OTTAVIO CALIGARIS

ELISABETTA FERRANDO

PIETRO OLIVA

ELEMENTI DI ALGEBRA LINEARE E GEOMETRIA ANALITICA

APRIL 10, 2014

1. I numeri Complessi

1.1 L'introduzione dei numeri complessi

Nonostante i numeri complessi risolvano il problema di trovare soluzioni di un'equazione algebrica di secondo grado anche nel caso in cui il discriminante sia negativo, essi furono introdotti solo in seguito ai tentativi di risolvere un'equazione di terzo grado.

Già nel 1225, Leonardo da Pisa (Fibonacci) trovò la soluzione dell'equazione di terzo grado

$$x^3 + 2x^2 + 10x = 20$$

che gli era stata proposta da un matematico alla corte di Federico II in Sicilia inoltre verso la fine del '300 era noto che mediante il cambio di variabile $t = x + \frac{1}{3}a$ si può ridurre una generica equazione di terzo grado ad una più semplice.

Infatti da

$$\begin{cases} x^3 + ax^2 + bx = c \\ t = x + \frac{1}{3}a \end{cases}$$

si ottiene

$$t^3 - \left(\frac{1}{3}a^2 - b\right) t = c + \frac{ab}{3} - \frac{2}{27}a^3$$

quindi un'equazione, che si definisce in forma ridotta, del tipo

$$x^3 + px = q$$

Il primo a risolvere un'equazione in forma ridotta fu, nella prima metà del '500, Scipione del Ferro che confidò la formula risolutiva al suo allievo Antonio Maria Fiore sul letto di morte.

Tartaglia riscoprì indipendentemente la formula risolutiva e la comunicò senza dargli la dimostrazione a Cardano che fu in grado di ricostruirla per suo conto e di confrontarla con il il lavoro di del Ferro al quale aveva avuto accesso.

Cardano pubblicò la formula nella sua Ars Magna nella seconda metà del '500 riconoscendo in esso il contributo di Tartaglia e del Ferro.

Per risolvere un'equazione del tipo

$$x^3 + px = q$$

si cerca di determinare u, v tali che

$$x = u + v$$

risolva l'equazione.

Si ha

$$x^3 + px = (u+v)^3 + p(u+v) = u^3 + v^3 + (3uv + p)(u+v) = q$$

e pertanto possiamo trovare x se siamo in grado di trovare u, v in modo che

$$\begin{cases} 3uv = -p \\ u^3 + v^3 = q \end{cases} \qquad \text{ovvero} \qquad \begin{cases} u^3 v^3 = -\left(\frac{p}{3}\right)^3 \\ u^3 + v^3 = q \end{cases}$$

Avremo quindi che u^3, v^3 sono le soluzioni dell'equazione di secondo grado $z^2 - qz + \left(\frac{p}{3}\right)^3 = 0$ che possiamo risolvere.

Accade tuttavia che l'equazione di secondo grado non abbia soluzioni reali (in linguaggio moderno, quando il suo discriminante è negativo); Cardano non trattò questo caso definendolo *'casus irreducibilis'* e forse sottintendendo che non ci sono soluzioni, ma Bombelli, verso la fine del '500, osservò che l'equazione

$$x^3 = 15x + 4$$

ammette la soluzione $x = 4$ nonostante il procedimento di Cardano conduca ad un *'casus irreducibilis'*con valori di u, v dati da

$$u = \sqrt[3]{2 + \sqrt{-121}} \qquad , \qquad v = \sqrt[3]{2 - \sqrt{-121}}$$

Bombelli, in sostanza, definì $\imath = \sqrt{-1}$ e si pose il problema di determinare a, b in modo che

$$u = \sqrt[3]{2 + \sqrt{-121}} = \sqrt[3]{2 + \imath 11} = a + \imath b$$

cioè

$$2 + \imath 11 = (a + \imath b)^3$$

Svolgendo il cubo ed usando le regole di moltiplicazione che conseguono naturalmente dalla definizione di $\imath$

$$\imath^2 = -1 \qquad , \qquad \imath 1 = \imath$$

otteniamo

$$a(a^2 - 3b^2) + \imath(3a^2 - b^2)b = 2 + \imath 11$$

$$x^3 - 15x = 4$$
$$(u+v)^3 - 15(u+v) = 4$$
$$u^3 + v^3 + 3uv(u+v) - 15(u+v) = 4$$
$$\begin{cases} u^3 + v^3 = 4 \\ 3uv = 15 \end{cases} \qquad \begin{cases} u^3 + v^3 = 4 \\ u^3 v^3 = 125 \end{cases}$$
$$z^2 - 4z + 125 = 0$$
$$z = 2 \pm \sqrt{4 - 125} = 4 \pm \sqrt{-121}$$

per cui deve essere

$$\begin{cases} a(a^2 - 3b^2) = 2 \\ (3a^2 - b^2)b = 11 \end{cases}$$

e si vede che una soluzione è $a = 2$ e $b = 1$.

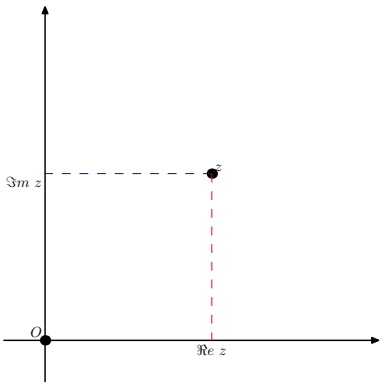
Quindi $u = 2 + \imath$, in maniera del tutto simile, $v = 2 - \imath$ ed infine $x = u + v = 4$.

Questo convinse Bombelli della bontà della sua definizione; egli non usò tuttavia la notazione $\imath$ che fu introdotta solo nel '700 da Eulero cui si deve anche l'idea di rappresentare i numeri complessi mediante coppie di numeri reali. Il nome *n*umero immaginario è invece di Cartesio , prima metà del '600, mentre il termine *n*umero complesso fu introdotto da Gauss agli inizi dell'800 .

1.2 *Definizione e proprietà dei numeri complessi*

Definizione 1.1 *Chiamiamo numero complesso una coppia ordinata z di numeri reali (x, y) in cui x si chiama parte reale e si indica con $\Re_e z$ mentre y si chiama parte immaginaria del numero complesso e si indica con $\Im_m z$. L'insieme dei numeri complessi è denotato con $\mathbb{C}$. Se $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$ definiamo*

$$(x_1, y_1) + (x_2 + y_2) = (x_1 + x_2, y_1 + y_2)$$
$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$$



In sostanza quindi $\mathbb{C}$ è l' insieme di coppie di numeri reali, dotato di opportune operazioni, e quindi è naturale rappresentare un numero complesso come un punto del piano cartesiano $\mathbb{R}^2$.

Vista la definizione di somma di due numeri complessi la struttura di $\mathbb{C}$ è quella (come vedremo) di uno spazio vettoriale bidimensionale inoltre in $\mathbb{C}$ è definita l' operazione di moltiplicazione che (vedremo) può essere interpretata graficamente nel piano .

La struttura vettoriale di $\mathbb{C}$ consente di esprimere ogni numero complesso come

$$z = (x, y) = x(1, 0) + y(0, 1)$$

e dal momento che $(a, 0)$ rappresenta il numero complesso con parte reale a e parte immaginaria 0, quindi in sostanza un numero reale; convenzionalmente scriviamo a per $(a, 0)$. In particolare scriviamo 1 per $(1, 0)$ ed analogamente $\imath$ per $(0, 1)$.

Pertanto avremo che

$$z = x\,1 + y\,\imath = x + \imath y = x \cdot 1 + y \cdot i$$

Possiamo verificare allora che la definizione data di moltiplicazione di due numeri complessi consente di operare sull'espressione $x + \imath y$ con le usuali regole del calcolo letterale.

Si verifica in particolare che

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

mentre

$$\imath y = (0,1) \cdot (y,0) = (0,y)$$

e di questo ci si serve per eseguire facilmente operazioni tra numeri complessi.

Chiamiamo $x = \Re_{\mathfrak{e}} z$ ed $y = \Im_{\mathfrak{m}} z$ parte reale e parte immaginaria del numero complesso z.

Si verifica facilmente che le operazioni di somma e di prodotto definite in precedenza sono riconducibili alle usuali regole di calcolo algebrico semplicemente tenendo conto della regola per cui $\imath^2 = -1$.

A titolo di esempio osserviamo che

$$\begin{aligned} z_1 + z_2 = x_1 + \imath y_1 + x_2 + \imath y_2 &= (x_1 + x_2) + \imath(y_1 + y_2) \\ z_1 z_2 = (x_1 + \imath y_1)(x_2 + \imath y_2) &= x_1 x_2 + \imath^2 y_1 y_2 + \imath(x_1 y_2 + y_1 x_2) = \\ &= x_1 x_2 - y_1 y_2 + \imath(x_1 y_2 + y_1 x_2) \end{aligned}$$

Dal momento che abbiamo visto come sia conveniente identificare un numero complesso con una coppia di numeri reali, possiamo rappresentare $\mathbb{C}$ mediante un piano in cui è fissato un sistema di coordinate cartesiane ortogonali monometriche, che viene usualmente indicato come piano di Argand-Gauss. È importante peró osservare che la struttura di $\mathbb{C}$ è più ricca della semplice struttura vettoriale di $\mathbb{R}^2$

Definizione 1.2 *Diciamo che:*

- *diciamo che due numeri complessi sono uguali se sono uguali la loro parte reale e parte immaginaria;*
- $0 = (0,0) \in \mathbb{C}$ *è un elemento neutro perchè* $z + 0 = z \; \forall z \in \mathbb{C}$ *;*
- $1 = (1,0)$ *è un elemento neutro rispetto alla moltiplicazione perchè* $z1 = z$ $\forall z \in \mathbb{C}$*;*
- $w \in \mathbb{C}$ *tale che* $z + w = 0$ *si chiama inverso di* z *rispetto alla somma e si denota con* $-z$*;*
- $w \in \mathbb{C}$ *tale che* $zw = 1$ *si chiama inverso di* z *rispetto alla moltiplicazione e si denota con* $\frac{1}{z}$*;*
- *se* $z = x + \imath y \in \mathbb{C}$ *il suo complesso coniugato* $\bar{z}$ *è definito da*

$$\bar{z} = x - \imath y$$

- *definiamo il modulo* $|z|$ *di un numero complesso mediante la*

$$|z| = \sqrt{x^2 + y^2}$$

- *definiamo argomento di* z, $\arg z$, *mediante la*

$$\tan(\arg(z)) = \frac{y}{x}$$

in accordo con i segni di x *e di* y.

Poichè $\arg z$ *è definito a meno di multipli di* 2π, *definiamo inoltre* $\operatorname{Arg} z$, *l'argomento principale di* $\arg z$, *come quel valore di* $\arg(z)$ *che è compreso in* $[0, 2\pi)$.

Osserviamo che il problema della definizione di $\arg z$ è lo stesso incontrato quando si considerano le coordinate polari nel piano. In particolare **non è sempre vero che**

$$\arg(z) = \arctan \frac{y}{x}$$

in quanto, in tal modo, si otterrebbero soltanto valori compresi in $(-\pi/2, \pi/2)$.

È allora necessario tenere anche conto del segno di x e di y, aggiungendo π ad $\arctan \frac{y}{x}$ nel caso in cui $x < 0$ e definendo $\arg(z) = \pm\pi/2$ nel caso in cui $x = 0$ e, rispettivamente, $y \gtrless 0$

Più precisamente se $z = x + \imath y$,

$$\operatorname{Arg}(z) = \begin{cases} \arctan \frac{y}{x} & \text{se } x > 0 \\ \arctan \frac{y}{x} + \pi & \text{se } x < 0 \\ \arctan \frac{y}{x} + \infty & \text{se } x = 0 \text{ e } y > 0 \\ \arctan \frac{y}{x} + \pi & \text{se } x = 0 \text{ e } y > 0 \end{cases}$$

È immediato verificare che

- $$\overline{z + w} = \bar{z} + \bar{w}$$

- $$\overline{zw} = \bar{z}\bar{w}$$

- $$z\bar{z} = |z|^2$$

ed osservare che per calcolare il quoziente di due numeri complessi si può ricorrere alla formula

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{|z_2|^2}$$

se $z = a + \imath b$ e $w = c + \imath d$ allora

$$\overline{z + w} = \overline{(a + b) + \imath(c + d)} = \\ = (a + b) - \imath(c + d) = \bar{z} + \bar{w}$$

Inoltre

$$\overline{zw} = (ac - bd) - \imath(ad + bc) = \\ = (ac - (-b)(-d)) + \imath(a(-d) + (-b)c) = \\ = (a - \imath b)(c - \imath d) = \bar{z}\bar{w}$$

A proposito del modulo di un numero complesso si può verificare che

$$|z_1 z_2| = |z_1||z_2|$$
$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$$
$$|z_1 + z_2| \leq |z_1| + |z_2|$$
$$||z_1| - |z_2|| \leq |z_1 - z_2|$$

$$|z_1 z_2|^2 = |(x_1 + \imath y_1)(x_2 + \imath y_2)|^2 =$$
$$= |x_1 x_2 + \imath^2 y_1 y_2 + \imath(x_1 y_2 + y_1 x_2)|^2 =$$
$$= |x_1 x_2 - y_1 y_2 + \imath(x_1 y_2 + y_1 x_2)|^2 =$$
$$= (x_1 x_2)^2 + (y_1 y_2)^2 - 2x_1 y_1 x_2 y_2 +$$
$$+ (x_1 y_2)^2 + (y_1 x_2)^+ 2x_1 y_1 x_2 y_2 =$$
$$= (x_1^2 + y_1^2)(x_2^2 + y_2^2) =$$
$$= |z_1|^2 |z_2|^2$$

$$\left|\frac{z_1}{z_2}\right| = \left|\frac{z_1 \bar{z}_2}{|z_2|^2}\right| =$$
$$\frac{|z_1||\bar{z}_2|}{|z_2|^2|} = \frac{|z_1||z_2|}{|z_2|^2} = \frac{|z_1|}{|z_2|} =$$

Come nel piano euclideo, anche nel piano di Argand-Gauss si può usare un riferimento polare (ρ, θ) e si verifica subito che in tal caso si ha

$$\rho = |z| \qquad , \qquad \theta = \operatorname{Arg} z$$

Pertanto un numero complesso $z = x + \imath y$ può essere rappresentato anche nella forma

$$z = |z|\,(\cos(\arg z) + \imath \sin(\arg z)) = \rho(\cos\theta + \imath \sin\theta)$$

Tale rappresentazione si dice forma polare di un numero complesso.
Si può dimostrare che vale la seguente formula di Eulero

$$e^{\imath\theta} = \cos(\theta) + \imath \sin(\theta)$$

Allo scopo si esprimono e^θ, $\sin\theta$ e $\cos\theta$ mediante serie di potenze

$$e^\theta = \sum_{k=0}^{+\infty} \frac{\theta^n}{n!} \quad , \quad \sin\theta = \sum_{k=0}^{+\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!} \quad , \quad \cos\theta = \sum_{k=0}^{+\infty} (-1)^n \frac{\theta^{2n}}{(2n)!}$$

e si osserva che

$$e^{\imath\theta} = \sum_{k=0}^{+\infty} \frac{(\imath\theta)^n}{n!} = \sum_{k=0}^{+\infty} \frac{(\imath\theta)^{2n}}{(2n)!} + \sum_{k=0}^{+\infty} \frac{(\imath\theta)^{2n+1}}{(2n+1)!} =$$
$$= \sum_{k=0}^{+\infty} \frac{(\imath^2)^n(\theta)^{2n}}{(2n)!} + \sum_{k=0}^{+\infty} \frac{\imath(\imath^2)^n(\theta)^{2n+1}}{(2n+1)!} =$$
$$\sum_{k=0}^{+\infty} (-1)^n \frac{\theta^{2n}}{(2n)!} + \imath \sum_{k=0}^{+\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!} = \cos\theta + \imath \sin\theta$$

La formula di Eulero consente di rappresentare un numero complesso in forma esponenziale mediante la

$$z = |z|\,(\cos(\arg z) + \imath \sin(\arg z)) = |z| e^{\imath\theta}$$

Possiamo anche vedere che

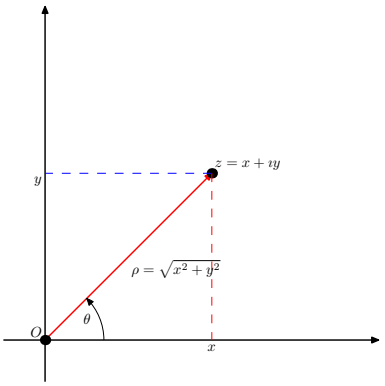
$$e^z = e^{x+\imath y} = e^x(\sin y + \imath \cos y) = e^x e^{\imath y}$$

$$|z_1| = |z_1 - z_2 + z_2| \leq |z_1 - z_2| + |z_2|$$
$$|z_2| = |z_2 - z_1 + z_1| \leq |z_1 - z_2| + |z_1|$$

perciò

$$-|z_1 - z_2| \leq |z_1| - |z_2| \leq |z_1 - z_2|$$

$$|z_1 + z_2|^2 = (z_1 + z^2)(z_1 \bar{+} z_2) = (z_1 + z^2)(\bar{z}_1 + \bar{z}_2) =$$
$$= |z_1|^2 + z_1\bar{z}_2 + z_2\bar{z}_1 + |z_2|^2 =$$
$$= |z_1|^2 + 2\Re z_1\bar{z}_2 + |z_2|^2 \leq$$
$$= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 =$$
$$= (|z_1| + |z_2|)^2$$



Vale anche la pena di osservare che la circonferenza unitaria nel piano complesso, cioè

$$\{z \in \mathbb{C} \ : \ |z| = 1\}$$

può essere rappresentata come

$$\{z \in \mathbb{C} \ : z = \cos(\theta) + \imath \sin(\theta) \qquad \theta \in [0, 2\pi]\}$$

ed anche come

$$\{z \in \mathbb{C} \ : z = e^{\imath\theta} \qquad \theta \in [0, 2\pi]\}$$

Usando la forma esponenziale è facile anche calcolare z^w con $z, w \in \mathbb{C}$, infatti se

$$z = a + \imath b = ue^{\imath\alpha} \qquad , \qquad w = c + \imath d = ve^{\imath\beta}$$

si può calcolare

$$z^w = (ue^{\imath\alpha})^{c+\imath d} = (e^{\log(u)+\imath\alpha})^{c+\imath d}$$

da cui

$$z^w = e^{(c\log(u)-\alpha d)+\imath(\alpha c+d\log(u))}$$

e

$$z^w = e^{(c\log(u)-\alpha d)}\left(\cos(\alpha c + d\log(u)) + \imath\sin(\alpha c + d\log(u))\right)$$

Teorema 1.1 - De Moivre - *Se* $z_1, z_2 \in \mathbb{C}$,

$$\begin{aligned} z_1 &= \rho_1(\cos\theta_1 + \imath\sin\theta_1) \\ z_2 &= \rho_2(\cos\theta_2 + \imath\sin\theta_2) \end{aligned}$$

allora

$$\begin{aligned} z_1 z_2 &= \rho_1\rho_2(\cos(\theta_1+\theta_2) + \imath\sin(\theta_1+\theta_2)) \\ \frac{z_1}{z_2} &= \frac{\rho_1}{\rho_2}(\cos(\theta_1-\theta_2) + \imath\sin(\theta_1-\theta_2)) \end{aligned}$$

La dimostrazione del teorema di De Moivre si riduce ad un semplice calcolo che coinvolge le usuali identità trigonometriche. Infatti

$$\begin{aligned} z_1 z_2 = \rho_1(\cos\theta_1 + \imath\sin\theta_1)\rho_2(\cos\theta_2 + \imath\sin\theta_2) = \\ = \rho_1\rho_2\left((\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + \imath(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)\right) = \\ = \rho_1\rho_2(\cos(\theta_1+\theta_2) + \imath\sin(\theta_1+\theta_2)) \end{aligned}$$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\frac{\cos\theta_1 + \imath\sin\theta_1}{\cos\theta_2 + \imath\sin\theta_2} =$$
$$= \frac{\rho_1}{\rho_2}\frac{(\cos\theta_1 + \imath\sin\theta_1)(\cos\theta_2 - \imath\sin\theta_2)}{\cos^2\theta_2 + \sin^2\theta_2} =$$
$$= \frac{\rho_1}{\rho_2}\left((\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) + \imath(\cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2)\right) =$$
$$= \frac{\rho_1}{\rho_2}(\cos(\theta_1 - \theta_2) + \imath\cos(\theta_1 - \theta_2))$$

Definizione 1.3 *Se $z \in \mathbb{C}$ diciamo che $w \in \mathbb{C}$ è una radice $n-$esima di z se*

$$w^n = z$$

Scriviamo in tal caso

$$w = z^{1/n}$$

Dal teorema di De Moivre, se

$$w = \rho_w(\cos(\theta_w) + \imath\sin(\theta_w))$$

si ha che

$$w^n = \rho_w^n(\cos(n\theta_w) + \imath\sin(n\theta_w))$$

e pertanto si può affermare che se

$$z = \rho_z(\cos\theta_z + \imath\sin\theta_z)$$

allora

$$\rho_w^n = \rho_z \quad , \quad n\theta_w = \theta_z + 2k\pi \qquad k \in \mathbb{Z}$$

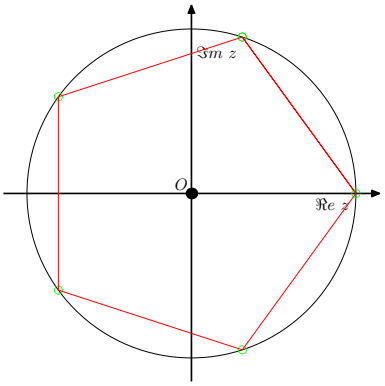
e

$$\rho_w = \rho_z^{1/n} \quad , \quad \theta_w = \frac{\theta_z + 2k\pi}{n} \qquad k \in \mathbb{Z}$$

per cui

$$z^{1/n} = \rho_z^{1/n}\left(\sin\left(\frac{\theta_z + 2k\pi}{n}\right) + \imath\cos\left(\frac{\theta_z + 2k\pi}{n}\right)\right) \qquad k \in \mathbb{Z}$$

Chiaramente nella formula precedente avremo valori distinti solo se $k = 1, 2, \ldots, n-1$.



