{1}{2}y = 0 \end{cases}$$

$$\Rightarrow x = y.$$

Allora $\begin{pmatrix} x \\ x \end{pmatrix}$ è l'autospazio relativo a λ_1 e una base è, ad esempio, $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Per $\lambda_2 = 2$, si ha

$$\begin{pmatrix} \frac{3}{2}-2 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2}-2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -\frac{1}{2}x - \frac{1}{2}y = 0 \\ -\frac{1}{2}x - \frac{1}{2}y = 0 \end{cases}$$

$$\Rightarrow x = -y.$$

Allora l'autospazio V_2 è costituito da vettori del tipo $\begin{pmatrix} x \\ -x \end{pmatrix}$. Una base è, ad esempio, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Ora possiamo scrivere M, mettendo in colonna v_1 e v_2:

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad M^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

e si ha

$$A' = M^{-1}AM = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

2) In $\mathbb{R}^3$ consideriamo l'applicazione $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ tale che

$$\begin{aligned} \varphi(\overrightarrow{e}_1) &= \overrightarrow{e}_2; \\ \varphi(\overrightarrow{e}_2) &= \overrightarrow{e}_1; \\ \varphi(\overrightarrow{e}_3) &= \overrightarrow{e}_3. \end{aligned}$$

La matrice associata a φ è $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$$P(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)(\lambda^2 - 1).$$

$$P(\lambda) = 0 \Leftrightarrow \lambda = \begin{cases} \nearrow \ 1 \text{ con molteplicità } 2 \\ \searrow \ -1 \text{ con molteplicità } 1 \end{cases}$$

Autospazi: $V_1 = \{v : Av = v\}$; $V_2 = \{v : Av = -v\}$

$$Av = v \Leftrightarrow (A - I)v = 0 \Leftrightarrow \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Rightarrow r(A) = 1$, quindi eliminiamo due equazioni e fissiamo due parametri. V_1 è allora il piano di equazione $x - y = 0$, cioè $x = y$ e $v \in V_1$ è del tipo $v = \begin{pmatrix} x \\ x \\ z \end{pmatrix}$. Una base di V_1 è, ad esempio, $u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

$$Av = -v \Leftrightarrow (A + I)v = 0 \Leftrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$\Rightarrow r(A) = 2$, quindi eliminiamo un'equazione e fissiamo un parametro. V_2 è allora la retta di equazioni $\begin{cases} x + y = 0 \\ z = 0 \end{cases}$ e una base è data da $u_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

Una matrice diagonalizzante è, ad esempio,

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

e si ha $A' = M^{-1}AM = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

3) L'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ data da

$$f(\overrightarrow{e}_1) = -\overrightarrow{e}_2;$$
$$f(\overrightarrow{e}_2) = \overrightarrow{e}_1;$$
$$f(\overrightarrow{e}_3) = \overrightarrow{e}_3,$$

non è diagonalizzabile. Infatti, la matrice associata è

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

e il polinonio caratteristico è $P(\lambda) = (1 - \lambda)(\lambda^2 + 1)$. Un solo autovalore è reale $\lambda_1 = 1$. Allora V_1 ha dimensione 1 e un vettore non basta per formare una base in $\mathbb{R}^3$. A **è invece diagonalizzabile su** $\mathbb{C}$, cioè come endomorfismo di $\mathbb{C}^3$, perchè in $\mathbb{C}$ ha tre autovalori distinti: $\lambda_1 = 1$, $\lambda_2 = i$, $\lambda_3 = -i$ e i corrispondenti autospazi hanno tutti dimensione 1.

4) La matrice

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \end{pmatrix}$$

è a blocchi: $M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 6 \end{pmatrix}, M_2 = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$. Il primo blocco ha autovalori 1 e 6 (la verifica è immediata). Il secondo blocco ha rango 1, quindi la dimensione del nucleo è 2 ($\dim\ker M_2 = 3 - r(M_2)$) e 0 è un autovalore di molteplicità geometrica 2 (infatti, $\dim\ker M_2 = \dim\ker(M_2 - 0I) = d(0)$). Inoltre, riusciamo a trovare subito un altro autovalore, osservando che