Un caso particolarmente interessante si trova considerando l'equazione

$$z^n = 1$$

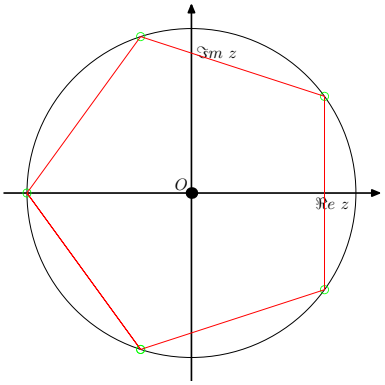
le cui soluzioni si dicono radici $n-$esime dell'unità.

Si verifica subito che le radici $n-$esime dell'unità hanno modulo uguale ad 1 e pertanto si trovano tutte su una circonferenza unitaria, centrata nell'origine, del piano complesso; inoltre, dal momento che $z = 1 + \imath 0$ risolve l'equazione, una delle radici coincide con il punto $(1, 0)$ del piano, mentre le altre sono collocate sui vertici di un poligono

regolare di n lati inscritto nella circonferenza unitaria con un vertice coincidente con $(1,0)$.

È anche interessante considerare l'equazione

$$z^{2n} = -1$$



In questo caso le radici hanno ancora modulo uguale ad 1 e pertanto si trovano tutte su una circonferenza unitaria, centrata nell'origine, del piano complesso; inoltre, dal momento che $z = 0 + \imath$ una delle radici coincide con il punto $(0,1))$ del piano, mentre le altre sono collocate sui vertici di un poligono regolare di n lati inscritto nella circonferenza unitaria con un vertice coincidente con $(0,1)$.

1.3 *Il teorema fondamentale dell'algebra*

Sia

$$p_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

un polinomio di grado n definito per $z \in \mathbb{C}$ con coefficienti $a_n \in \mathbb{C}$.

Il problema di determinare gli zeri di p_n e quindi di risolvere l'equazione

$$p_n(z) = 0$$

è facilmente risolubile se $n = 1, 2$.

1.3.1 *Equazioni di Primo Grado*

Se consideriamo

$$az + b = 0$$

avremo

$$z = -\frac{b}{a} = \frac{b\bar{a}}{a\bar{a}} = \frac{b\bar{a}}{|a|^2}$$

1.3.2 *Equazioni di Secondo Grado*

Consideriamo

$$az^2 + bz + c = a\left(z^2 + \frac{b}{a}z + \frac{c}{a}\right) = a\left(\left(z + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + c\right)$$

Avremo

$$az^2 + bz + c = 0$$

se e solo se

$$\left(z + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + c = 0$$

da cui

$$z + \frac{b}{2a} = \left(\frac{b^2}{4a^2} + c\right)^{1/2}$$

e

$$z = \frac{b}{2a} + \left(\frac{b^2}{4a^2} + c\right)^{1/2}$$

Ovviamente $\left(\frac{b^2}{4a^2} + c\right)^{1/2}$ assumerà due valori complessi , eventualmente coincidenti.

Le equazioni di secondo e di terzo grado possono essere risolte anche se con maggiore difficoltà

1.3.3 Equazioni di terzo grado.

Si tratta di equazioni del tipo

$$ax^3 + bx^2 + cx + d = 0$$

Se poniamo

$$x = y - \frac{b}{3a}$$

possiamo riscrivere l'equazione data come:

$$ay^3 + \left(c - \frac{b^2}{3a}\right)y + \frac{2b^3}{27a^2} - \frac{bc}{3a} + d = 0$$

posto

$$p = \frac{c}{a} - \frac{b^2}{3a^2} \qquad , \qquad q = \frac{2b^3}{27a^3} - \frac{bc}{3a^2} + \frac{d}{a}$$

otteniamo unèquazione equivalente a quella data, ma più semplice, che chiamiamo equazione ridotta.

$$y^3 + py + q = 0$$

Se ora osserviamo che

$$(u+v)^3 = 3uv(u+v) + (u^3 + v^3)$$

e poniamo $y = u + v$, otteniamo

$$y^3 + py = (u+v)^3 + p(u+v) = u^3 + v^3 + (3uv + p)(u+v) = q$$

e pertanto possiamo trovare y se siamo in grado di trovare u, v in modo che

$$\begin{cases} 3uv = -p \\ u^3 + v^3 = q \end{cases} \qquad \text{ovvero} \qquad \begin{cases} u^3 v^3 = -\left(\frac{p}{3}\right)^3 \\ u^3 + v^3 = q \end{cases}$$

pertanto u^3 e v^3 devono essere soluzioni dell'equazione

$$z^2 - qz - \frac{p^3}{27} = 0$$

Allora

$$u^3 = \frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \quad , \quad v^3 = \frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

Per concludere occorre estrarre le radici cubiche complesse e scegliere u e v in modo che uv sia reale e considerare $u + v$.

Poichè cerchiamo esattamente 3 soluzioni mentre le radici cubiche di u e di v sono 3 e le scelte possibili sono, di conseguenza, 9, dovremo agire con un po' di attenzione e, nel caso i coefficienti siano reali, possiamo ottenere qualche semplificazione.

Supponiamo allora che i coefficienti siano reali.

In tal caso è utile osservare che u^3 e v^3 sono soluzioni di un'equazione di secondo grado a coefficienti reali e quindi sono l'uno il coniugato dell'altro. Se

$$ae^{\imath\alpha}, ae^{-\imath\alpha}$$

è la loro espressione polare, possiamo affermare che u può assumere uno dei seguenti valori

$$u_0 = \sqrt[3]{a}e^{\imath\alpha/3} \quad , \quad u_1 = \sqrt[3]{a}e^{\imath(\alpha/3+2/3\pi)} \quad , \quad u_2 = \sqrt[3]{a}e^{\imath(\alpha/3+4/3\pi)}$$

mentre v può assumere uno dei seguenti valori

$$v_0 = \sqrt[3]{a}e^{-\imath\alpha/3} \quad , \quad v_1 = \sqrt[3]{a}e^{-\imath(\alpha/3+2/3\pi)} \quad , \quad v_2 = \sqrt[3]{a}e^{-\imath(\alpha/3+4/3\pi)}$$

Pertanto

$$u_h v_k = \sqrt[3]{a^2}e^{\imath(\alpha/3-\alpha/3+2/3(k+h)\mathcal{P})i} = \sqrt[3]{a^2}e^{\imath(2/3(k+h)\pi)}$$

con $h, k = 0, 1, 2$ e $(h + k = 0, 1, 2, 3, 4)$

dovendo essere $u_h v_k = -p/3$, occorre scegliere le radici in modo che $\sqrt[3]{a^2}e^{\imath(2/3(k+h)\pi)} \in \mathbb{R}$ e ciò avviene per

$$(h, k) = (0, 0) \quad , \quad (h, k) = (1, 2) \quad , \quad (h, k) = (2, 1)$$

Le soluzioni corrispondenti saranno quindi:

$$u_0 v_0 = \frac{2}{3}\sqrt[3]{a^2}\cos(\alpha/3) \quad , \quad u_1 v_2 = \sqrt[3]{a^2}\cos(\alpha/3)(-1+\imath\sqrt{3}) \quad , \quad u_2 v_1 = \sqrt[3]{a^2}\cos(\alpha/3)(-1-\imath\sqrt{3})$$

1.3.4 Equazioni di quarto grado.

Si tratta di equazioni della forma

$$ax^4 + bx^3 + cx^2 + dx + e = 0$$

Se poniamo

$$x = y - \frac{b}{4a}$$

Ci riduciamo all'equazione

$$y^4 + py^2 + qy + r = 0$$

dove si sia definito

$$p = \frac{c}{a} - \frac{3b^2}{8a^2} \quad , \quad q = \frac{b^3}{8a^3} - \frac{bc}{2a^2} + \frac{d}{a}$$
$$r = \frac{b^2c}{16a^3} - \frac{3b^4}{256a^4} - \frac{bd}{4a^2} + \frac{e}{a}$$

Se poniamo

$$x = u + v + w$$

avremo, elevando al quadrato,

$$x^2 - (u^2 + v^2 + w^2) = 2(uv + uw + vw)$$

ed ancora, sempre elevando al quadrato,

$$x^4 - 2(u^2 + v^2 + w^2)x^2 + (u^2 + v^2 + w^2)^2 = 4[(u^2v^2 + u^2w^2 + v^2w^2) + 2uvw(u + v + w)]$$

da cui

$$x^4 - 2(u^2 + v^2 + w^2)x^2 + 8uvw(u + v + w) + (u^2 + v^2 + w^2)^2 - 4(u^2v^2 + u^2w^2 + v^2w^2) = 0$$

e

$$x^4 - 2(u^2 + v^2 + w^2)x^2 + 8uvwx + (u^2 + v^2 + w^2)^2 - 4(u^2v^2 + u^2w^2 + v^2w^2) = 0$$

Possiamo asserire che tale equazione è equivalente all'equazione data se

$$\begin{cases} u^2 + v^2 + w^2 = -\frac{p}{2} \\ uvw = -\frac{q}{8} \\ (u^2 + v^2 + w^2)^2 - 4(u^2v^2 + u^2w^2 + v^2w^2) = r \end{cases}$$

ovvero se

$$\begin{cases} u^2 + v^2 + w^2 = -\frac{p}{2} \\ u^2v^2w^2 = \frac{q^2}{64} \\ (u^2v^2 + u^2w^2 + v^2w^2) = \frac{p^2 - 4r}{16} \end{cases}$$

pertanto u^2, v^2 e w^2 dovranno essere scelti tra le soluzioni dell'equazione di terzo grado

$$z^3 + \frac{p}{2}z^2 + \left(\frac{p^2}{16} - \frac{r}{4}\right)z - \frac{q^2}{64} = 0$$

in quanto è noto che i coefficienti di grado $2, 1, 0$ sono la somma, la somma dei prodotti a due a due e la somma dei prodotti tre a tre delle soluzioni.

Potremo infine ricavare y e poi x avendo la sola cura di scegliere u, v e w in modo che uvw abbia segno opposto a quello di q.

Alternativamente possiamo osservare che da

$$x^4 + x^3 + bx^2 + cx + d = 0$$

si ha

$$x^4 + x^3 + \frac{a^2}{4}x^2 = -(bx^2 + cx + d) + \frac{a^2}{4}x^2$$

e possiamo ottenere

$$\left(x^2 + \frac{a}{2}x\right)^2 = \left(\frac{a^2}{4} - b\right)x^2 - cx - d$$

Qualora il secondo membro sia esso pure il quadrato di binomio di secondo grado possiamo ridurci a due equazioni di secondo grado, in caso contrario possiamo aggiungere ad ambo i membri $(x^2 + ax/2)z + z^2/4$ in modo da ottenere

$$\left(x^2 + \frac{a}{2}x + \frac{z}{2}\right)^2 = \left(\frac{a^2}{4} - b + z\right)x^2 - \left(c - \frac{a}{2}z\right)x - d + \frac{1}{2}z^2$$

Possiamo ora scegliere z in modo che il secondo membro sia il quadrato di un trinomio di secondo grado, imponendo che il suo discriminante sia nullo. Tale condizione si riduce alla seguente equazione di terzo grado

$$z^3 - bz^2 + (ac - 4d)z + 4bd - a^2d - c^2 = 0$$

Trovata una soluzione reale di tale equazione ci si riduce ad una eguaglianza di quadrati e quindi ad una coppia di equazioni di secondo grado.

Per le equazioni di quinto grado o di grado superiore non è possibile trovare una formula risolutiva, possiamo però dimostrare che ammettono un numero di soluzioni complesse pari al loro grado. È utile a questo scopo provare che vale il seguente risultato.

1.3.5 Fattorizzazione di Polinomi

Lemma 1.1 *Se p è un polinomio a coefficienti complessi definito in $\mathbb{C}$ di grado n e se $z_0 \in \mathbb{C}$ annulla p, cioè se $p(z_0) = 0$, allora esiste un polinomio q definito in $\mathbb{C}$ di grado $n-1$ tale che*

$$p(z) = (z - z_0)q(z)$$

DIMOSTRAZIONE. L'algoritmo di divisione dei polinomi consente di affermare che esistono un polinomio q di grado $n-1$ e $r \in \mathbb{C}$ tali che

$$p(z) = (z - z_0)q(z) + r$$

Pertanto per $z = z_0$ si ha $r = 0$ e la tesi. □

Il lemma precedente permette di riscrivere il polinomio come somma di potenze centrate in $z_0 \in \mathbb{C}$; sia p un polinomio di grado n a coefficienti complessi e sia $z_0 \in \mathbb{C}$. poichè $p(z) - p(z_0)$ si annulla per $z = z_0$ avremo.

$$p(z) - p(z_0) = (z - z_0)p_1(z)$$

con p_1 polinomio di grado $n-1$.

Quindi

$$\begin{aligned} p(z) - p(z_0) &= (z - z_0)(p_1(z) - p_1(z_0) + p_1(z_0)) = \\ &= (z - z_0)p_1(z_0) + (z - z_0)(p_1(z) - p_1(z_0)) \end{aligned}$$

Similmente

$$p_1(z) - p_1(z_0) = (z - z_0)p_2(z)$$

e sostituendo

$$\begin{aligned} p(z) - p(z_0) &= (z - z_0)p_1(z_0) + (z - z_0)(p_1(z) - p_1(z_0)) = \\ &= (z - z_0)p_1(z_0) + (z - z_0)^2 p_2(z) \end{aligned}$$

con p_2 polinomio di grado $n-2$. Procedendo per n passi si ottiene infine

$$\begin{aligned} p(z) - p(z_0) = (z - z_0)p_1(z_0) &+ (z - z_0)^2 p_2(z_0) + \\ &+ (z - z_0)^3 p_3(z_0) + \cdots + (z - z_0)^n p_n(z) \end{aligned}$$

essendo p_n un polinomio di grado 0 e quindi una costante.

Abbiamo così riscritto $p(z)$ nella forma

$$p(z) = \sum_{k=0}^{n} b_k (z - z_0)^k$$

dove $b_k \in \mathbb{C}$ e almeno $b_n \neq 0$

È possibile dimostrare che:

Teorema 1.2 - Teorema Fondamentale dell'Algebra *Ogni polinomio di grado n ha almeno una radice complessa.*

DIMOSTRAZIONE. Sia

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + ... + a_1 z + a_0$$

con $a_n \in \mathbb{C}$, $a_n \neq 0$ e un polinomio a coefficienti complessi di grado $n \geq 1$; si ha

$$|p(z)| \ = \ |z|^n \left| a_n + \frac{a_{n-1}}{z} + ... + \frac{a_0}{z^n} \right|$$

e quindi

$$\lim_{|z| \to +\infty} |p(z)| \ = \ +\infty$$

Ne segue che $|p(z)|$ ammette minimo in $\mathbb{C}$; sia $z_0 \in \mathbb{C}$ il punto in cui il minimo è assunto, dimostriamo che $p(z_0) = 0$.

Infatti se fosse $|p(z_0)| \neq 0$, potremmo trovare valori di $|p(z)|$ minori di $|p(z_0)|$ come vediamo qui di seguito.

Dal momento che $p(z) - p(z_0)$ si annulla in z_0 , possiamo applicare il lemma precedente ed ottenere

$$p(z) - p(z_0) = (z - z_0) \left(\sum_{i=0}^{n-1} \beta_i z^i \right) = (z - z_0) \left(\sum_{i=0}^{n-1} b_i (z - z_0)^i \right)$$

per cui si ha,

$$p(z) = p(z_0) + \sum_{i=1}^{n} b_i (z - z_0)^i$$

dove $b_i \in \mathbb{C}$ ed i coefficienti b_i non sono tutti nulli; sia, $b_k \neq 0$ il primo coefficiente non nullo, $k \geq 1$

Se $\lambda > 0$ e $u \in \mathbb{C}$,

$$p(z_0 + \lambda u) = p(z_0) + b_k \lambda^k u^k + \lambda^k \omega(\lambda)$$

con ω infinitesimo per $\lambda \to 0$.

Scelto u tale che $b_k u^k = -p(z_0)$, risulta (poichè $p(z_0) \neq 0$)

$$p(z_0 + \lambda u) = p(z_0) - \lambda^k p(z_0) + \lambda^k p(z_0) \frac{\omega(\lambda)}{p(z_0)}$$

e

$$p(z_0 + \lambda u) = p(z_0)(1 - \lambda^k + \lambda^k \omega(\lambda))$$

avendo ancora indicato con $\omega(\lambda)$ il rapporto $\omega(\lambda)/p(z_0)$.

Pertanto

$$|p(z_0 + \lambda u)| \ = \ |p(z_0)| \, |1 - \lambda^k + \lambda^k \omega(\lambda)|$$

e, per $\lambda < 1$

$$|1 - \lambda^k + \lambda^k \omega(\lambda)| \ \leq \ |1 - \lambda^k| + \lambda^k |\omega(\lambda)| \ = \\ = 1 - \lambda^k + \lambda^k |\omega(\lambda)| \ = 1 - \lambda^k (1 - |\omega(\lambda)|)$$

Ne segue che $1 - |\omega(\lambda)|$ si mantiene strettamente positivo in un intorno di zero, (perché $\omega \to 0$) e quindi tale quantità è strettamente minore di 1 in un intorno (destro) di zero.

Allora in tale intorno

$$|p(z_0 + \lambda u)| \ < \ |p(z_0)|$$

e questo è assurdo.

□

Come conseguenza di questo risultato e della fattorizzazione di un polinomio possiamo allora enunciare anche il seguente

Teorema 1.3 *Se p è un polinomio a coefficienti complessi definito in $\mathbb{C}$ di grado n allora esistono $z_1, z_2, \ldots, z_n \in \mathbb{C}$ non necessariamente distinti tali che*

$$p(z) = \alpha(z - z_1)(z - z_2) \cdots (z - z_n)$$

Ovviamente $z_1, z_2, \ldots, z_n \in \mathbb{C}$ sono tutte e sole le soluzioni dell'equazione $p(z) = 0$; nel caso in cui z_k compaia in più di un fattore diciamo che è una soluzione multipla per l'equazione $p(z) = 0$ e chiamiamo molpteplicità di z_k il numero dei fattori in cui compare.

Esprimiamo questo concetto dicendo che un polinomio a coefficienti complessi definito in $\mathbb{C}$ di grado n ha esattamente n radici complesse contate con la loro molteplicità.

È anche utile osservare che qualora

$$p_n(z) = \sum_{k=0}^{n} a_k z^k$$

sia un polinomio di grado n i cui coefficienti a_k sono reali ($a_k \in \mathbb{R}$) e z_0 sia una radice del polinomio, si abbia cioè

$$p_n(z_0) = 0$$

allora anche il suo complesso coniugato $\bar{z}_0$ è radice del polinomio.

Infatti sia

$$\sum_{k=0}^{n} a_k z_0^k$$

Poichè $\bar{a}_k = a_k$,

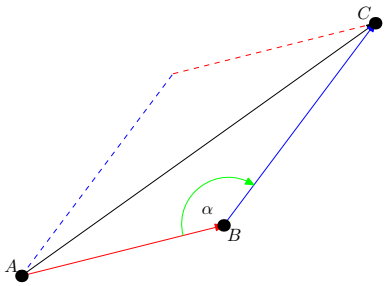
$$\sum_{k=0}^{n} a_k \bar{z}_0^k = \sum_{k=0}^{n} \bar{a}_k \bar{z}_0^k = \sum_{k=0}^{n} \overline{(a_k z_0^k)} = \overline{\left(\sum_{k=0}^{n} a_k z_0^k \right)} = 0$$

in quanto coniugato di

$$\sum_{k=0}^{n} a_k z_0^k = 0$$

2. *Introduzione al concetto di vettore*

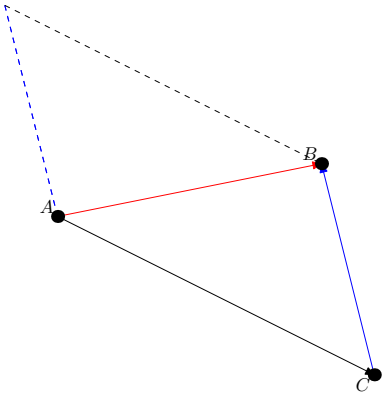
Il concetto di vettore nasce tra il 1545, quando Cardano affrontò per la prima volta problematiche associate a quelli che ora chiamiamo numeri complessi ed il 1846, quando Hamilton introdusse per la prima volta i termini scalare e vettore nella sua trattazione dei quaternioni.



L'esigenza che portò sua alla nascita era quella di rappresentare concetti geometrici e fisici quali lo spostamento, la velocità, l'accelerazione o la forza mediante uno strumento che ne permettesse la manipolazione così come l'algebra consentiva di manipolare grandezze numeriche.

Il primo a sentire questa esigenza fu Leibniz nel 1679, seguito da Newton nel 1687 cui si deve l'idea di parallelogrammo delle forze per descrivere la risultante di due forze che agiscono sullo stesso corpo.

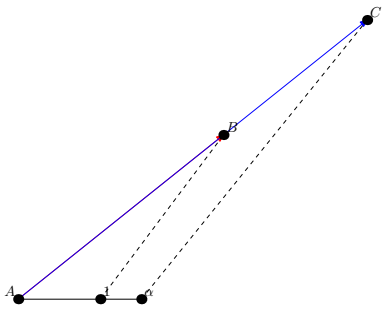
Forze, spostamenti, velocità, accelerazione hanno in comune il fatto che sono individuabili , ad esempio nel piano, mediante una direzione ed una lunghezza: per spostarci da un punto ad un'altro ci muoviamo nella direzione che parte dal primo punto e va verso il secondo e percorriamo una lunghezza pari alla distanza che c'è tra i due punti.



Ad esempio, se ci trasferiamo da un punto A ad un punto B nel piano, possiamo descrivere lo spostamento mediante il segmento AB orientato da A a B e se poi ci trasferiamo in C il nuovo spostamento è identificato dal segmento BC e si vede subito che lo spostamento risultante dai due spostamenti in sequenza si può rappresentare mediante il segmento AC che è la diagonale del parallelogrammo di lati AB e BC.

È evidente quindi che la somma di due spostamenti si puo' trovare usando l'idea di parallelogrammo introdotta da Newton per comporre l'azione di due forze.

La stessa idea può anche essere usata per velocità, accelerazioni ed altro ancora.

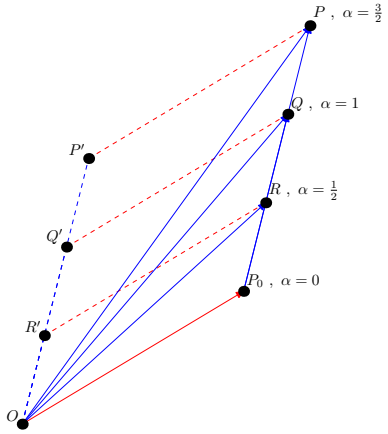


Se consideriamo due vettori AB e AC possiamo anche definire il vettore differenza BC mediante la regola del parallelogrammo. Possiamo infatti trovare la differenza BC tra i vettori AB e AC semplicemente considerando il lato BC del parallelogramma avente AB come diagonale ed un lato coincidente con AC . Possiamo infine identifica-

re il prodotto di un vettore AB per un numero reale α, con il vettore $AC = \alpha AB$, che ha la stessa direzione e verso di AB e lunghezza pari alla lunghezza di AB moltiplicata per α.

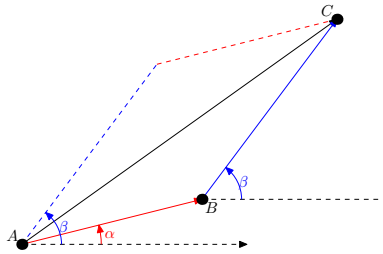
Con semplici operazioni di questo tipo possiamo inoltre individuare i punti di una retta nel piano come l'insieme di quei punti a cui si può arrivare partendo da un punto fissato P_0 e muovendosi da lì nella direzione verso un punto Q percorrendo una distanza proporzionale, secondo un fattore α alla distanza che intercorre tra P_0 e Q

$$OP = OP_0 + \alpha P_0 Q$$



Tornando al considerare la somma di spostamenti, è evidente che il primo spostamento da A a B è individuato dai punti iniziale e finale, così come anche il secondo da B a C , tuttavia il secondo potrebbe essere anche individuato assegnando la lunghezza del segmento BC e l'angolo che il segmento BC forma con il segmento AB.

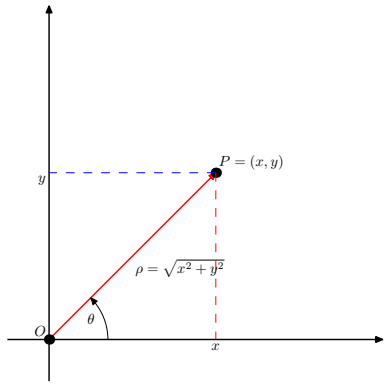
In altre parole partendo da A possiamo solo dire che ci recheremo in B mentre una volta giunti in B possiamo dire che ci muoveremo per una certa distanza nella direzione che forma un certo angolo α rispetto alla direzione precedente.



L'apparente dissimmetria può essere eliminata se introduciamo esplicitamente un riferimento, ad esempio, assegnando una semiretta nel piano rispetto al quale misurare gli angoli che definiscono la direzione dello spostamento.

Nel piano, da Cartesio e Fermat (1637) in poi, è naturale usare un riferimento, che usualmente si chiama cartesiano, che permette di individuare ogni punto mediante le sue coordinate.

In un riferimento cartesiano un punto è identificato da una coppia di numeri (x, y) ma e' facile vedere che se definiamo $\rho = \sqrt{x^2 + y^2}$ e θ l'angolo formato dalla semiretta dall'origine per il punto (x, y) con il semiasse positivo delle x .

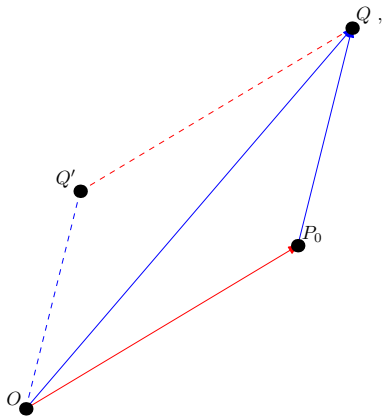


Si ha $(x, y) = (\rho \cos\theta, \rho \sin\theta)$ ed e' evidente l'analogia con quanto detto in precedenza per il vettore che individua uno spostamento: ne segue che i vettori che identificano spostamenti dall'origine sono rappresentabili con i punti del piano cartesiano, cioè con una coppia di numeri.

Potrebbe sembrare una limitazione il fatto che i vettori così definiti individuano solo spostamenti dall'origine, tuttavia si vede subito che ciò non è vero.

Infatti sia P un punto del piano che si possa raggiungere partendo da P_0 e muovendosi verso Q, sia, cioè, P un generico punto della retta passante per P_0 e per Q. Possiamo individuare P mediante la

$$OP = OP_0 + \alpha P_0 Q \qquad \forall \alpha \in \mathbb{R}$$



dove OP è un vettore che indica uno spostamento dal'origine, mentre P_0Q identifica uno spostamento da P_0 verso Q

D'altra parte, se OQ' ha la stessa direzione e la stessa lunghezza di P_0Q il parallelogramma generato da OP_0 e da P_0Q coincide con quello generato da OP_0 e da OQ' e quindi

$$OP = OP_0 + \alpha OQ'$$

Ciò mostra che è sufficiente introdurre un formalismo che contempli solo vettori del tipo OP essendo O l'origine del piano cartesiano.

È anche utile osservare che, in questo modo, un vettore è individuato da un punto P del piano cartesiano e quindi da una coppia di numeri reali, che ne costituiscono le coordinate cartesiane ($P = (x, y)$).

Se $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, usando la regola del parallelogrammo, si vede che le coordinate del vettore $OP = OP_1 + OP_2$ si ricavano da quelle di P_1 e P_2 mediante una somma componente per componente.

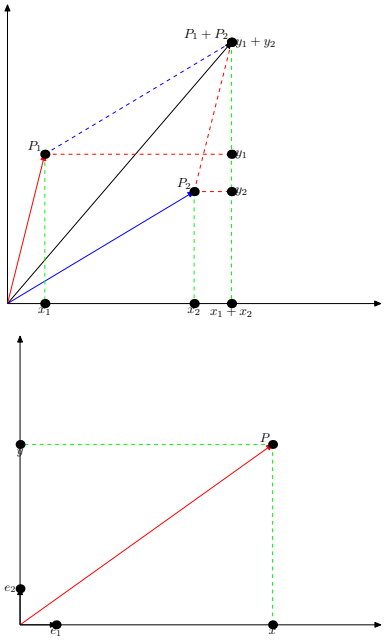
$$(x, y) = (x_1 + x_2, y_1 + y_2)$$

Infine possiamo anche vedere che

$$P = (x, y) = x(1, 0) + y(0, 1) = xe_1 + ye_2$$

essendo e_1, e_2 i vettori di lunghezza 1 lungo le direzioni degli assi cartesiani, è evidente il vantaggio che si può ottenere legando il concetto di vettore alla struttura cartesiana del piano e al concetto di somma di due vettori e di prodotto di un vettore per uno scalare.

Tutto ciò può essere inquadrato nel concetto di spazio vettoriale che definiamo della sezione successiva.



3. Spazi vettoriali

La linearità è una proprietà essenziale e molto semplice il cui uso è largamente diffuso in Matematica, ed il concetto di spazio vettoriale è introdotto proprio allo scopo di definirla e di codificarne le principali proprietà.

3.1 Definizione di Spazio vettoriale

Definizione 3.1 *Diciamo che $\mathcal{F}$ è un campo se è un insieme su cui sono definite due operazioni $+$, che chiamiamo somma, e $\cdot$, che chiamiamo prodotto, tali che valgano le seguenti proprietà:*

- $\forall a, b \in \mathcal{F}$ *si ha che* $a + b, a \cdot b \in \mathcal{F}$
  (Chiusura rispetto $a +$ e $a \cdot$)
- $\forall a, b, c \in \mathcal{F}$ *si ha che* $a + (b + c) = (a + b) + c$ *e* $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
  (Associatività di $+$ e $\cdot$)
- $\forall a, b \in \mathcal{F}$ *si ha che* $a + b = b + a$ *e* $a \cdot b = b \cdot a$
  (Commutatività di $+$ e $\cdot$)
- $\forall a, b, c \in \mathcal{F}$ *si ha che* $a \cdot (b + c) = a \cdot b + a \cdot c$
  (Distributività di $\cdot$ rispetto a $+$)
- *esiste* $0 \in \mathcal{F}$ *tale che* $a + 0 = a$
  (Esistenza di un elemento neutro rispetto a $+$)
- *esiste* $1 \in \mathcal{F}$ *tale che* $a \cdot 1 = a$
  (Esistenza di un elemento neutro rispetto a $\cdot$)
- $\forall a \in \mathcal{F}$ *esiste* $x \in \mathcal{F}$ *tale che* $a + x = 0$
  (Esistenza di un inverso rispetto a $+$)
- $\forall a \in \mathcal{F}$ *esiste* $x \in \mathcal{F}$ *tale che* $a \cdot x = 0$
  (Esistenza di un inverso rispetto a $\cdot$)

Sono esempi importanti di campi l'insieme $\mathbb{R}$ e l'insieme $\mathbb{C}$.

Per quel che riguarda il seguito assumeremo sempre che $\mathcal{F}$ sia $\mathbb{R}$ o $\mathbb{C}$, tuttavia le definizioni possono essere date per un generico campo.

Definizione 3.2 *Diciamo che $\mathcal{V}$ è uno spazio vettoriale sul campo $\mathcal{F}$ (chiamiamo scalari gli elementi in $\mathcal{F}$ e vettori gli elementi in $\mathcal{V}$) se è un insieme su cui sono definite due operazioni*

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$$

che chiamiamo somma (di vettori), e

$$\cdot : \mathcal{F} \times \mathcal{V} \to \mathcal{V}$$

che chiamiamo prodotto (di un vettore per uno scalare), tali che valgano le seguenti proprietà:

- $\forall u, v \in \mathcal{V}$ *e* $\forall a, b \in \mathcal{F}$ *si ha che* $a \cdot u + b \cdot v \in \mathcal{V}$
  (Chiusura rispetto a somma e a prodotto per uno scalare)
- $\forall u, v, w \in \mathcal{V}$ *e* $\forall a, b \in \mathcal{F}$ *si ha che* $u + (v + w) = (u + v) + w$ *e* $a \cdot (b \cdot v) = (a \cdot b) \cdot v$
  (Associatività di somma di vettori e prodotto per uno scalare)
- $\forall u, v \in \mathcal{V}$ *si ha che* $u + v = v + u$
  (Commutatività della somma di vettori)
- $\forall u, v, \in \mathcal{V}$ $\forall a, b, \in \mathcal{F}$ *si ha che* $a \cdot (u + v) = a \cdot u + a \cdot v$ *e* $(a + b) \cdot u = a \cdot u + b \cdot u$
  (Distributività del prodotto per uno scalare)
- *esiste* $0 \in \mathcal{V}$ *tale che* $u + 0 = u$
  (Esistenza di un elemento neutro)
- $\forall u \in \mathcal{V}$ *esiste* $x \in \mathcal{V}$ *tale che* $u + x = 0$
  (Esistenza di un inverso)
- $\forall u \in \mathcal{V}$ *si ha* $1 \cdot u = u$

Ovviamente $\mathcal{V} = \{0\}$ è uno spazio vettoriale che indichiamo con il nome di spazio vettoriale banale e altrettanto ovviamente supporremo sempre che esista almeno un elemento non nullo in $\mathcal{V}$.

Definizione 3.3 *Diciamo che un sottoinsieme $\mathcal{U}$ non vuoto di $\mathcal{V}$ è un sottospazio vettoriale di $\mathcal{V}$ se*

- $\forall u, v \in \mathcal{U}$ *si ha* $u + v \in \mathcal{U}$
- $\forall u \in \mathcal{U}$, $\forall \alpha \in \mathcal{F}$ *si ha* $\alpha u \in \mathcal{U}$

Ovviamente ogni sottospazio vettoriale dello spazio vettoriale V è esso stesso spazio vettoriale.

Definizione 3.4 *Se I è un insieme finito di vettori, indichiamo con $\#I$ il numero dei suoi elementi.*

Definizione 3.5 *Chiamiamo combinazione lineare dei vettori $v_1, v_2, \dots, v_n \in \mathcal{V}$ a coefficienti $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathcal{F}$ il vettore*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in \mathcal{V}$$

Definizione 3.6 *Diciamo che i vettori $v_1, v_2, \dots, v_n \in \mathcal{V}$ sono linearmente indipendenti se da*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

si può dedurre che

$$\alpha_k = 0 \ , \ \forall k = 1, 2, \dots n$$

Chiamiamo $\mathcal{I}$ la famiglia di tutti gli insiemi di vettori linearmente indipendenti in $\mathcal{V}$.

In altre parole i vettori $v_1, v_2, \dots, v_n \in \mathcal{V}$ sono linearmente indipendenti se non è possibile scrivere uno di essi come combinazione lineare dei rimanenti.

Per contro

I vettori $v_1, v_2, \dots, v_n \in \mathcal{V}$ sono linearmente dipendenti se esistono $\alpha_k \in \mathcal{F}$, $k = 1, 2, \dots n$ non tutti nulli tali che

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

Teorema 3.1 *Valgono i seguenti fatti:*

- *Se $\mathcal{I}$ è un insieme di vettori linearmente indipendenti allora $0 \notin \mathcal{I}$*
- *Due vettori sono linearmente dipendenti se e solo se uno è multiplo dell'altro.*
- *Almeno uno degli elementi di un insieme finito di vettori linearmente dipendenti può essere ottenuto come combinazione lineare degli altri.*

Infatti se $\mathcal{I} = \{v_1, v_2, \dots, v_n\}$, e $v_k = 0$ allora $1 \cdot 0 + \sum 0 \cdot v_k$ è una combinazione lineare nulla a coefficienti non tutti nulli.

infatti se $av_1 + bv_2 = 0$ con a e b non entrambi nulli, si ha, supponendo ad esempio $a \neq 0$, $v_1 = \frac{b}{a} v_2$.

infatti se $\{v_1, v_2, \dots, v_n\}$ sono linearmente dipendenti esiste una combinazione lineare nulla $\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_n v_n = 0$ a coefficienti non tutti nulli. Se supponiamo che $\alpha_1 \neq 0$ possiamo scrivere che

$$v_1 = -\alpha_2/\alpha_1 v_2 + \cdots \alpha_n v/\alpha_1 v_n$$

Nello stesso modo possiamo esprimere in funzione degli altri ogni elemento che figura nella combinazione lineare con coefficiente diverso da 0.

Definizione 3.7 *Consideriamo un insieme di vettori $B \subset \mathcal{V}$; diciamo che $\mathbb{S}(B)$ è il sottospazio generato da B se contiene tutte e sole le combinazioni lineari di elementi di B cioè:*

$$\mathbb{S}(B) = \{u \in \mathcal{V} \ : \ \exists \alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R} \ \ con \ \ u = \sum_i \alpha_i v_i\}$$

Definizione 3.8 *Diciamo che $B = \{v_1, v_2, \dots, v_n\} \subset \mathcal{V}$ è un insieme di generatori per $\mathcal{V}$ se $\mathbb{S}(B) = \mathcal{V}$.*

Chiamiamo $\mathcal{S}$ la famiglia di tutti gli insiemi di generatori di $\mathcal{V}$.

Definizione 3.9 *Definiamo base di uno spazio vettoriale $\mathcal{V}$ un sottoinsieme $B \subset \mathcal{V}$ tale che*

- *gli elementi di B siano linearmente indipendenti*

- $\mathbb{S}(B) = \mathcal{V}$

Quindi B è una base per lo spazio $\mathcal{V}$ se le combinazioni lineari di elementi di B descrivono completamente $\mathcal{V}$ ed inoltre se non è possibile esprimere nessun elemento di B, come combinazione lineare dei rimanenti.

Questa affermazione è formalizzata nel teorema seguente.

Teorema 3.2 *$B \subset \mathcal{V}$ è una base per $\mathcal{V}$ se e solo se ogni $u \in \mathcal{V}$ si può esprimere in maniera univoca come combinazione lineare di elementi di B*

DIMOSTRAZIONE. Se B è una base, allora $\mathbb{S}(B) = \mathcal{V}$ e quindi ogni elemento di $\mathcal{V}$ si può esprimere come combinazione lineare di elementi di B; resta da verificare l'univocità. Se fosse

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \qquad \text{e} \qquad u = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

con $\alpha_k, \beta_k \in \mathbb{R}$ e $v_k \in B$ si avrebbe

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = 0$$

e dal momento che $v_k \in B$ sono linearmente indipendenti avremmo

$$\alpha_k = \beta_k \ , \ \forall k = 1, 2, \ldots, n$$

Viceversa se ogni $u \in \mathcal{V}$ si può esprimere come combinazione lineare di elementi di B, allora B genera $\mathcal{V}$.

Inoltre poichè

$$0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \qquad \text{se} \qquad \alpha_k = 0, \quad \forall k$$

essendo la rappresentazione di ogni elemento, e quindi anche di 0, univoca si deduce che

$$0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \qquad \text{se e solo se} \qquad \alpha_k = 0, \quad \forall k$$

e la lineare indipendenza di B □

Definizione 3.10 *Uno spazio vettoriale $\mathcal{V}$ si dice di dimensione finita se esiste un sottoinsieme finito $B \subset \mathcal{V}$ tale che $\mathbb{S}(B) = \mathcal{V}$.*

Teorema 3.3 *Per ogni $S \in \mathcal{S}$ e per ogni $I \in \mathcal{I}$ definiamo*

$$n_s = \min_{S \in \mathcal{S}} \#S \qquad , \qquad n_i = \max_{I \in \mathcal{I}} \#I$$

Allora

$$n_s = n_i$$

DIMOSTRAZIONE.

Cominciamo con il dimostrare che per ogni $S \in \mathcal{S}$ e per ogni $I \in \mathcal{I}$ si ha $\#I \leq \#S$

Se $\#I \leq \#S$ esisterebbero $S \in \mathcal{S}$ e $I \in \mathcal{I}$ con $S = \{v_1, v_2, \ldots, v_n\} \in \mathcal{S}$ e $I = \{w_1, w_2, \ldots, w_{n+1}\} \in \mathcal{I}$.