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} = 6 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

per cui 6 è un autovalore. L'autovalore 1 ha molteplicità algebrica e geometrica uguale a 1. Per quanto riguarda l'autovalore 6, questo ha molteplicità algebrica uguale a 2, mentre si deve trovare la sua molteplicità geometrica:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ u \\ v \end{pmatrix} = \begin{pmatrix} x+y \\ 6y \\ 2z+2u+2v \\ 2z+2u+2v \\ 2z+2u+2v \end{pmatrix} = \begin{pmatrix} 6x \\ 6y \\ 6z \\ 6u \\ 6v \end{pmatrix}$$

$$\Rightarrow \begin{cases} x = \dfrac{t}{5} \\ y = t \\ z = r \\ u = r \\ v = r \end{cases} \Rightarrow \begin{pmatrix} x \\ y \\ z \\ u \\ v \end{pmatrix} = \frac{1}{5}t \begin{pmatrix} 1 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} + r \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

da cui una base per l'autospazio dell'autovalore 6 è:

$$\left\{ \begin{pmatrix} 1 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

quindi 6 è regolare. Infine, 0 ha molteplicità geometrica uguale a 2, come detto sopra. Anche la molteplicità algebrica è uguale a 2, perchè la matrice A è di ordine 5, perciò la somma delle molteplicità algebriche deve dare 5: $m(0) = 5 - m(6) - m(1) = 5 - 2 - 1 = 2$. Poichè tutti gli autovalori sono regolari, la matrice A è diagonalizzabile.

- TEOREMA DI CAYLEY-HAMILTON

Sia A una matrice. Abbiamo visto che $P(x) = \det(A - xI)$ è un polinomio di grado n nella variabile x e gli zeri di $P(x)$ sono gli autovalori di A. Consideriamo il seguente importante:

Teorema di Cayley-Hamilton. Ogni matrice è radice del suo polinomio caratteristico.

ESEMPIO.
Sia data la matrice

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

Polinomio caratteristico: $P(x) = \det(A - xI) = \det \begin{pmatrix} 1-x & 1 \\ 1 & 3-x \end{pmatrix}$ cioè $P(x) = x^2 - 3x - x + 3 - 1 = x^2 - 4x + 2$.

Il teorema dice che, se al posto di x mettiamo A e al posto di 1 mettiamo l'identità I, otteniamo la matrice nulla, cioè

$$P(A) = A^2 - 4A + 2I = O,$$

con O matrice con tutti 0. Infatti

$$\begin{aligned} A^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 4 & 10 \end{pmatrix} \\ -4A &= \begin{pmatrix} -4 & -4 \\ -4 & -12 \end{pmatrix} \\ 2I &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ A^2 - 4A + 2I &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Per le matrici 2×2 è sempre

$$P(x) = x^2 - xTrA + \det A,$$

dove TrA è la **traccia** di A, cioè la somma degli elementi sulla diagonale.

- Conseguenza interessante del teorema di Cayley-Hamilton è il calcolo della matrice inversa (se $\det A \neq 0$).

Sia $P(x) = (-1)^n x^n + ... + k_1 x + k_0$ il polinomio caratteristico.
(Si ricordi che il coefficiente di x^n è **sempre** $(-1)^n$ e il termine noto, cioè il termine di potenza 0, è **sempre** $k_0 = \det A$).

$$\begin{gathered} P(A) = O \Leftrightarrow (-1)^n A^n + ... + k_1 A + \det A \cdot I = O \\ \Leftrightarrow A\left[(-1)^n A^{n-1} + k_{n-1}A^{n-2} + ... + k_1 I\right] = -\det A \cdot I \\ \Leftrightarrow A\left[(-1)^n A^{n-1} + k_{n-1}A^{n-2} + ... + k_1 I\right] = -\det A \cdot AA^{-1} \end{gathered}$$

da cui

$$\boxed{A^{-1} = -\frac{1}{\det A}\left[(-1)^n A^{n-1} + k_{n-1}A^{n-2} + ... + k_1 I\right]}$$

ESEMPIO.
Data

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

si ha $P(x) = -x^3 + 6x^2 - 9x + 4$, quindi

$$A^{-1} = -\frac{1}{4}\left(-A^2 + 6A - 9I\right) = \frac{1}{4}\left(A^2 - 6A + 9I\right) = \frac{1}{4}\begin{pmatrix} 4 & -1 & -2 \\ 0 & 3 & -2 \\ 0 & -1 & -2 \end{pmatrix}$$