Osserviamo innanzi tutto che, dal momento che $I \in \mathcal{I}$, $w_k \neq 0 \ \forall k$.

Se esistesse k tale che $w_k = 0$, preso $w_h \in \mathcal{I}$ avremmo che $0 \cdot w_h + 1 * w_k = 0$ e quindi esisterebbe una combinazione lineare nulla a coefficienti non tutti nulli di elementi di I

Consideriamo l'insieme di vettori

$$S_1 = \{w_1, v_2, \ldots, v_n\}$$

Poichè S è un insieme di generatori possiamo esprimere w_1 come combinazione lineare di elementi di S, cioè esistono $c_i \in \mathbb{R}$ tali che $w_1 = c_1 v_1 + c_2 v_2 + \cdots, c_n v_n$; inoltre, poichè $w_1 \neq 0$, i coefficienti d c_i non sono tutti nulli .

Possiamo supporre che c_1 sia non nullo e possiamo quindi esprimere v_1 come combinazione lineare di $\{w_1, v_2, \cdots, v_n\}$. Pertanto S_1 è un insieme di generatori .

$v_1 = \frac{1}{c_1}(w_1 - c_2 v_2 - \cdots - v_n)$

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \alpha_1 \frac{1}{c_1}(w_1 - c_2 v_2 - \cdots - v_n) + \alpha_2 v_2 + \cdots + \alpha_n v_n$

Consideriamo ora

$$S_2 = \{w_1, w_2, \ldots, v_n\}$$

poichè possiamo esprimere w_2 come combinazione lineare a coefficienti non tutti nulli di elementi di S_1, esistono $c_i \in \mathbb{R}$ non tutti nulli tali che

$$w_2 = c_1 w_1 + c_2 v_2 + \cdots + c_n v_n$$

Dal momento che w_1 e w_2 sono linearmente indipendenti esiste un coefficiente $c_i \neq 0$, $c_i \neq c_1$. Sia ad esempio c_2 tale coefficiente. Ne segue che v_2 si può esprimere come combinazione lineare di $\{w_1, w_2, \cdots, v_n\}$ e pertanto,come prima si può affermare che S_2 è un insieme di generatori.

Se $c_i = 0$ per tutti gli indici $i = 2, \ldots n$ si avrebbe $w_1 = c_1 w_2$ il che è in contrasto con l'ipotesi che w_1 e w_2 siano linearmente indipendenti.

Proseguendo per n passi, possiamo affermare che

$$S_n = \{w_1, w_2, \ldots, w_n\}$$

è un insieme di generatori e, quindi possiamo affermare che w_{n+1} si può esprimere come combinazione lineare di $\{w_1, w_2, \cdots, w_n\}$ il che è in contrasto col fatto che i vettori di I sono linearmente indipendenti.

Possiamo pertanto affermare che $n_i \leq n_s$.

Viceversa Sia $S \in \mathcal{S}$, $S = \{v_1, v_2, \ldots, v_{n_s}\} \subset \mathcal{V}$ tale che $\#S = n_s$; allora $S \in \mathcal{I}$ in quanto se così non fosse esisterebbero $\alpha_1, \alpha_2, \ldots, \alpha_{n_s}$ non tutti nulli tali che

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n_s} v_{n_s} = 0$$

Supponendo che sia, ad esempio $\alpha_1 \neq 0$ potremmo quindi ricavare

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \cdots - \frac{\alpha_{n_s}}{\alpha_1} v_{n_s}$$

e quindi sarebbe possibile generare $\mathcal{V}$ con $n_s - 1$ vettori, in contrasto con la definizione di n_s Ne segue che $n_s \leq n_i$.

□

Teorema 3.4 *Sia B una base di $\mathcal{V}$, allora B contiene $n_s = n_i = n$ elementi.*

In altre parole B contiene il massimo numero di vettori linearmente indipendenti di $\mathcal{V}$ ed anche il numero minimo di vettori necessario a generare $\mathcal{V}$.

Si definisce pertanto n dimensione dello spazio $\mathcal{V}$

DIMOSTRAZIONE. Dal momento che B è una base di $\mathcal{V}$ si ha che $B \in \mathcal{S}$ e $B \in \mathcal{I}$ per cui se $n = \#B$ si ha

$$n \leq n_i = n_s \leq n$$

□

Corollario 3.1 *Sia $\mathcal{V}$ uno spazio vettoriale non banale di dimensione finita e sia B un sottoinsieme finito di $\mathcal{V}$ che genera $\mathcal{V}$; allora esiste una base di $\mathcal{V}$ i cui elementi appartengono a B.*

DIMOSTRAZIONE. Se $\#B$ è uguale alla dimensione dello spazio $\mathcal{V}$, B è una base. In caso contrario i vettori di B non sono linearmente indipendenti. Pertanto esiste una combinazione lineare nulla a coefficienti non tutti nulli

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{n_s} v_{n_s} = 0$$

Supponendo che sia, ad esempio $\alpha_1 \neq 0$ potremmo quindi ricavare

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \cdots - \frac{\alpha_{n_s}}{\alpha_1} v_{n_s}$$

e $B \setminus \{v_1\}$ è ancora un insieme di generatori di $\mathcal{V}$.

Procedendo in quasto modo possiamo ridurre il numero di elementi di B alla dimensione di $\mathcal{V}$ ed ottenere una base. □

Corollario 3.2 *Sia $\mathcal{V}$ uno spazio vettoriale non banale di dimensione finita e sia B un sottoinsieme finito di $\mathcal{V}$ tale che $\mathbb{S}(B)$ non è tutto $\mathcal{V}$; allora esiste una base di $\mathcal{V}$ che contiene B.*

DIMOSTRAZIONE. Supponiamo che B contenga $m < n$ ove n è la dimensione di $\mathcal{V}$. Evidentemente esiste $v \in \mathcal{V} \setminus \mathbb{S}(B)$ e risulta che v è linearmente indipendente dagli elementi di B. Pertanto $B \cup \{u\}$ contiene $m + 1$ vettori linearmente indipendenti. Se $m + 1 = n$ possiamo concludere, in caso contrario procediamo come prima per un numero finito di passi. □

Definizione 3.11 *Siano $\mathcal{V}$ e $\mathcal{U}$ spazi vettoriali sul campo $\mathcal{F}$. Diciamo che la funzione $L : \mathcal{V} \to \mathcal{U}$, $\mathcal{V} \ni x \mapsto L(x) \in \mathcal{U}$ è lineare (è un omomorfismo) se*

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y) \qquad \forall x, y \in \mathcal{V} \ , \ \forall \alpha, \beta \in \mathcal{F}$$

Diciamo che L è iniettiva se $L(x) = L(y)$ implica $x = y$
Diciamo che L è surgettiva se $\forall y \in \mathcal{U}$ esiste $x \in \mathcal{V}$ tale che $y = L(x)$.
Diciamo che L è un isomorfismo se è iniettiva e surgettiva.
Chiamiamo rango di L, o immagine di $\mathcal{V}$ secondo L, l'insieme

$$R(L) = Im(L) = \{y \in \mathcal{U} \ : \ \exists x \in \mathcal{V} \ : \ L(x) = y\} = L(\mathcal{V})$$

Chiamiamo nucleo (kernel) di L l'insieme

$$\ker(L) = \{x \in \mathcal{V} \ : \ L(x) = 0\}$$

Diciamo che $\mathcal{V}$ e $\mathcal{U}$ sono isomorfi se esiste un isomorfismo L da $\mathcal{V}$ in $\mathcal{U}$

Chiaramente L è surgettiva se e solo se

$$L(\mathcal{V}) = \mathcal{U}$$

e si vede anche subito che, vista la linearità, L è iniettiva se e solo se $\ker(L) = \{0\}$ Infatti

$$L(x) = L(y) \iff L(x - y) = 0$$

e, ovviamente,

$$x = y \iff x - y = 0$$

Definizione 3.12 *Siano $\mathcal{V}$ e $\mathcal{U}$ spazi vettoriali sul campo $\mathcal{F}$. Diciamo che la funzione $L : \mathcal{V} \to \mathcal{U}$ è invertibile se esiste $\Lambda : \mathcal{U} \to \mathcal{V}$ tale che*

$$\Lambda(L(v)) = v \qquad \forall v \in \mathcal{V}$$

$$L(\Lambda(u)) = u \qquad \forall u \in \mathcal{U}$$

In tal caso chiamiamo

$$\Lambda = L^{-1}$$

Si vede subito che

Teorema 3.5 *Siano $\mathcal{V}$ e $\mathcal{U}$ spazi vettoriali sul campo $\mathcal{F}$ e sia $L : \mathcal{V} \to \mathcal{U}$; L è invertibile se e solo se L è iniettiva e surgettiva.*

DIMOSTRAZIONE. Se L è invertibile allora $\forall u \in \mathcal{U}$ si ha $L^{-1}(u) \in \mathcal{V}$ e $u = L(L^{-1}(u))$ per cui L è surgettiva.

Inoltre se $L(x) = L(y)$, per $x.y \in \mathcal{V}$, avremo

$$x = L^{-1}(L(x)) = L^{-1}(L(y)) = y$$

e quindi L è iniettiva.

Viceversa sia $u \in \mathcal{U}$ poichè L è surgettiva esiste un elemento $v \in \mathcal{V}$ tale che $u = L(v)$ e dal momento che L è iniettiva tale elemento è unico.

Se definiamo $v = L^{-1}(u)$ avremo subito che

$$v = L^{-1}(L(v)) \ \forall v \in \mathcal{V} \qquad \text{e anche} \qquad u = L(L^{-1}(u)) \ \forall u \in \mathcal{U}$$

□

Del teorema precedente è opportuno sottolineare che l'inversa di una funzione da $\mathcal{V}$ in $\mathcal{U}$ è caratterizzata dalla seguente equivalenza:

$$v = L^{-1}(u) \iff u = L(v)$$

che permette di determinare, talora esplicitamente, l'inversa di L.

Teorema 3.6 *Siano $\mathcal{V}$ e $\mathcal{U}$ spazi vettoriali sul campo $\mathcal{F}$ e sia $L : \mathcal{V} \to \mathcal{U}$; L invertibile e lineare. Allora anche L^{-1} è lineare.*

DIMOSTRAZIONE. Occorre mostrare che

$$L^{-1}(\alpha x + \beta y) = \alpha L^{-1}(x) + \beta L^{-1}(y) \qquad \forall x, y \in \mathcal{U} \ , \ \forall \alpha, \beta \in \mathcal{F}$$

Siano $u, v \in \mathcal{V}$ tali che

$$x = L(u) \qquad y = L(v)$$

o, equivalentemente,

$$L^{-1}(x) = u \qquad L^{-1}(y) = v$$

Avremo

$$\alpha x + \beta y = L(\alpha u + \beta v) \text{ e quindi } \alpha u + \beta v = L^{-1}(\alpha x + \beta y)$$

Ne segue

$$L^{-1}(\alpha x + \beta y) = \alpha u + \beta v = \alpha L^{-1}(x) + \beta L^{-1}(y)$$

□

Teorema 3.7 *Siano $\mathcal{V}$ e $\mathcal{U}$ due spazi vettoriali isomorfi, sia $B = \{v_1, v_2, \ldots, v_n\}$ una base di $\mathcal{V}$ e sia $L : \mathcal{V} \to \mathcal{U}$ un isomorfismo da $\mathcal{V}$ in $\mathcal{U}$, allora $L(B)$ è una base di $\mathcal{U}$.*

DIMOSTRAZIONE. Vediamo innanzi tutto che $L(B)$ genera $\mathcal{U}$. Sia $u \in \mathcal{U}$, allora esiste $v \in \mathcal{V}$ tale che $u = L(v)$ e, dal momento che B è una base per $\mathcal{V}$, possiamo trovare $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathcal{F}$ tali che

$$v = \sum_1^n \alpha_i v_i$$

per cui

$$u = L(v) = L\left(\sum_1^n \alpha_i v_i\right) = \sum_1^n \alpha_i L(v_i)$$

D'altro canto se

$$0 = \sum_1^n \alpha_i L(v_i) = L\left(\sum_1^n \alpha_i v_i\right)$$

si ha

$$\sum_1^n \alpha_i v_i = 0$$

e $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Pertanto i vettori di $L(B)$ sono linearmente indipendenti. □

Ovviamente se L è un isomorfismo da $\mathcal{V}$ in $\mathcal{U}$ allora L^{-1} è un isomorfismo da $\mathcal{U}$ in $\mathcal{V}$ e quindi la controimmagine di una base di $\mathcal{U}$ è una base di $\mathcal{V}$.

Corollario 3.3 *Due spazi isomorfi hanno la stessa dimensione.*

È inoltre importante osservare che per definire un'applicazione lineare su uno spazio vettoriale $\mathcal{V}$ è sufficiente assegnare il trasformato $L(v_i)$ degli elementi v_i di una base di $\mathcal{V}$, infatti, se $v \in \mathcal{V}$,

$$L(v) = L\left(\sum_0^n \alpha_i v_i\right) = \sum_0^n \alpha_i L(v_i)$$

Definizione 3.13 *Siano $\mathcal{F}$ e $\mathcal{G}$ campi di scalari, e sia*

$$\mathcal{F}^n = \{(x_1, x_2, \ldots, x_n) \ : \ x_i \in \mathcal{F}\}$$

l'insieme di tutte le $n-$ple ordinate di elementi di $\mathcal{F}$ e supponiamo che sia definito il prodotto di uno scalare di $\mathcal{F}$ e di uno scalare di $\mathcal{G}$

Se $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathcal{F}^n$, definiamo la somma mediante la

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

ed il prodotto per uno scalare $\alpha \in \mathcal{G}$ mediante la

$$\alpha x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$$

In particolare può essere $\mathcal{F} = \mathcal{G}$ e si verifica subito che

- $\mathcal{F}^n$ è uno spazio vettoriale (su $\mathcal{F}$ o su $\mathcal{G}$)
- posto

$$\begin{cases} e_1 = (1, 0, \ldots, 0) \\ e_2 = (0, 1, \ldots, 0) \\ \ldots \\ e_n = (0, 0, \ldots, 1) \end{cases}$$

si ha che $E = \{e_1, e_2, \ldots, e_n\}$ è una base per $\mathcal{F}^n$

Possiamo dimostrare un risultato di notevole importanza.

Teorema 3.8 *Ogni spazio vettoriale di dimensione n, finita, è isomorfo allo spazio $\mathcal{F}^n$.*

DIMOSTRAZIONE. Sia $B = \{v_1, v_2, \ldots, v_n\}$ una base dello spazio vettoriale $\mathcal{V}$ possiamo definire un'applicazione lineare

$$L : \mathcal{F}^n \to \mathcal{V}$$

mediante la

$$L(e_1) = v_1, L(e_2) = v_2, \ldots, L(e_n) = v_n$$

L è surgettiva in quanto se $v \in \mathcal{V}$ si ha

$$v = \sum_{k=1}^{n} \alpha_k v_k = \sum_{k=1}^{n} \alpha_k L(e_k) = L\left(\sum_{k=1}^{n} \alpha_k e_k\right)$$

e $\sum_{k=1}^{n} \alpha_k e_k \in \mathcal{F}^n$.

L è iniettiva in quanto se $L(x) = 0$ si ha

$$0 = L(x) = L\left(\sum_{k=1}^{n} \alpha_k e_k\right) = \sum_{k=1}^{n} \alpha_k L(e_k) = \sum_{k=1}^{n} \alpha_k v_k$$

e poichè B è una base di $\mathcal{V}$ possiamo affermare che $\alpha_k = 0$ per ogni k e quindi $x = 0$

Ne segue che L è un isomorfismo.

□

Se $x \in \mathcal{F}^n$ si ha

$$x = x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + x_n(0, 0, \ldots, 1)$$

per cui, relativamente alla base E, $x_1, x_2, \ldots, x_n$ sono i coefficienti della combinazione lineare che definisce x.

Chiameremo tali coefficienti coordinate di x.

Normalmente lo spazio $\mathcal{F}$ degli scalari è $\mathbb{R}$ o $\mathbb{C}$, piu' frequentemente $\mathbb{R}$, pertanto considereremo d'ora innanzi $\mathcal{F} = \mathbb{R}$ e $\mathcal{F}^n = \mathbb{R}^n$ essendo la perdita di generalità piccola in quanto. come abbiamo appena visto, ogni spazio vettoriale di dimensione n è isomorfo ad $\mathbb{R}^n$.

Infatti, se due spazi vettoriali sono tra loro isomorfi, hanno la stessa struttura vettoriale e quel che si verifica in uno, riguardo alla struttura lineare, può essere verificato anche nell'altro per mezzo dell'isomorfismo.

Definizione 3.14 *Si definisce norma in $\mathbb{R}^n$ una funzione che si indica con*

$$\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$$

che verifica le seguenti proprietà:

1. $\|x\| \geq 0 \qquad \forall x \in \mathbb{R}^n$

2. $\|x\| = 0 \Leftrightarrow x = 0$

3. $\|\alpha x\| = |\alpha|\|x\| \qquad \forall \alpha \in \mathbb{R}\ ,\ \forall x \in \mathbb{R}^n$

4. $\|x + y\| \leq \|x\| + \|y\| \qquad \forall x, y \in \mathbb{R}^n.$

Definizione 3.15 *Si definisce prodotto scalare in* $\mathbb{R}^n$ *una funzione*

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

tale che

1. $\langle x, x \rangle \geq 0 \qquad \forall x \in \mathbb{R}^n$

2. $\langle x, y \rangle = \langle y, x \rangle \qquad \forall x, y \in \mathbb{R}^n$

3. $\langle x, x \rangle = 0 \quad \Leftrightarrow \quad x = 0$

4. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle \qquad \forall x, y, z \in R^n\ ,\ \forall \alpha, \beta \in \mathbb{R}.$

Avendo definito in $\mathbb{R}^n$ un prodotto scalare, possiamo definire

$$|x|^2 = \langle x, x \rangle$$

Si verifica facilmente che $|\cdot|$ soddisfa le condizioni 1), 2) e 3) della definizione di norma ed inoltre possiamo verificare che anche 4) è vera utilizzando il seguente lemma.

Lemma 3.1 *Siano* $x, y \in \mathbb{R}^n$*, allora*

- -**Disuguaglianza di Schwarz**-

$$\langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\|\|y\|$$

- -**Disuguaglianza triangolare**-

$$\|x + y\| \leq \|x| + \|y\|$$

DIMOSTRAZIONE. La disuguaglianza di Schwarz può essere riscritta come

$$|\langle x, y \rangle| \leq \|x\|\|y\|$$

per cui, osservando che, $\forall t \in \mathbb{R}$

$$0 \leq \|x + ty\|^2 = \langle x + ty, x + ty \rangle = t^2\|y\|^2 + 2t\langle x, y \rangle + \|x\|^2$$

possiamo affermare che $t^2\|y\|^2 + 2t\langle x, y \rangle + \|x\|^2$ è un polinomio di secondo grado in t sempre positivo e deve quindi avere discriminante negativo o nullo.

Se ne deduce

$$\langle x, y \rangle^2 - \|x\|^2\|y\|^2 \ \leq 0$$

e la disuguaglianza di Schwarz.

La disuguaglianza triangolare segue da

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x| + \|y\|)^2$$

da cui

$$\|x + y\| =\leq \|x| + \|y\|$$

□

Osserviamo che

$$|\langle x, y\rangle| = \|x\|\|y\|$$

se e solo se esiste $t \in \mathbb{R}$ tale che $x + ty = 0$, ovvero x e y sono uno multiplo dell'altro.

Si definisce usualmente prodotto scalare in $\mathbb{R}^n$ mediante la

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k y_k$$

e ne segue che la norma ad esso associata è

$$\|x\|^2 = \langle x, x\rangle = \sum_{sk=1}^{n} x_k^2$$

$\|\cdot\|$ è detta norma euclidea in quanto $\|x\|$ coincide con la distanza euclidea di x dall'origine.

Possiamo verificare, come nel caso del valore assoluto per i numeri reali, che si ha

$$|\|x\| - \|y\|| \leq \|x - y\|$$

$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$
$\|y\| = \|y - x + x\| \leq \|x - y\| + \|x\|$
perciò

$$-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\|$$

Si vede anche subito che

$$\langle P_1, P_2\rangle = \|P_1\|\|P_2\|\cos(\theta_2 - \theta_1)$$

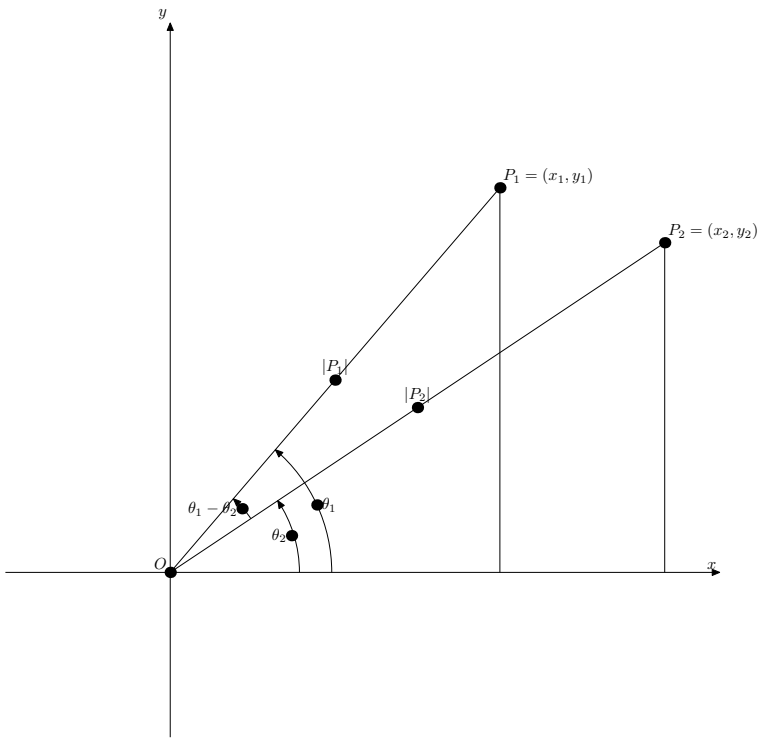
Infatti, facendo riferimento ad $\mathbb{R}^2$ e alla figura 3.1, si ha

$$\begin{aligned}\langle P_1, P_2\rangle &= \\ = x_1x_2 + y_1y_2 = \|P_1\|\|P_2\|\cos(\theta_2)\cos(\theta_1) + \|P_1\|\|P_2\|\sin(\theta_2)\sin(\theta_1) &= \\ = \|P_1\|\|P_2\|\cos(\theta_2 - \theta_1)\end{aligned}$$

L'osservazione appena fatta giustifica il fatto che

Diciamo che due vettori $x, y \in \mathbb{R}^n$ sono ortogonali se $\langle x, y\rangle = 0$.

Diciamo che sono paralleli se esiste $\lambda \in \mathbb{R}$ tale che $x = \lambda y$. Se x ed y sono paralleli $\langle x, y\rangle = \|x\|\|y\|$


y
$P_1 = (x_1, y_1)$
$P_2 = (x_2, y_2)$
$|P_1|$
$|P_2|$
$\theta_1 - \theta_2$
θ_1
θ_2
O
x

4. Applicazioni Lineari e matrici

Come già detto useremo nel seguito solo spazi vettoriali euclidei $\mathbb{R}^n$ essendo ogni spazio vettoriale di dimensione n ad essi isomorfo. Indicheremo con $E_n = \{e_i : i = 1, 2, \ldots, n\}$ la base canonica di $\mathbb{R}^n$ per modo che

$$x = \sum_{i=0}^{n} x_i e_i$$

e ci riferiremo ad x_i come alla componente $i-$esima di x.

Sia, ad esempio, $L : \mathbb{R}^2 \to \mathbb{R}^3$ un'applicazione lineare . Avremo che

$$y = L(x) = L\left(x_1 e_1 + x_2 e_2\right) = x_1 L(e_1) + x_2 L(e_2)$$

per cui si ha:

$$\begin{cases} y_1 = x_1 L(e_1)_1 + x_2 L(e_2)_1 \\ y_2 = x_1 L(e_1)_2 + x_2 L(e_2)_2 \\ y_2 = x_1 L(e_1)_3 + x_2 L(e_2)_3 \end{cases} \tag{4.1}$$

essendo $L(e_i)_j \in \mathbb{R}$ la $j-$esima componente del vettore $L(e_i)$ immagine di e_i secondo L.

L è univocamente determinata da $L(e_i)_j$ al variare di $i = 1, 2$ e $j = 1, 2, 3$ e possiamo disporre tali valori secondo una tabella con 3 righe e 2 colonne, come segue:

$$A = \begin{pmatrix} L(e_1)_1 & L(e_2)_1 \\ L(e_1)_2 & L(e_2)_2 \\ L(e_1)_3 & L(e_2)_3 \end{pmatrix} = \begin{pmatrix} L(e_1) & L(e_2) \end{pmatrix}$$

D'altro canto anche x ed y possono essere scritti come tabelle aventi una sola colonna e 2 o 3 righe

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

che chiamiamo vettori colonna oppure, scambiando righe con colonne con un'operazione che chiamiamo trasposizione ed indichiamo con ^{T}, come

$$x^T = (x_1, x_2) \qquad , \qquad y^T = (y_1, y_2, y_3)$$

che chiamiamo vettori riga.

Possiamo allora scrivere la 4.1 come

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} L(e_1)_1 & L(e_2)_1 \\ L(e_1)_2 & L(e_2)_2 \\ L(e_1)_3 & L(e_2)_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax \tag{4.2}$$

ove intendiamo l'uguaglianza in $\mathbb{R}^3$, quindi componente per componente, e il prodotto fatto righe per colonne, cioè definendo ogni riga della matrice prodotto come il risultato del prodotto scalare tra la corrispondente riga della matrice A ed il vettore x .

Le righe della matrice A sono vettori della stessa dimensione di x

Pertanto

$$Ax = \begin{pmatrix} L(e_1) & L(e_2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax = \begin{pmatrix} \langle (L(e_1)_1, L(e_2)_1), x \rangle \\ \langle (L(e_1)_2, L(e_2)_2)), x \rangle \\ \langle (L(e_1)_3, L(e_2)_3), x \rangle \end{pmatrix}$$

ove si intenda che la

$$\begin{pmatrix} L(e_1) & L(e_2) \end{pmatrix}$$

è una tabella di 3 righe con 2 colonne costituite da $L(e_1)$ e $L(e_2)$.

Quanto sopra esposto mostra come possa essere utile un formalismo che consenta di trattare tabelle come A con lo scopo di descrivere applicazioni lineari tra spazi euclidei.

Inoltre lo stesso formalismo tornerà comodo nello studio dei sistemi di equazioni algebriche lineari come ad esempio:

$$\begin{cases} a_{11}x_1 + a_{21}x_2 = b_1 \\ a_{12}x_1 + a_{22}x_2 = b_2 \\ a_{13}x_1 + a_{23}x_2 = b_3 \end{cases} \tag{4.3}$$

che chiameremo sistema di equazioni lineare ed assoceremo alla tabella

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

scrivendo, in forma vettoriale,

$$Ax = b$$

D'altro canto data la tabella A possiamo definire un'applicazione lineare $L : \mathbb{R}^2 \to \mathbb{R}^3$ mediante la

$$L(x) = Ax$$

e si vede quindi che risolvere il sistema di equazioni lineari

$$Ax = b$$

significa trovare i vettori x, in questo caso di $\mathbb{R}^2$, la cui immagine secondo L è b.

Nel caso in cui $b = 0$ il sistema si dice omogeneo.

Usando le notazioni prima introdotte, risolvere un sistema lineare omogeneo significa trovare

$$\ker(L) = \{x \in \mathbb{R}^2 \ : \ L(x) = 0\} = \{x \in \mathbb{R}^2 \ : \ Ax = 0\}$$

4.1 Matrici

Diciamo che è data una matrice A di ordine $m \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{m,n} \end{pmatrix} = \{a_{ij}\}$$

se è data una tabella avente m righe ed n colonne i cui elementi a_{ij} si identificano mediante due indici i, j di cui i indica la riga e j la colonna.

Chiamiamo $\mathcal{M}^{m\times n}$ l'insieme delle matrici $m \times n$ e definiamo la somma di due matrici $A, B \in \mathcal{M}^{m\times n}$ mediante la

$$A + B = \{a_{ij}\} + \{b_{ij}\} = \{a_{ij} + b_{ij}\}$$

ed il prodotto di una matrice A per uno scalare α , in $\mathbb{R}$ o in $\mathbb{C}$, mediante la

$$\alpha A = \alpha\{a_{ij}\} = \{\alpha a_{ij}\}$$

Si verifica che, con queste operazioni, $\mathcal{M}^{m\times n}$ è uno spazio vettoriale di dimensione $p = nm$ e che una base è data dall'insieme delle matrici che hanno tutti gli elementi nulli tranne uno che vale 1; in altre parole una base per lo spazio $\mathcal{M}^{m\times n}$ è costituita da tutte le matrici $M_{h,k}$ tali che $a_{ij} = 0$ se $(i, j) \neq (h, k)$ e $a_{hk} = 1$;

$$M_{hk} = \begin{array}{c} \\ h \end{array} \begin{array}{c} \begin{array}{ccccc} & & k & & \end{array} \\ \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & 1 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \end{array}$$

Si vede anche che lo spazio $\mathcal{M}^{m\times n}$ è isomorfo ad uno spazio euclideo $\mathbb{R}^p$ con $p = nm$, essendo l'isomorfismo definito dalla corrispondenza della matrice $M_{h,k} \in \mathcal{M}_{m\times n}$ con il vettore $e_{n(h-1)+k}$ della base canonica di $\mathbb{R}^p$.

In $\mathcal{M}^{m\times n}$ si definisce anche un'operazione di prodotto;

se $A, B \in \mathcal{M}^{m\times n}$ definiamo

$$AB = C$$

dove $\forall i = 1, \dots m$, $\forall j = 1, \dots n$ si pone

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Tale prodotto si indica come prodotto righe per colonne di due matrici e può essere calcolato a condizione che le colonne della prima matrice siano in numero uguale alle righe della seconda, inoltre il prodotto di due matrici non è commutativo neppure nel caso in cui le matrici abbiano ciascuna lo stesso numero di righe e colonne e cioè siano matrici quadrate.

Lo spazio $\mathcal{M}^n$ delle matrici quadrate di ordine n, con le operazioni appena definite, ha una struttura di algebra non commutativa.

Vale la pena ricordare che, come abbiamo già osservato, una matrice quadrata di ordine n definisce un'applicazione lineare di $\mathbb{R}^n$ in se stesso e che ogni applicazione lineare di $\mathbb{R}^n$ in se stesso individua una matrice che la caratterizza completamente.

Il prodotto righe per colonne di due matrici trova utile applicazione nella rappresentazione della composizione di due applicazioni lineari. Infatti se, ad esempio,

$$L_1 : \mathbb{R}^2 \to \mathbb{R}^3 \qquad \text{e} \qquad L_2 : \mathbb{R}^3 \to \mathbb{R}$$

essendo

$$\mathbb{R}^2 \ni x \mapsto L_1(x) = Ax \in \mathbb{R}^3 \qquad \text{e} \qquad \mathbb{R}^3 \ni y \mapsto L_2(y) = By \in \mathbb{R}$$

avremo che $B \in \mathcal{M}^{3\times 2}$, $A \in \mathcal{M}^{1\times 3}$ mentre $C = AB \in \mathcal{M}^{1\times 2}$ rappresenta l'applicazione lineare $L_2(L_1(\cdot)) : \mathbb{R}^3 \to \mathbb{R}$. infatti

$$\begin{aligned} L_2(L_1(x)) &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} L_1(x) = \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{3,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \\ &= \begin{pmatrix} \sum_{k=1}^{3} a_{1k} b_{k1} & \sum_{k=1}^{3} a_{1k} b_{k2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

Definiamo matrice identica in $\mathcal{M}^n$, e la indichiamo con I_n, la matrice tale che

$$a_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Si ottiene

$$AI = A = IA$$

il che giustifica il nome.

4.2 Spazio delle Applicazioni Lineari

Definizione 4.1 *Definiamo $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ come l'insieme delle applicazioni lineari su $\mathbb{R}^n$ a valori in $\mathbb{R}^m$.*

Se $n = m$ scriviamo semplicemente $\mathcal{L}^n$ per indicare lo spazio delle applicazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}^n$

Gli elementi di $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ possono essere messi in corrispondenza biunivoca con le matrici aventi m righe ed n colonne, $\mathcal{M}^{m\times n}$.

Più precisamente

Teorema 4.1 *$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ed $\mathcal{M}^{m\times n}$ sono isomorfi*

DIMOSTRAZIONE. Infatti possiamo definire $\Lambda : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathcal{M}^{m\times n}$ mediante la

$$\Lambda(L) = A = \begin{pmatrix} L(e_1) & L(e_2) & \cdots & L(e_n) \end{pmatrix}$$

ed anche $\Phi : \mathcal{M}^{m\times n} \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ dove $L = \Phi(A)$ è definita dall'uguaglianza

$$L(x) = \Phi(A)(x) = Ax$$

Si verifica che

$$\Lambda(\Phi(A)) = A \quad \text{e} \quad \Phi(\Lambda(L)) = L$$

in quanto si ha $\Phi(A) = L$ se

$$L(x) = Ax$$

per cui

$$\Lambda(\Phi(A)) = \Lambda(L) = \begin{pmatrix} L(e_1) & L(e_2) & \cdots & L(e_n) \end{pmatrix} = \begin{pmatrix} Ae_1 & Ae_2 & \cdots & Ae_n \end{pmatrix} = AI = A$$

viceversa

$$\Phi(\Lambda(L))(x) = \Phi\left(\begin{pmatrix} L(e_1) & L(e_2) & \cdots & L(e_n) \end{pmatrix}\right) x =$$
$$= \begin{pmatrix} L(e_1) & L(e_2) & \cdots & L(e_n) \end{pmatrix} x = \sum_1^n x_k L(e_k) = L\left(\sum_1^n x_k e_k\right) = L(x)$$

Avremo quindi che Λ è invertibile e Φ è la sua inversa e che Λ (e $\Phi = \Lambda^{-1}$) è sia iniettiva che surgettiva;

Poichè

$$\Lambda(\alpha L_1 + \beta L_2) =$$
$$= \begin{pmatrix} \alpha L_1(e_1) + \beta L_2(e_1) & \alpha L_1(e_2) + \beta L_2(e_n) & \cdots & \alpha L_1(e_n) + \beta L_2(e_n) \end{pmatrix} =$$
$$= \alpha \begin{pmatrix} L_1(e_1) & L_1(e_2) & \cdots & L_1(e_n) \end{pmatrix} + \beta \begin{pmatrix} L_2(e_1) & L_2(e_2) & \cdots & L_2(e_n) \end{pmatrix} =$$
$$= \alpha\Lambda(L_1) + \beta\Lambda(L_2)$$

Sia ad esempio $n = 2$ ed $m = 3$ ed $L : \mathbb{R}^2 \to R^3$ definita da

$$L(x, y) = (a_1x + b_1y, a_2x + b_2y, a_3x + b_3y)$$

Allora

$$L(e_1) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \; L(e_2) = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$L(x, y) = xL(e_1) + yL(e_2) = x \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + y \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} =$$
$$= \begin{pmatrix} a_1x + b_1y \\ a_2x + b_2y \\ a_3x + b_3y \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$
$$= \begin{pmatrix} L(e_1) & L(e_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Pertanto L definisce in maniera univoca la matrice

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} = \begin{pmatrix} L(e_1) & L(e_2) \end{pmatrix}$$

e

$$\Lambda(L) = A$$

Viceversa $A \in \mathcal{M}^{3\times 2}$,

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$$

definisce, in modo univoco, una applicazione lineare $L : \mathbb{R}^2 \to R^3$ mediante

$$L(x, y) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1x + b_1y \\ a_2x + b_2y \\ a_3x + b_3y \end{pmatrix}$$

e

$$\Phi(A) = L$$

Si ha quindi che

$$\Lambda(\Phi(A)) = A \quad \text{e} \quad \Phi(\Lambda(L)) = L$$

Λ è una applicazione lineare e quindi $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ed $\mathcal{M}^{m\times n}$ sono isomorfi □

Possiamo pertanto identificare L con la matrice A per la quale risulta

$$L(x) = Ax$$

Chiamiamo A matrice di rappresentazione di L. È evidente, per come abbiamo visto essere costituita, che la matrice A è identificata univocamente dalle basi scelte in $\mathbb{R}^n$ ed $\mathbb{R}^m$.

Per estensione, che se L è l'applicazione lineare associata alla matrice A, chiamiamo rango (in Inglese range) di A,

$$R(A) = \{y \in \mathbb{R}^m \ : \ \exists x \in \mathbb{R}^n \ , \ y = Ax\} = R(L)$$

inoltre definiamo nucleo di A,

$$\ker(A) = \{x \in \mathbb{R}^n \ \ : \ \ Ax = 0\} = \ker(L)$$

Osserviamo esplicitamente che il termine rango (=range) di A è anche usato, per indicare la dimensione dello spazio vettoriale Immagine dell'applicazione lineare L identificata da A. In questo caso però il termine inglese corrispondente è rank mentre in italiano spesso si usa in questo caso il termine caratteristica di una matrice.

Viste le definizioni, si ha che

- Il sistema lineare
$$L(x) = Ax = b$$
ha soluzioni per ogni $b \in \mathbb{R}^m$ se e solo se $R(L) = \mathbb{R}^m$; indichiamo questo fatto dicendo che *L (oppure A) ha dimensione massima (in inglese "full range")*.

- Per ogni $b \in \mathbb{R}^m$, il sistema lineare
$$L(x) = Ax = b$$
ha soluzione unica se e solo se $\ker(L) = \{0\}$; indichiamo questo fatto dicendo che *L (oppure A) ha nucleo nullo*.

- Sia
$$\mathcal{S} = \{x \in \mathbb{R}^n \ : \ Ax = b\}$$
lo spazio delle soluzioni del sistema $Ax = b$; sia inoltre $\bar{x} \in \mathbb{R}^n \in \mathcal{S}$ sia cioè $A\bar{x} = b$. Allora
$$\mathcal{S} = \bar{x} + \ker(A)$$
Infatti se $x \in \mathcal{S}$ allora $Ax = A\bar{x} = b$ e quindo $A(\bar{x} - x) = 0 \in \ker(A)$.

Possiamo provare che

Teorema 4.2 *Sia $L : \mathbb{R}^n \to \mathbb{R}^m$ una applicazione lineare. allora*

$$n = \dim \mathbb{R}^n = \dim(\ker(L)) + \dim(R(L))$$

DIMOSTRAZIONE. Dal momento che $\ker(L)$ è uno sottospazio vettoriale di $\mathbb{R}^n$ possiamo trovare una base $\{u_1, u_2, \ldots, u_k\}$ di $\ker(L)$ e ovviamente risulterà $\dim(\ker(L)) = k \leq n$.

Possiamo quindi trovare $n-k$ vettori $\{v_1, v_2, \ldots, v_{n-k}\}$ in modo che $B = \{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_{n-k}\}$ sia una base di $\mathbb{R}^n$.

Per dimostrare il teorema è sufficiente provare che $A = \{L(v_1), L(v_2), \ldots, L(v_{n-k})\}$ è una base per $R(L)$.

Se infatti $w \in R(L)$ si ha $w = L(v)$ per qualche $v \in \mathbb{R}^n$ e quindi

$$w = L(v) = L\left(\sum_1^k \alpha_i u_i + \sum_1^{n-k} \beta_i v_i\right) = L\left(\sum_1^{n-k} \beta_i v_i\right)$$

in quanto $L(u_i) = 0$ dal momento che $u_i \in \ker(L)$.

Ciò basta per affermare che A genera $R(L)$.

Inoltre se avessimo

$$0 = \sum_1^{n-k} \beta_i L(v_i) = L\left(\sum_1^{n-k} \beta_i v_i\right)$$

ne seguirebbe che

$$\sum_1^{n-k} \beta_i v_i \in \ker(L)$$

e quindi

$$\sum_1^{n-k} \beta_i v_i = \sum_1^k \alpha_i u_i$$

Avremmo quindi trovato una combinazione lineare nulla di elementi di B e dal momento che B è una base di $\mathbb{R}^n$ ne seguirebbe che $\alpha_i = \beta_i = 0$ per cui i vettori $L(v_1), L(v_2), \ldots, L(v_{n-k})$ sono linearmente indipendenti. □

Corollario 4.1 *Sia $L : \mathbb{R}^n \to \mathbb{R}^m$ una applicazione lineare. Si ha che*

- *se $n > m$ allora L non è iniettiva*

  Infatti da $n - \dim(\ker(L)) = \dim(R(L)) \leq m < n$ *segue* $\dim(\ker(L)) > 0$

- *se $n < m$ allora L non è surgettiva*

  Infatti da $\dim(R(L))) + \dim(\ker(L) = n < m$ *segue* $\dim(\mathbb{R}(L)) < m$

Chiaramente quindi non possiamo parlare di invertibilità nel caso in cui $n \neq m$. e quindi, studiando l'invertibilità di L, supporremo

$$L : \mathbb{R}^n \to \mathbb{R}^n$$

Dal momento che ogni applicazione lineare è individuata univocamente da una matrice, possiamo sempre scrivere Ax in luogo di $L(x)$.

È opportuno comunque sottolineare che, così facendo, non omettiamo semplicemente l'uso delle parentesi, ma bensì utilizziamo il fatto che gli spazi $\mathcal{L}^n$ e $\mathcal{M}^n$ sono tra loro isomorfi.

Dal teorema 4.2 segue che:

Corollario 4.2 *Sia A una matrice in $\mathcal{M}^n$ e sia $L \in \mathcal{L}^n$ l'applicazione lineare ad essa associata.*

Il sistema lineare

$$L(x) = Ax = b$$

ha soluzioni per ogni $b \in \mathbb{R}^n$ se e solo se ammette un'unica soluzione.

Definizione 4.2 *Sia $A \in \mathcal{M}^n$*

Diciamo che A è invertibile a destra se esiste $B \in \mathcal{M}^n$ tale che

$$AB = I_n$$

Diciamo che A è invertibile a sinistra se esiste $C \in \mathcal{M}^n$ tale che

$$CA = I_n$$

Diciamo che A è invertibile se è invertibile sia a destra che a sinistra. In tal caso avremo che $B = C$ infatti

$$B = I_n B = (CA)B = C(AB) = CI_n = C$$

Definiamo $B = A^{-1}$ matrice inversa di A.

Ricordando che con $R(A)$ e $\ker(A)$ intendiamo il rango e il nucleo dell'applicazione lineare associata ad A, Possiamo provare che

Teorema 4.3 *Sia $A \in \mathcal{M}^n$, sono fatti equivalenti.*

- $R(A) = \mathbb{R}^n$
- $\ker(A) = \{0\}$
- *A ammette inversa a destra*

DIMOSTRAZIONE. L'equivalenza delle prime due affermazioni risulta dal teorema 4.2 Proviamo ora che la prima e la terza sono equivalenti.

Sia $x_i \in \mathbb{R}^n$ il vettore colonna tale che

$$Ax_i = e_i$$

e sia

$$B = (x_1\ x_2\ \cdots\ x_n)$$

la matrice le cui colonne sono i vettori x_i. Evidentemente

$$AB = A(x_1\ x_2\ \cdots\ x_n) = (e_1\ e_2\ \cdots\ e_n) = I_n$$

Se viceversa B è un'inversa a destra di A, per ogni $x \in \mathbb{R}^n$ possiamo considerare Bx e si ha

$$ABx = I_n x = x$$

per cui $x \in R(A)$ □

Corollario 4.3 *Se $A \in \mathcal{M}^n$ ammette inversa a destra B, allora B è anche inversa a sinistra per A.*

DIMOSTRAZIONE. Si ha che

$$ABA = I_n A = A$$

da cui

$$0 = ABA - I_n A = ABA - AI_n = A(BA - I_n)$$

Allora, dal momento che $\ker(A) = \{0\}$

$$BA - I_n = 0$$

□

Corollario 4.4 *Se $A \in \mathcal{M}^n$ ammette inversa a sinistra C allora* $\ker(A) = \{0\}$.

DIMOSTRAZIONE. Sia infatti $x \in \ker(A)$ allora $Ax = 0$ e quindi

$$x = CAx = C0 = 0$$

□

Ne consegue che se $A \in \mathcal{M}^n$ si ha:

$$\begin{array}{ccc} A \text{ ammette inversa a destra} & \Longleftrightarrow & R(A) = \mathbb{R}^n \\ \Downarrow & & \Updownarrow \\ A \text{ ammette inversa a sinistra} & \Longrightarrow & \ker(A) = \{0\} \end{array}$$

e quindi :

$$\begin{array}{ccc} A \text{ ammette inversa a destra} & \Longleftrightarrow & R(A) = \mathbb{R}^n \\ \Updownarrow & & \Updownarrow \\ A \text{ ammette inversa a sinistra} & \Longleftrightarrow & \ker(A) = \{0\} \end{array}$$

Inoltre se B, C sono inverse di A allora

$$C = CAB = B$$

e l'inversa è unica.

Chiamiamo matrice trasposta di $A = \{a_{ij}\} \in \mathcal{M}^n$ la matrice $A^T = \{a'_{ij}\}$ per la quale risulta

$$a'_{ij} = a_{ji}$$

.

Evidentemente

$$A = (A^T)^T$$

Diciamo che $A \in \mathcal{M}^n$ è simmetrica se

$$A = A^T$$

Definizione 4.3 *Sia $A = \{a_{ij}\} \in \mathcal{M}^{m\times n}$, e siano $C_1, C_2, \dots, C_n$ le sue colonne e $R_1, R_2, \dots, R_m$ le sue righe, per modo che*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} C_1 & C_2 & \cdots & C_n \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ \cdots \\ R_m \end{pmatrix}$$

Essendo

$$i = 1, 2, \dots n \ , \ j = 1, 2, \dots m$$

Indichiamo con $\mathcal{S}_c(A)$ e con $\mathcal{S}_r(A)$, rispettivamente, lo spazio vettoriale generato dai vettori $C_1, C_2, \dots, C_n$ e $R_1, R_2, \dots, R_m$.

Chiamiamo caratteristica (o rango) per colonne di A, ρ^c_A, *la dimensione di $\mathcal{S}_c(A)$ e* caratteristica (o rango) per righe di A, ρ^r_A , *la dimensione di $\mathcal{S}_r(A)$.*

In altre parole

$$\rho^c_A = \dim(\mathcal{S}_c(A)) \qquad , \qquad \rho^r_A = \dim(\mathcal{S}_r(A))$$

Teorema 4.4 *Sia $A \in \mathcal{M}^{m\times n}$, allora*

$$\dim(\mathcal{S}_c(A)) = \dim(\mathcal{S}_r(A))$$

e quindi il rango per colonne di A è uguale al rango per righe di A ed é minore sia di n che di m.

DIMOSTRAZIONE. Supponiamo ad esempio che $\dim(\mathcal{S}_c(A)) = r$ e siano $\gamma_1, \gamma_2, \dots, \gamma_r \in \mathbb{R}^m$ gli elementi di una sua base.

Allora $\forall j = 1, \dots, n$ esistono $\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jr} \in \mathbb{R}$ tali che

$$C_j = \sum_{k=1}^{r} \alpha_{kj}\gamma_k$$

Sia ad esempio $n = 2$, $m = 3$.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} C_1 & C_2 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} \qquad C_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$$

$$R_1 = \begin{pmatrix} a_{11} & a_{12} \end{pmatrix}$$
$$R_2 = \begin{pmatrix} a_{21} & a_{22} \end{pmatrix}$$
$$R_3 = \begin{pmatrix} a_{31} & a_{32} \end{pmatrix}$$

Se, ad esempio, $r = 2$ e

$$\gamma_1 = \begin{pmatrix} \gamma 11 \\ \gamma 21 \\ \gamma 31 \end{pmatrix} \qquad \gamma_2 = \begin{pmatrix} \gamma 12 \\ \gamma 22 \\ \gamma 32 \end{pmatrix}$$

$$C_1 = \alpha_{11}\gamma_1 + \alpha_{21}\gamma_2$$
$$C_2 = \alpha_{12}\gamma_1 + \alpha_{22}\gamma_2$$

$$C_1 = \begin{pmatrix} \alpha_{11}\gamma_{11} + \alpha_{21}\gamma_{12} \\ \alpha_{11}\gamma_{21} + \alpha_{21}\gamma_{22} \\ \alpha_{11}\gamma_{31} + \alpha_{21}\gamma_{32} \end{pmatrix}$$

$$C_2 = \begin{pmatrix} \alpha_{12}\gamma_{11} + \alpha_{22}\gamma_{12} \\ \alpha_{12}\gamma_{21} + \alpha_{22}\gamma_{22} \\ \alpha_{12}\gamma_{31} + \alpha_{22}\gamma_{32} \end{pmatrix}$$

ovvero

$$C_j = \begin{pmatrix} \alpha_{1j}\gamma_{11} + \alpha_{2j}\gamma_{12} \\ \alpha_{1j}\gamma_{21} + \alpha_{2j}\gamma_{22} \\ \alpha_{1j}\gamma_{31} + \alpha_{2j}\gamma_{33} \end{pmatrix}$$

e

$$a_{ij} = \alpha_{1j}\gamma_{i1} + \alpha_{2j}\gamma_{i2}$$

Se

$$\alpha_1 = (\alpha_{11}, \alpha_{12})$$
$$\alpha_2 = (\alpha_{21}, (\alpha_{22})$$

allora $\alpha_1, \alpha_2 \in \mathbb{R}^2$ e si ha

$$R_1 = \alpha_1\gamma_{11} + \alpha_2\gamma_{12}$$
$$R_2 = \alpha_1\gamma_{21} + \alpha_2\gamma_{22}$$
$$R_3 = \alpha_1\gamma_{31} + \alpha_2\gamma_{32}$$

cioè

$$R_i = \alpha_1\gamma_{i1} + \alpha_2\gamma_{i2}$$

e

$$a_{ij} = \sum_{k=1}^{r} \alpha_{kj}\gamma_{ik} \tag{4.4}$$

Consideriamo ora gli r vettori riga $\alpha_1 = (\alpha_{k1}, \alpha_{k2}, \ldots, (\alpha_{kn}) \in \mathbb{R}^n$ allora la 4.4 si riduce a

$$R_i = \sum_{k=1}^{r} \alpha_k \gamma_{ki}$$

Quindi ogni riga è generata da al piu' r vettori e quindi $\dim(\mathcal{S}_c(A)) \leq \dim(\mathcal{S}_c(A))$.

Scambiando il ruolo di righe e colonne si dimostra anche la disuguaglianza opposta e quindi l'uguaglianza.

□

Definizione 4.4 *Il teorema precedente ci consente di definire caratteristica (o rango) ρ_A di una matrice A, equivalentemente, come la dimensione dello spazio generato dalle sue righe o dalle sue colonne.*

In altre parole

$$\rho_A = \dim(\mathcal{S}_c(A)) = \dim(\mathcal{S}_r(A))$$

4.3 Sistemi di equazioni lineari

Cominciamo a studiare il problema di risolvere un sistema algebrico di equazioni lineari considerando il caso di un sistema di due equazioni in due incognite.

$$\begin{cases} ax + by = \alpha \\ cx + dy = \beta \end{cases} \tag{4.5}$$

Per a, b, c, d e $\alpha, \beta \in \mathbb{R}$ fissati, intendiamo porci il problema di trovare $x, y \in \mathbb{R}$ tali che sostituiti nelle due precedenti uguaglianze le rendano vere.

Più formalmente, assegnati a, b, c, d e $\alpha, \beta \in \mathbb{R}$, risolvere il sistema lineare 4.5 significa trovare $x, y \in \mathbb{R}$ tali che 4.5 sia verificato.

Se definiamo

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad , \quad B = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{e} \quad u = \begin{pmatrix} x \\ y \end{pmatrix}$$

possiamo usare le notazioni matriciali per scrivere il sistema 4.5 in forma compatta come

$$Au = B$$

Il modo più semplice per risolvere 4.5 prevede che:

- si ricavi x dalla prima equazione

- lo si sostituisca nella seconda ottenendo un'equazione che contiene solo l'incognita y
- da quest'ultima si ricavi y
- si ricavi x utilizzando l'espressione ottenuta al primo passo che fornisce x in funzione di y

Esplicitamente, dal sistema

$$\begin{cases} ax + by = \alpha \\ cx + dy = \beta \end{cases}$$

ricavando x dalla prima equazione e sostituendo nella seconda:

$$\begin{cases} x = \frac{\alpha}{a} - \frac{b}{a}y \\ c\left(\frac{\alpha}{a} - \frac{b}{a}y\right) + dy = \beta \end{cases} \quad \begin{cases} x = \frac{\alpha}{a} - \frac{b}{a}y \\ \left(d - \frac{bc}{a}\right)y = \beta - \frac{\alpha c}{a} \end{cases}$$

da cui ricavando y e sostituendo nella prima

$$\begin{cases} x = \frac{\alpha}{a} - \frac{b}{a}y \\ y = \frac{\beta - \frac{\alpha c}{a}}{d - \frac{bc}{a}} \end{cases} \quad \begin{cases} x = \frac{\alpha}{a} - \frac{b}{a}y \\ y = \frac{a\beta - \alpha c}{ad - bc} \end{cases} \quad \begin{cases} x = \frac{\alpha}{a} - \frac{b}{a}\frac{a\beta - \alpha c}{ad - bc} \\ y = \frac{a\beta - \alpha c}{ad - bc} \end{cases}$$

Ma

$$x = \frac{\alpha}{a} - \frac{b}{a}\frac{a\beta - \alpha c}{ad - bc} = \frac{\alpha ad - \alpha bc - ba\beta + bc\alpha}{a(ad - bc)} = \frac{\alpha ad - ba\beta}{ad - bc} = \frac{\alpha d - \beta b}{ad - bc}$$

e le soluzioni del sistema sono date da

$$\begin{cases} x = \frac{\alpha d - \beta b}{ad - bc} \\ y = \frac{a\beta - \alpha c}{ad - bc} \end{cases}$$

Il procedimento che abbiamo usato per la soluzione del sistema, che è noto come procedimento di eliminazione Gaussiana, può essere formalizzato usando i soli elementi della matrice A e del vettore B.

A questo scopo si considera la matrice aumentata del sistema che si ottiene affiancando alla matrice dei coefficienti A il vettore B

$$M = \left(\begin{array}{cc|c} a & b & \alpha \\ c & d & \beta \end{array}\right)$$

A tal proposito è opportuno sottolineare il diverso ruolo che A e B ricoprono; la matrice A definisce l'applicazione lineare mentre il vettore B è un elemento di cui si vuole stabilire se appartiene all'immagine, $R(A)$, dell'applicazione lineare stessa.

È facile riconoscere che il sistema di partenza è espresso dall'uguaglianza

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

e che mediante opportune manipolazioni ci si è ricondotti ad una formulazione del tipo

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\alpha d - \beta b}{ad-bc} \\ \frac{a\beta - \alpha c}{ad-bc} \end{pmatrix}$$

Vediamo allora che il sistema nella sua formulazione finale è associato ad una matrice completa che è:

$$M^{rref} = \left(\begin{array}{cc|c} 1 & 0 & \frac{\alpha d - \beta b}{ad-bc} \\ 0 & 1 & \frac{a\beta - \alpha c}{ad-bc} \end{array}\right)$$

Possiamo passare da M ad M^{rref} mediante una serie di operazioni del tipo:

- scambiare due righe della matrice (significa cambiare l'ordine con cui si considerano le equazioni).
- moltiplicare per una costante k una riga della matrice. (significa moltiplicare per k ambo i membri di un'equazione)
- sottrarre da una riga un'altra riga moltiplicata per una costante k. (significa ricavare una variabile da un'equazione e sostituirla in un'altra equazione)

Evidentemente, a meno dell'operazione di scambio di righe, le operazioni considerate forniscono la possibilità di sostituire ad una riga una combinazione lineare di tutte le righe ed inoltre ad essa si riducono.

Se R_i ed R_j stanno ad indicare due generiche righe di M (e quindi individuano, ciascuna, un'equazione del sistema), possiamo usare la seguente simbologia per rappresentare, nell'ordine, le operazioni precedentemente elencate.

- $R_i \longleftrightarrow R_j$
- $kR_i \to R_i$
- $R_j + kR_i \to R_j$

Riportiamo di seguito i passaggi per ottenere M^{rref} da M.

$$M = \left(\begin{array}{cc|c} a & b & \alpha \\ c & d & \beta \end{array}\right) \quad \frac{1}{a}R_1 \to R_1 \quad \left(\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{\alpha}{a} \\ c & d & \beta \end{array}\right)$$

$$R_2 - cR_1 \to R_2 \quad \left(\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{\alpha}{a} \\ 0 & d - \frac{bc}{a} & \beta - \frac{\alpha c}{a} \end{array}\right)$$

$$R_2 \frac{1}{d - \frac{bc}{a}} \to R_2 \quad \left(\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{\alpha}{a} \\ 0 & 1 & \frac{\beta - \frac{\alpha c}{a}}{d - \frac{bc}{a}} \end{array}\right) = M^{rr}$$

$$R_1 - R_2\frac{b}{a} \to R_1 \quad \left(\begin{array}{cc|c} 1 & 0 & \frac{\alpha}{a} - \frac{b}{a}\frac{\beta-\frac{\alpha c}{a}}{d-\frac{bc}{a}} \\ 0 & 1 & \frac{\beta-\frac{\alpha c}{a}}{d-\frac{bc}{a}} \end{array}\right)$$

e

$$\left(\begin{array}{cc|c} 1 & 0 & \frac{\alpha}{a} - \frac{b}{a}\frac{\beta-\frac{\alpha c}{a}}{d-\frac{bc}{a}} \\ 0 & 1 & \frac{\beta-\frac{\alpha c}{a}}{d-\frac{bc}{a}} \end{array}\right) = \left(\begin{array}{cc|c} 1 & 0 & \frac{\alpha d-\beta b}{ad-bc} \\ 0 & 1 & \frac{a\beta-\alpha c}{ad-bc} \end{array}\right) = M^{rref}$$

Nel corso del procedimento abbiamo individuato dapprima una matrice che abbiamo chiamato M^{rr} e che definiamo ***matrice ridotta per righe (row reduced)*** e siamo infine pervenuti alla matrice che abbiamo indicato con M^{rref} e che chiamiamo ***matrice ridotta a scala (row reduced echelon form)***.

Ciascuna delle trasformazioni sulle righe può essere ottenuta semplicemente moltiplicando a sinistra A per una matrice che chiamiamo ***matrice elementare***

Nel caso 2×2 si verifica subito che

- la matrice
$$T_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
che chiamiamomatrice di permutazione, opera lo scambio della prima e della seconda riga;

- la matrice
$$T_2(k) = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$$
opera la moltiplicazione della seconda riga per k;

- la matrice
$$T_{12}(k) = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$$
sostituisce alla seconda riga la somma tra la seconda riga stessa e la prima riga moltiplicata per k.

Si vede subito che T_{12}, $T_{12}(k)$ e per $k \neq 0$, $T_2(k)$ sono invertibili e che $T_{12}^{-1} = T_{12}$, $T_2(k)^{-1} = T_2(1/k)$ e $T_{12}(k)^{-1} = T_{12}(-k)$.

Con riferimento al caso precedente, considerando la matrice A in luogo della matrice completa M, possiamo trovare una matrice A^{rref} che chiamiamo matrice ridotta a scala (***row reduced echelon-form matrix***) mediante le stesse manipolazioni sulle righe:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \frac{1}{a}R_1 \to R_1 \quad \begin{pmatrix} 1 & \frac{b}{a} \\ c & d \end{pmatrix}$$

$$R_2 - cR_1 \to R_2 \quad \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & d - \frac{bc}{a} \end{pmatrix}$$

$$R_2 \frac{1}{d - \frac{bc}{a}} \to R_2 \quad \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix} = U$$

$$R_1 - R_2 \frac{b}{a} \to R_1 \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A^{rref}$$

Dal momento che abbiamo applicato esclusivamente operazioni elementari sulle righe, possiamo ottenere lo stesso risultato semplicemente moltiplicando a sinistra per opportune matrici elementari

$$U = T_2\left(d - \frac{bc}{a}\right) T_{21}(c) T_1\left(\frac{1}{a}\right) A$$

Essendo le matrici elementari utilizzate triangolari inferiori ed essendo il prodotto di matrici triangolari inferiori ancora triangolare inferiore, si ha che $T_2(d - \frac{bc}{a})T_{21}(c)T_1(\frac{1}{a})$ è essa stessa una matrice triangolare inferiore.

Inoltre

$$\begin{aligned} &\left(T_2\left(d - \frac{bc}{a}\right) T_{21}(c)\, T_1\left(\frac{1}{a}\right)\right)^{-1} = \\ &\quad = T_1\left(\frac{1}{a}\right)^{-1} (T_{21}(c))^{-1} \left(T_2\left(d - \frac{bc}{a}\right)\right)^{-1} = \\ &\quad = T_1(a)\,(T_{21}(-c))\left(T_2\left(\frac{1}{d - \frac{bc}{a}}\right)\right) = L \end{aligned}$$

è ancora una matrice triangolare inferiore.

Se osserviamo che U è triangolare superiore e che risulta

$$LU = A$$

vediamo che la matrice A è decomposta nel prodotto di due matrici, L triangolare inferiore ed U triangolare superiore.

Osserviamo esplicitamente che abbiamo supposto che

$$a \neq 0 \qquad , \qquad d - \frac{bc}{a} \neq 0$$

e quindi non abbiamo avuto bisogno di operare scambi di righe per completare il procedimento.

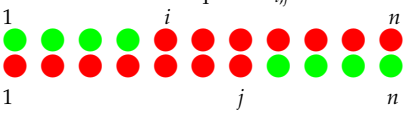
Nel caso generale quindi potrebbe non essere possibile scrivere una matrice come prodotto di due matrici triangolari tuttavia, se la matrice A non è singolare, si può trovare una matrice di permutazione P tale che $PA = LU$.

Per illustrare il procedimento consideriamo il seguente esempio.

Sia

Il prodotto di due matrici triangolari inferiori è a sua volta triangolare inferiore, infatti siano $A = \{a_{ij}\}$, $B = \{b_{ij}\}$ matrici triangolari inferiori, sia cioè $a_{ij} = 0$ e $b_{ij} = 0$ se $i < j$.

Consideriamo il prodotto $C = AB = c_{ij} = \sum_{k=0}^{n} a_{ik} b_{kj}$; poichè $a_{ik} = 0$ se $k > i$ e $b_{kj} = 0$ se $k < j$, nel caso in cui $i < j$ avremo che per ogni k almeno uno dei due fattori è nullo e quindi $c_{i,j} = 0$.



$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 12 & 14 \\ 1 & 4 & 22 & 4 \\ 1 & 4 & 13 & -2 \end{pmatrix}$$

e procediamo mediante le seguenti operazioni:

$$\begin{cases} R_2 - R_1 \to R_2 \\ R_3 - R_1 \to R_3 \\ R_4 - R_1 \to R_4 \end{cases} \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 9 & 10 \\ 0 & 2 & 19 & 0 \\ 0 & 2 & 10 & -6 \end{pmatrix}$$

$$R_3 - R_4 \to R_3 \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 9 & 6 \\ 0 & 2 & 10 & -6 \end{pmatrix}$$

$$R_2 - R_3 \to R_2 \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 9 & 6 \\ 0 & 2 & 10 & -6 \end{pmatrix}$$

$$\frac{1}{4}R_2 \to R_2 \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 9 & 6 \\ 0 & 2 & 10 & -6 \end{pmatrix}$$

$$\frac{1}{9}R_3 \to R_3 \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{6}{9} \\ 0 & 2 & 10 & -6 \end{pmatrix}$$

$$\frac{1}{2}R_4 \to R_4 \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{6}{9} \\ 0 & 1 & 5 & -3 \end{pmatrix} = U^0$$

Fin qui abbiamo operato con trasformazioni elementari e quindi abbiamo trovato una matrice Λ, prodotto di tutte le matrici elementari che operano le trasformazioni che abbiamo applicato, tale che

$$U^0 = \Lambda A$$

U^0 non è tuttavia triangolare superiore ma si vede subito che permutandone le righe si può ottenere una matrice triangolare superiore. Più precisamente occorre operare lo scambio $R_2 \longleftrightarrow R_4$.

La permutazione può essere effettuata moltiplicando per la matrice

$$P^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

che si ottiene dallo scambio $R_2 \longleftrightarrow R_4$. operato sulla matrice identica ed opera lo scambio tra 2 e la quarta riga.

Se allora consideriamo $P^0 A$ ed procediamo come visto in precedenza, otteniamo che

$$P^0 A = LU$$

e

$$A = PLU$$

essendo P l'inversa di P^0.

Tornando al caso 2×2, se definiamo il determinante della matrice A come

$$\begin{pmatrix} a & & b \\ & \searrow \swarrow & \\ & \swarrow \searrow & \\ c & & d \end{pmatrix}$$

$$\det(A) = ad - bc$$

e poniamo

$$A^1 = \begin{pmatrix} \alpha & b \\ \beta & d \end{pmatrix} \quad \text{e} \quad A^2 \begin{pmatrix} a & \alpha \\ b & \beta \end{pmatrix}$$

vediamo che le soluzioni del sistema lineare 4.5 possono essere scritte come

$$\begin{cases} x = \frac{\det(A^1)}{\det(A)} \\ x = \frac{\det(A^2)}{\det(A)} \end{cases}$$

4.4 *Il caso generale*

Consideriamo ora una generica matrice quadrata in $\mathcal{M}^n$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

Il caso in cui la prima colonna sia tutta di elementi nulli non è significativo in quanto, in tal caso, la prima variabile del sistema associato non compare nel sistema stesso.

Supponiamo quindi che esista almeno un elemento della prima colonna non nullo e supponiamo che sia $a_{11} \neq 0$ (se così non fosse potremmo operare uno scambio di righe).

Sottraendo dalla $j-$esima riga la prima moltiplicata per $\frac{a_{j1}}{a_{11}}$, operando cioè la trasformazione

$$R_j - \frac{a_{j1}}{a_{11}} R_1$$

otteniamo

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} - \frac{a_{21}}{a_{11}} a_{11} & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & a_{23} - \frac{a_{21}}{a_{11}} a_{13} & \cdots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ a_{31} - \frac{a_{31}}{a_{11}} a_{11} & a_{32} - \frac{a_{31}}{a_{11}} a_{12} & a_{33} - \frac{a_{31}}{a_{11}} a_{13} & \cdots & a_{3n} - \frac{a_{31}}{a_{11}} a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} - \frac{a_{n1}}{a_{11}} a_{11} & a_{n2} - \frac{a_{n1}}{a_{11}} a_{12} & a_{n3} - \frac{a_{n1}}{a_{11}} a_{13} & \cdots & a_{nn} - \frac{a_{n1}}{a_{11}} a_{1n} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & a_{23} - \frac{a_{21}}{a_{11}} a_{13} & \cdots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ 0 & a_{32} - \frac{a_{31}}{a_{11}} a_{12} & a_{33} - \frac{a_{31}}{a_{11}} a_{13} & \cdots & a_{3n} - \frac{a_{31}}{a_{11}} a_{1n} \\ 0 & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2} - \frac{a_{n1}}{a_{11}} a_{12} & a_{n3} - \frac{a_{n1}}{a_{11}} a_{13} & \cdots & a_{nn} - \frac{a_{n1}}{a_{11}} a_{1n} \end{pmatrix} =$$

Possiamo formalizzare quanto fatto come segue in modo da condensare in un'unica notazione le operazioni eseguite sulle $n-1$ righe dopo la prima.

Scriviamo A nella forma

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

dove $a_{11} \in \mathbb{R}$, A_{12} è un vettore riga $(n-1)-$dimensionale , A_{21} è un vettore colonna $(n-1)-$dimensionale e A_{22} è una matrice quadrata $(n-1)-$dimensionale.

Diciamo che la matrice A è scritta in forma partizionata ***a blocchi***; le sue righe sono costituite da blocchi di diverse dimensioni, ai quali si possono però applicare le operazioni di somma e di prodotto pur di tenere nel giusto conto le dimensioni dei blocchi.

Ragionando come nell'esempio riportato a margine possiamo affermare che, utilizzando solo operazioni elementari rappresentate da matrici invertibili, (che sono triangolari inferiori a meno che non si tratti di permutazioni di righe), si ottiene:

$$A = \begin{pmatrix} a_{11} & A_{12} \\ 0 & A_{22} - \frac{1}{a_{11}} A_{21} A_{12} \end{pmatrix}$$

essendo $A_{21}A_{12}$ una matrice quadrata $(n-1)-$dimensionale.

Se la prima colonna di $A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$ ha elementi non nulli, che possiamo portare nelle prime righe con permutazioni, ripetiamo l'operazione ed otteniamo una seconda colonna con tutti elementi nulli tranne, eventualmente, i primi due, questo risultato essendo già

Se ad esempio

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

dove

$$A_{21} = \begin{pmatrix} a_{21} \\ a_{31} \end{pmatrix} \quad A_{12} = \begin{pmatrix} a_{12} & a_{13} \end{pmatrix}$$

e

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

Operando su A con $R_i - \frac{a_{i1}}{a_{11}} R_1 \longrightarrow R_i$ otteniamo

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

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} - \frac{a_{21}a_{12}}{a_{11}} & a_{23} - \frac{a_{21}a_{13}}{a_{11}} \\ 0 & a_{32} - \frac{a_{31}a_{13}}{a_{11}} & a_{33} - \frac{a_{31}a_{13}}{a_{11}} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & A_{12} \\ 0 & A_{22} - \frac{1}{a_{11}} A_{21} A_{12} \end{pmatrix}$$

ottenuto nel caso in cui la prima colonna di $A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$ abbia elementi tutti nulli.

Iterando possiamo ricondurci ad una matrice, i cui elementi sotto la diagonale principale sono tutti nulli, che indichiamo con A^{rr} e chiamiamo matrice ridotta per righe.

Tale matrice avrà sulla diagonale principale n elementi che chiamiamo ***pivot*** che possono essere nulli.

Osserviamo che A^{rr} si ottiene da A moltiplicando a sinistra per matrici elementari che risultano invertibili e che, a meno di quelle che rappresentano scambio di righe, sono triangolari inferiori.

Nel caso in cui il numero di righe n sia superiore al numero delle colonne m si otterranno almeno $n - m$ righe nulle. Osserviamo anche che il numero di righe nulle può essere strettamente più grande di $n - m$.

Più precisamente, se $r(\leq n)$ è la caratteristica della matrice A, cioè se r è la dimensione dello spazio vettoriale generato dalle righe di A, rimarranno solo r righe non nulle ciascuna delle quali avrà un primo elemento diverso da 0 (da sinistra verso destra) che chiamiamo pivot.

Nel caso in cui $r = n$ il sistema lineare

$$Ax = B$$

associato alla matrice avrà un'unica soluzione per ogni assegnato B.

Qualora invece sia $r < n$ l'esistenza di una soluzione dipende dalla scelta di B.

Inoltre è garantita l'unicità nel caso e solo nel caso in cui non ci siano righe nulle e gli elementi sulla diagonale principale, cioè i pivot, siano tutti diversi da 0.

Ad ogni singolo passo il procedimento trasforma una matrice $A^{(n)}$, (l'indice (n) indica l'ordine della matrice ma tiene anche traccia dei passi eseguiti) nella forma :

$$\begin{pmatrix} a_{11}^{(n)} & A_{12}^{(n)} \\ 0 & A^{(n-1)} \end{pmatrix}$$

dove

$$A^{(n-1)} = A_{22}^{(n)} - \frac{1}{a_{11}^{(n)}} A_{21}^{(n)} A_{12}^{(n)}$$

Se chiamiamo $a_{ij}^{(k)}$ gli elementi della matrice $A^{(k)}$ possiamo affermare che vale la seguente regola di ricorrenza

$$a_{ij}^{(n-1)} = a_{i+1\,j+1}^{(n)} - \frac{1}{a_{11}^{(n)}} a_{(i+1)\,1}^{(n)} a_{1\,j+1}^{(n)} = \frac{a_{11}^{(n)} a_{i+1\,j+1}^{(n)} - a_{(i+1)\,1}^{(n)} a_{1\,j+1}^{(n)}}{a_{11}^{(n)}} \tag{4.6}$$

Si verifica subito che nel caso di una matrice 2×2

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{23} \end{pmatrix} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(2)} & a_{23}^{(2)} \end{pmatrix}$$

ci si riduce a

$$\begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ 0 & a_{11}^{1} \end{pmatrix}$$

dove

$$a_{11}^{1} = \frac{a_{11}^{(2)} a_{22}^{(2)} - a_{21}^{(2)} a_{1\,2}^{(2)}}{a_{11}^{(2)}}$$

Quanto illustrato motra che si può agire su una matrice con operazioni sulle righe, che chiamiamo operazioni elementari, che ci consentono di ottenere una matrice più semplice associata ad un sistema lineare equivalente a quello di poartenza, cioè con le stesse soluzioni.

Come abbiamo già osservato occupandoci del caso 2×2 le operazioni elementari sulle righe sono le seguenti:

- $R_i \longleftrightarrow R_j$
- $kR_i \to R_i$
- $R_j + kR_i \to R_j$

e si applicano moltiplicando a sinistra per una matrice che indichiamo con

- T_{ij}
- $T_i(k)$
- $T_{ij}(k)$

che si ottiene eseguendo l'operazione in esame sulla matrice identica.

Naturalmente quanto detto per le operazioni sulle righe può essere esteso alle operazioni sulle colonne.

4.5 Il determinante di una matrice

Definizione 4.5 *Chiamiamo funzione determinante. una funzione*

$$D : \mathcal{M}^n \to \mathbb{R}$$

definita sullo spazio delle matrici quadrate di ordine n

$$\mathcal{M}^n \ni A = \begin{pmatrix} C_1 & C_2 & \dots C_n \end{pmatrix} \mapsto D(A) = D(C_1, C_2, \dots, C_n) \in \mathbb{R}$$

soddisfacente, le seguenti condizioni:

- *D è lineare rispetto a ciascuna colonna della matrice suo argomento: in altre parole:*

$$D(C_1, C_2, \ldots, \alpha C' + \beta C'', \ldots, C_n) = \alpha D(C_1, C_2, \ldots, C', \ldots, C_n) + \beta D(C_1, C_2, \ldots, C'', \ldots, C_n)$$

- *Scambiando due colonne il valore di D cambia segno:*

$$D(C_1, C_2, \ldots, C', \ldots, C'', \ldots, C_n) = -D(C_1, C_2, \ldots, C'', \ldots, C', \ldots, C_n)$$

- *Il determinante della matrice identica è 1:*

$$D(I_n) = 1$$

Teorema 4.5 *Affermare che*

$$D(C_1, C_2, \ldots, C', \ldots, C'', \ldots, C_n) = -D(C_1, C_2, \ldots, C'', \ldots, C', \ldots, C_n)$$

è equivalente ad affermare che

$$D(C_1, C_2, \ldots, C', \ldots, C', \ldots, C_n) = 0$$

DIMOSTRAZIONE.
Poniamo per comodità di notazioni

$$D(C', C'') = D(C_1, C_2, \ldots, C', \ldots, C'', \ldots, C_n)$$

Supponiamo che

$$D(C', \ldots, C'') = -D(C'', \ldots, C')$$

e siano $C' = C''$ due colonne uguali in A, si ha

$$\begin{aligned} D(A) = D(C_1, C_2, \ldots, C', \ldots, C'', \ldots, C_n) = \\ = -D(C_1, C_2, \ldots, C'', \ldots, C', \ldots, C_n) = \\ = -D(C_1, C_2, \ldots, C', \ldots, C'', \ldots, C_n) = -D(A) \end{aligned}$$

da cui $D(A) = 0$
Viceversa se

$$D(C_1, C_2, \ldots, C', \ldots, C', \ldots, C_n) = 0$$

si ha

$$\begin{aligned} D(C', C'') = D(C', C'') + D(C', C') = D(C', C' + C'') = \\ = D(-C'', C' + C'') + D(C' + C'', C' + C'') = \\ = -D(C'', C' + C'') = -D(C'', C') - D(C'', C'') = -D(C'', C') \end{aligned}$$

□

Nel caso si disponga di una funzione determinante è possibile risolvere un sistema lineare di equazioni algebriche.

Sia infatti

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n \end{cases}$$

un sistema di n equazioni lineari in n incognite e sia

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,m} \\ a_{21} & a_{22} & \cdots & a_{2,m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,m} \end{pmatrix} = \{a_{ij}\} = \begin{pmatrix} C_1 & C_2 & \ldots & C_n \end{pmatrix}$$

la matrice dei suoi coefficienti, per la quale supponiamo che $D(A) \neq 0$.
Sia inoltre

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} \qquad \text{e} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}$$

per modo che il sistema può essere scritto nella forma

$$Ax = y$$

Definiamo ancora

$$A^j = \begin{pmatrix} C_1 & C_2 & \ldots & C_{j-1} & b & C_{j+1} & \ldots & C_n \end{pmatrix}$$

Avremo che $Ax = b$ se solo se

$$D(A^j) = D\left(\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & b_2 & a_{2(j+1)} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & b_n & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}\right) =$$

$$= D\left(\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & \sum_k a_{1k}x_k & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & \sum_k a_{2k}x_k & a_{2(j+1)} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & \sum_k a_{nk}x_k & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}\right) =$$

$$= \sum_k x_k D\left(\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1k} & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2k} & a_{2(j+1)} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{nk} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}\right) =$$

$$= x_j D\left(\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1j} & a_{1(j+1)} & \cdots & a_{1,m} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2j} & a_{2(j+1)} & \cdots & a_{2,m} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \cdots & a_{n,m} \end{pmatrix}\right) =$$

$$= x_j D(A)$$

in quanto nella somma compaiono i determinanti di matrici aventi due colonne uguali, tranne che nel caso in cui $k = j$

Possiamo pertanto affermare che sono fatti equivalenti:

1. $Ax = b$
2. $D(A^j) = x_j D(A)$ per ogni $j = 1.\ldots n$

Ne segue che

1. se $D(A) \neq 0$ allora il sistema ha una ed una sola soluzione
2. se invece $D(A) = 0$ allora il sistema o non ammette soluzioni (nel caso $D(A^j) \neq 0$ per qualche j) o ne ammette infinite (nel caso $D(A^j) = 0$ per ogni j).

Possiamo pertanto enunciare il seguente teorema.

Teorema 4.6 - **Regola di Cramer** *Se $A \in \mathcal{M}^n$ è una matrice quadrata Allora il sistema lineare $Ax = b$ ammette soluzione se e solo se $D(A) \neq 0$*

Inoltre si ha

$$x_j = \frac{D(A^j)}{D(A)}$$

Possiamo verificare che valgono i seguenti fatti:

Teorema 4.7 *Sia $A \in \mathcal{M}^n$ una matrice quadrata di ordine n e sia*

$$D : \mathcal{M}^n \to \mathbb{R}$$

una funzione determinante, allora

- $D(T_{ij}A) = -D(A)$
- $D(T_j(k)A) = kD(A)$
- $T_{ij}(k)A = kD(A)$

DIMOSTRAZIONE.

La prima affermazione discende dal fatto che D cambia segno se scambiamo due colonne.

La seconda è conseguenza della linearità di D rispetto alle colonne.

La terza si ottiene dalla linearità e dal fatto che D si annulla sulle matrici con due colonne uguali.

□

Vogliamo a questo punto trovare un'espressione esplicita per il calcolo del determinante di una matrice A

Definizione 4.6 *Sia $I = \{1, 2, 3, \ldots, n\}$ l'insieme dei primi n numeri naturali; chiamiamo permutazione ogni funzione $\pi : I \to I$ bigettiva .*

Chiamiamo segno della permutazione il valore $(-1)^s$ dove s è il numero di scambi tra elementi di I adiacenti necessario per passare da I a $\pi(I)$

Teorema 4.8 *Sia $A \in \mathcal{M}^n$ allora*

$$\det(A) = \sum_{\pi \in \mathcal{P}_n} \sigma(\pi) a_{1\pi(1)} a_{1\pi(1)} \cdots a_{n\pi(n)}$$

DIMOSTRAZIONE.

Sia ii $A \in \mathcal{M}^n$

$$A = \begin{pmatrix} C_1 & C_2 & \cdots, C_n \end{pmatrix}$$

con C_j vettori colonna in $\mathbb{R}^n$. Se indichiamo con E_j , $j = 1, 2, \ldots, n$ i vettori di una base canonica di $\mathbb{R}^n$ intesi come vettori colonna, possiamo scrivere:

$$C_j = \sum_{i=1}^{n} a_{i,j} E_i$$

per modo che

$$D(A) = D(C_1, C_2, \ldots, C_n) = D(\sum_{i_1=1}^{n} a_{i_1,1} E_{i_1}, \sum_{i_2=1}^{n} a_{i_2,2} E_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{i_n,n} E_{i_n}) =$$

$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} D(E_{i_1}, E_{i_2}, \ldots, E_{i_n}) =$$

Se definiamo

$$E = \begin{pmatrix} E_{i_1} & E_{i_2} & \cdot & , E_{i_n} \end{pmatrix}$$

possiamo constatare che $E \in \mathcal{M}^n$ è costituita da n vettori colonna scelti dalla base canonica di $\mathbb{R}^n$ ed inoltre si ha

Se E ha due colonne uguali $D(E) = 0$ per cui la somma deve essere estesa solo ai casi in cui nella matrice E sono presenti tutti i vettori E_j una sola volta

In questo caso E si può ottenere da una permutazione delle colonne della matrice identica I_n e pertanto, tenendo conto della seconda proprietà che definisce D,

$$D(E) = \sigma \det(I_n)$$

essendo $\sigma = \pm 1$ a seconda che il numero di scambi effettuati sia pari o dispari.

Poichè $D(I_n) = 1$ si vede subito che

$$D(A) = \sum_{\pi \in \mathcal{P}_n} \sigma(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$$

□

Le proprietà del determinante forniscono anche un metodo ricorsivo per il suo calcolo che, tuttavia, non è efficiente dal punto di vista computazionale.

Sia

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

una matrice quadrata di ordine n e definiamo

$$A_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \ddots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \ddots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$

come la matrice che si ottiene eliminando la $i-$esima riga e la $j-$esima colonna.

Sia, come al solito, la matrice A identificata dalle sue colonne C_j.

$$A = (C_1 C_2 \dots C_n)$$

Poichè

$$C_1 = a_{11}E_1 + a_{21}E_2 + \cdots + a_{n1}E_n$$

per la linearità del determinante rispetto alle colonne, si ha

$$\det(A) = a_{11}\det\begin{pmatrix} 1 & a_{12} & \cdots, & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix} +$$

$$+ a_{21}\det\begin{pmatrix} 0 & a_{12} & \cdots, & a_{1,n} \\ 1 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix} + \cdots +$$

$$+ a_{n1}\det\begin{pmatrix} 0 & a_{12} & \cdots, & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 1 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix}$$

$$A_1' = \begin{pmatrix} 1 & a_{12} & \cdots, & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix}$$

$$A_2' = \begin{pmatrix} 0 & a_{12} & \cdots, & a_{1,n} \\ 1 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix}$$

$$A_n' = \begin{pmatrix} 0 & a_{12} & \cdots, & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ 1 & a_{n2} & \cdots & ,a_{n,n} \end{pmatrix}$$

Coè

$$\det(A) = a_{11}\det(A_1') + a_{21}\det(A_2') + \cdots + a_{n1}\det(A_n')$$

Ma

$$\det(A_1') = 1 \sum_{\pi\in\mathcal{P}_n} \sigma(\pi)a_{2\pi(2)}a_{3\pi(3)}\cdots a_{n\pi(n)} = \det(A_{11})$$
$$\det(A_2') = 1 \sum_{\pi\in\mathcal{P}_n} \sigma(\pi)a_{1\pi(1)}a_{3\pi(3)}\cdots a_{n\pi(n)} = -\det(A_{21})$$
$$\det(A_n') = 1 \sum_{\pi\in\mathcal{P}_n} \sigma(\pi)a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n-1\pi(n-1)} = (-1)^n\det(A_{n1})$$

in quanto in ciascuna matrice la prima colonna ha un solo elemento non nullo ed uguale a 1 ed il segno delle permutazioni nel calcolo di $\det(A_{i,1})$ differisce da $\sigma(\pi)$ per un fattore $(-1)^i$.

In maniera del tutto analoga si può verificare che

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

4.6 Esistenza ed unicità delle soluzioni di un sistema lineare

Sia $A \in \mathcal{M}^{m\times n}$; A una matrice con n righe ed m colonne e consideriamo $B \in \mathbb{R}^m$.

Identifichiamo A mediante le sue m colonne $C_j \in \mathbb{R}^n$ per modo che

$$A = \begin{pmatrix} C_1 & C_2 & \cdots & C_m \end{pmatrix}$$

e consideriamo il sistema di equazioni lineari $Ax = B$ che possiamo scrivere nella forma

$$Ax = \begin{pmatrix} C_1 & C_2 & \cdots & C_m \end{pmatrix} x = B$$

o anche come

$$\sum_{j=1}^{m} x_j C_j = B$$

Dalla precedente formulazione si vede che il sistema ha soluzione se e solo se

$$B \in \mathcal{S}_c(A)$$

cioè se e solo se B appartiene al sottospazio di $\mathbb{R}^n$ generato dalle colonne $C = \{C_1, C_2, \ldots, C_m\}$ della matrice A. e ciò può avere luogo se e solo se lo spazio $\mathcal{S}_c(A)$ generato dalle colonne di A coincide con lo spazio $\mathcal{S}_c((A|B))$ generato dalle colonne di $(A|B)$.

Se

$$\dim \mathcal{S}_c(A) = \rho^c_A$$

è la caratteristica per colonne della matrice A e

$$\dim \mathcal{S}_c((A|B)) = \rho^c_{(A|B)}$$

è la caratteristica per colonne della matrice aumentata $(A|B)$, possiamo affermare che il sistema ha soluzioni se e solo se

$$\rho^c_A = \rho^c_{(A|B)}$$

Ma per ogni matrice la caratteristica per righe e la caratteristica per colonne sono uguali

$$\rho^c_A = \rho^r_{(A)} = \rho_A$$

e possiamo pertanto concludere che

Teorema 4.9 (di Rouché-Capelli) *Sia $A \in \mathcal{M}^{m\times n}$; A una matrice con n righe ed m colonne e $B \in \mathbb{R}^m$. Il sistema di equazioni lineari*

$$Ax = B$$

ammette soluzioni se e solo se

$$\rho_A = \rho_{(A|B)}$$

Consideriamo ora il sistema

$$Ax = B$$

e l'applicazione lineare

$$L_A : \mathbb{R}^n \to \mathcal{R}^m \qquad \text{definita da} \qquad L_A(x) = Ax$$

e supponiamo che il sistema abbia soluzioni per modo che risulta

$$\dim(R(L_A)) = \rho_A = \rho_{(A|B)} = \bar{\rho}$$

Poichè lo spazio vettoriale generato dalle righe di A (o di $(A|B)$) ha dimensione $\bar{\rho}$ avremo che

- se $\bar{\rho} = n$

  allora $\dim(R(L_A)) = n$ e di conseguenza $\dim(\ker(L_A)) = 0$.

  Ne segue che il sistema ammette la sola soluzione nulla, nel caso in cui $B = 0$ e, per la linearità, che il sistema ammette soluzione unica $\forall B \in \mathbb{R}^n$.

- se $\bar{\rho} < n$

  allora $\dim(\ker(L_A)) = n - \bar{\rho} > 0$.

  Pertanto il sistema $Ax = 0$ ha soluzioni che formano uno spazio vettoriale di dimensione $n - \bar{\rho}$ mentre il sistema $Ax = B$ ha soluzioni che formano uno spazio affine di dimensione $n - \bar{\rho}$

Ora, chiamiamo ***minore di ordine*** k di A ogni matrice $k \times k$ ottenuta scegliendo k righe r k colonne da A.

Dal momento che la matrice A ha caratteristica ρ_A allora se esistono ρ_A colonne, linearmente indipendenti che generano $R(A)$.

Sia $C \in \mathcal{M}^{n \times \rho_A}$ la matrice ottenuta considerando tali colonne.

Avremo che C ha ancora caratteristica ρ_A ed è quindi possibile trovare in C ρ_A righe linearmente indipendenti.

Possiamo allora considerare il minore $D \in \mathcal{M}^{\rho_A \times \rho_A}$ costituito da tali righe ed avremo che la sua caratteristica è ancora ρ_A.

Pertanto $\det(D) \neq 0$.

Viceversa, ogni minore di ordine di ordine più grande di ρ_A contiene righe (o colonne) linearmente dipendenti e quindi ha determinante nullo.

Ne segue che la caratteristica di una matrice è uguale al massimo ordine dei minori con determinante non nullo.

4.7 *Decomposizione SVD (Singular Value Decomposition) di una matrice*

Si consideri una matrice $A \in \mathcal{M}^{m \times n}$; A definisce una applicazione lineare da $\mathbb{R}^n$ in $\mathbb{R}^m$ che opera essenzialmente mediante rotazioni e

stiramenti, o allungamenti. Tali effetti possono essere evidenziati decomponendo la matrice nel prodotto di tre matrici, ciascuna con un effetto specifico.

Il tutto si basa sul seguente risultato.

Teorema 4.10 *Sia $A \in \mathcal{M}^{m\times n}$; è possibile trovare una base $\{b_1, b_2 \dots, b_n\}$ di $\mathbb{R}^n$ con $\|b_i\| = 1$ tale che $\{Ab_1, Ab_2, \dots, Ab_n\}$ sia un insieme di vettori a due a due perpendicolari.*

Dimostrazione. Siano $b_1, b_2, b_3, \dots, b_n \in \mathbb{R}^n$ tali che

$$\|Ab_1\|^2 = \max_{\substack{v\in\mathbb{R}^n \\ \|v\|=1}} \|Av\|^2 \qquad \text{e} \qquad S_2 = (\mathcal{S}(\{b_1\})^\perp$$

$$\|Ab_2\|^2 = \max_{\substack{v\in S_2 \\ \|v\|=1}} \|Av\|^2 \qquad \text{e} \qquad S_3 = (\mathcal{S}(\{b_1, b_2\})^\perp$$

$$\|Ab_3\|^2 = \max_{\substack{v\in S_3 \\ \|v\|=1}} \|Av\|^2 \qquad \text{e} \qquad S_4 = (\mathcal{S}(\{b_1, b_2, b_3\})^\perp$$

$$\|Ab_n\|^2 = \max_{\substack{v\in S_n \\ \|v\|=1}} \|Av\|^2$$

Evidentemente $\{b_1, b_2 \dots, b_n\}$ sono n vettori ortonormali in $\mathbb{R}^n$ e quindi costituiscono una base per $\mathbb{R}^n$ inoltre $\{Ab_1, Ab_2 \dots, Ab_n\}$ sono a due a due perpendicolari.

Se infatti $c_1b_1 + c_2b_2 + \dots c_nb_n = 0$, si ha

$$c_j\|b_j\|^2 = \langle c_1b_1 + c_2b_2 + \dots c_nb_n, b_j\rangle = \\ = \langle 0, b_j\rangle = 0$$

da cui $c_j = 0$.

Pertanto i b_k sono n vettori linearmente indipendenti in $\mathbb{R}^n$ e ne costituiscono una base.

Sia, infatti,

$$v(t) = b_i\cos(t) + b_j\sin(t)$$

Si ha

$$\|v(t)\|^2 = \langle b_i\cos(t) + b_j\sin(t), b_i\cos(t) + b_j\sin(t)\rangle = \\ = \|b_i\|^2\cos^2(t) + \|b_j\|^2\sin^2(t) + 2\langle b_i, b_j\rangle\cos(t)\sin(t) = \\ = \|b_i\|^2\cos^2(t) + \|b_j\|^2\sin^2(t) = \cos^2(t) + \sin^2(t) = 1$$

Dal momento che $b_i = v(0)$ e $\|Ab_i\|^2 = \max_{\substack{v\in S_i \\ \|v\|=1}} \|Av\|^2$, posto

$$f(t) = \|Av(t)\|^2$$

si ha che

$$f(0) = \max_t f(t)$$

per cui

$$f'(0) = 0$$

Ora

$$f(t) = \langle Ab_i\cos(t) + Ab_j\sin(t), Ab_i\cos(t) + Ab_j\sin(t)\rangle = \\ = \|Ab_i\|^2\cos^2(t) + \|Ab_j\|^2\sin^2(t) + 2\langle Ab_i, Ab_j\rangle\cos(t)\sin(t)$$

e

$$f'(t) = -2\|Ab_i\|^2 \cos(t)\sin(t) + 2\|Ab_j\|^2 \sin(t)\cos(t) + \\ + 2\langle Ab_i, Ab_j\rangle \left(\cos^2(t) - \sin^2(t)\right)$$

per cui

$$0 = f'(0) = 2\langle Ab_i, Ab_j\rangle$$

e

$$Ab_i \perp Ab_j$$

Chiaramente è possibile che $Ab_i = 0$ per qualche i ed anzi ciò si verifica certamente se $n > m$ per $n - m$ indici.

□

Sia ora $A \in \mathcal{M}^{m\times n}$, che rappresenta una trasformazione lineare da $\mathbb{R}^n$ in $\mathbb{R}^m$, consideriamo in $\mathbb{R}^n$ la base $\{b_1, b_2, \ldots, b_n\}$ determinata in precedenza ed i vettori $\{Ab_1, Ab_2, \ldots, Ab_n\}$ in $\mathbb{R}^m$ che risultano a due a due perpendicolari e che possono anche essere nulli.

Poniamo

$$h_i = \begin{cases} \frac{Ab_i}{\|Ab_i\|} & \text{se } Ab_i \neq 0 \\ 0 & \text{se } Ab_i = 0 \end{cases}$$

e $\sigma_i = \|Ab_i\|$ (da cui $Ab_i = \sigma_i h_i$).

Possiamo allora scrivere che

$$A = \begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \begin{pmatrix} b_1^T \\ b_2^T \\ \cdots \\ b_n^T \end{pmatrix} = HDB$$

Essendo H una matrice $m \times n$, D una matrice diagonale $n \times n$, e B una matrice $n \times n$.

Infatti si ha

$$Ab_i = HDBb_i = \begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix} \begin{pmatrix} 0 \\ \cdots \\ \|b_i\|^2 \\ \cdots \\ 0 \end{pmatrix} =$$

$$\begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix} \begin{pmatrix} 0 \\ \cdots \\ \sigma_i \\ \cdots \\ 0 \end{pmatrix} = \sigma_i h_i = Ab_i$$

B è una matrice ortonormale $n \times n$, quanto ad H si tratta di una matrice le cui colonne sono n vettori ortogonali, eventualmente nulli, in $\mathbb{R}^m$; osserviamo anche che nel caso in cui sia $h_i = 0$ si ha, in corrispondenza, $\sigma_i = 0$; inoltre, se $m > n$ possiamo aggiungere colonne nulle ad H e righe nulle a D in modo che

$$A = HDM$$

essendo

- H una matrice $m \times m$
- D una matrice $m \times n$ con elementi tutti nulli tranne che eventualmente sulla diagonale principale
- B una matrice $n \times n$

Infine possiamo considerare solo le colonne $h_i \neq 0$ di H , siano $\{h_1, h_2, \ldots, h_p\}$, trovare $\{\eta_1, \eta_2 \ldots, \eta_{m-p}\}$ in modo che

$$\{h_1, h_2, \ldots, h_p, \eta_1, \eta_2 \ldots, \eta_{m-p}\}$$

sia una base ortonormale di $\mathbb{R}^m$ e ridefinire H come

$$H = \begin{pmatrix} h_1 & h_2 & \ldots & h_p & \eta_1 & \eta_2 & \ldots & \eta_{m-p} \end{pmatrix}$$

Dal momento che le modifiche su H hanno riguardato colonne in corrispondenza delle quali le righe di D sono nulle possiamo allora affermare che la relazione

$$A = HDM$$

vale ancora essendo in più H una matrice ortonormale $m \times m$.

4.8 Autovalori ed Autovettori

4.8.1 Cambio di Base in $\mathbb{R}^n$

Ogni elemento di $\mathbb{R}^n$ può essere espresso in maniera univoca mediante combinazione lineare di elementi di una base; dal momento che possiamo considerare basi diverse è utile conoscere come si possa passare da una all'altra.

Siano pertanto $\{a_1, a_2, \ldots, a_n\}$ e $\{b_1, b_2, \ldots, b_n\}$ due basi in $\mathbb{R}^n$ e consideriamo le matrici A e B le cui colonne sono costituite dai vettori delle due basi:

$$A = (a_1, a_2, \ldots, a_n) \qquad B = (b_1, b_2, \ldots, b_n)$$

Ogni elemento a_i può essere espresso in maniera univoca mediante combinazione lineare degli elementi b_j e viceversa ogni elemento

b_j può essere espresso in maniera univoca mediante combinazione lineare degli elementi a_i, per cui possiamo trovare una matrice P tale che

$$A = PB \qquad \text{e} \qquad B = P^{-1}A$$

Osserviamo esplicitamente che l'invertibilità di P è garantita dal fatto che la rappresentazione di ogni vettore rispetto ad una base assegnata è sempre possibile (surgettività) in maniera univoca (iniettività).

Sia ora $L : \mathbb{R}^n \to \mathbb{R}^n$ una applicazione lineare; abbiamo già visto che si può trovare una matrice $A \in \mathcal{M}^n$, che dipende dalla base scelta in $\mathbb{R}^n$, tale che $L(x) = Ax$.

Consideriamo in $\mathbb{R}^n$ due diverse basi, $\{a_1, a_2, \ldots, a_n\}$ e $\{b_1, b_2, \ldots, b_n\}$, e la matrice P di passaggio tra A e B per modo che se x è espresso mediante i vettori della prima base allora $\xi = Px$ è la sua espressione in termini della seconda. Parimenti si ha $x = P^{-1}\xi$.

Supponiamo inoltre che, rispetto alla prima base, L sia rappresentata dalla matrice A per cui $L(x) = Ax$, mentre rispetto alla seconda base, L sia rappresentata dalla matrice B per cui $L(y) = By$.

Avremo che

$$L(Px) = BPx \qquad \text{ed anche} \qquad P^{-1}L(Px) = P^{-1}BPx$$

Pertanto

$$P^{-1}BPx = Ax \qquad \forall x \in \mathbb{R}^n$$

e

$$P^{-1}BP = A \qquad , \quad P^{-1}AP = B$$

Diremo che A e $P^{-1}AP$ sono matrici simili ed è evidente che matrici simili rappresentano la stessa applicazione lineare rispetto a basi diverse.

Siano ad esempio $\{e_1, e_2\}$ ed $\{\epsilon_1, \epsilon_2\}$ due basi di $\mathbb{R}^2$. È possibile, in maniera univoca, trovare α, β e γ, δ in $\mathbb{R}$ tali che

$$e_1 = \alpha\epsilon_1 + \beta\epsilon_2$$

$$e_2 = \gamma\epsilon_1 + \delta\epsilon_2$$

per cui, scrivendo le componenti:

$$\begin{aligned} e_{11} &= \alpha\epsilon_{11} + \beta\epsilon_{21} \\ e_{12} &= \alpha\epsilon_{12} + \beta\epsilon_{22} \\ e_{21} &= \gamma\epsilon_{11} + \delta\epsilon_{21} \\ e_{22} &= \gamma\epsilon_{12} + \delta\epsilon_{22} \end{aligned}$$

Da cui

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

o, alternativamente,

$$\begin{pmatrix} e_1 & e_2 \end{pmatrix} = \begin{pmatrix} \epsilon_1 & \epsilon_2 \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

4.8.2 *Autovalori ed autovettori*

Definizione 4.7 *Sia $L : \mathbb{C}^n \to \mathbb{C}^n$ una applicazione lineare; diciamo che $\lambda \in \mathbb{C}$ è un autovalore per L se esiste $x \in \mathbb{C}^n$, $x \neq 0$ tale che*

$$L(x) = \lambda x$$

in tal caso x si dice autovettore di L relativo all'autovalore λ.

Teorema 4.11 *Sia $L : \mathbb{C}^n \to \mathbb{C}^n$ una applicazione lineare e sia $A \in \mathcal{M}^n$ la sua matrice di rappresentazione.*

$\lambda \in \mathbb{C}$ è un autovalore per L se e solo se esiste $x \in \mathbb{C}^n$, $x \neq 0$ tale che

$$Ax = \lambda x$$

Poichè trovare $x \in \mathbb{C}^n$, $x \neq 0$ tale che $Ax = \lambda x$ è equivalente a trovare soluzioni non nulle del sistema lineare

$$(A - \lambda I_n)x = 0$$

si può affermare, usando il teorema di Rouchè- Capelli, che:

Teorema 4.12 *λ è un autovalore per A se e solo se*

$$\det(A - \lambda I_n) = 0$$

Pertanto sono autovalori dell'applicazione lineare L o, equivalentemente, di una sua matrice di rappresentazione A le radici del suo polinomio caratteristico p_A che è definito da

$$p_A(\lambda) = \det(A - \lambda I_n)$$

Quanto detto assicura inoltre che matrici simili hanno gli stessi autovalori, come si può anche verificare osservando che se

$$A\xi = \lambda\xi \iff APx = \lambda Px \iff P^{-1}APx = \lambda x$$

Il polinomio caratteristico di una matrice A ha grado n e, per il teorema fondamentale dell'algebra, ammette n radici complesse ciascuna contata con la sua molteplicità.

Definizione 4.8 *Chiamiamo molteplicità algebrica di un autovalore $\bar{\lambda}$ la molteplictà di $\bar{\lambda}$ come radice di p_A.*

Definizione 4.9 *Definiamo autospazio associato ad un autovalore λ, lo spazio vettoriale $E_{\bar{\lambda}}$ delle soluzioni dell'equazione*

$$\det(A - \bar{\lambda} I_n) = 0$$

Definiamo cioè

$$E_{\bar{\lambda}} = \ker(A - \bar{\lambda} I_n)$$

Chiamiamo moltepicità geometrica dell'autovalore $\bar{\lambda}$ la dimensione dello spazio $E_{\bar{\lambda}}$.

Si può dimostrare che

Teorema 4.13 *La molteplicità geometrica di un autovalore e' sempre minore o uguale alla sua molteplicità algebrica essendo possibile la disuguaglianza stretta.*

DIMOSTRAZIONE. Sia $\bar{\lambda}$ un autovalore della matrice A avente molteplicità geometrica r; e siano $v_1, v_2, \ldots, v_r \in E_{\bar{\lambda}}$ i vettori di una base di $E_{\bar{\lambda}}$

Sia $v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{n-r}$ una base di $\mathbb{C}^n$ ottenuta completando la base di $E_{\bar{\lambda}}$.

Ogni $x \in \mathbb{C}^n$ si può esprimere mediante i vettori v_i e w_j mediante la

$$x = \sum_1^r x_i v_i + \sum_{r+1}^n x_j w_j$$

per cui

$$L(x) = \sum_1^r x_i L(v_i) + \sum_{r+1}^n x_j L(w_j) =$$

$$= \left(\begin{array}{cccc|cccc} \bar{\lambda} & 0 & \cdots & 0 & (L(w_{(r+1)})_1 & (L(w_{(r+2)})_1 & \cdots & (L(w_n)_1 \\ 0 & \bar{\lambda} & \cdots & 0 & (L(w_{(r+1)})_2 & (L(w_{(r+2)})_2 & \cdots & (L(w_n)_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \bar{\lambda} & (L(w_{(r+1)})_r & (L(w_{(r+2)})_r & \cdots & L(w_n)_r \\ -- & -- & -- & -- & -- & -- & -- & -- \\ 0 & 0 & \cdots & 0 & (L(w_{(r+1)})_{r+1} & (L(w_{(r+2)})_{r+1} & \cdots & (L(w_n)_{r+1} \\ 0 & 0 & \cdots & 0 & (L(w_{(r+1)})_{r+2} & (L(w_{(r+2)})_{r+2} & \cdots & (L(w_n)_{r+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & (L(w_{(r+1)})_n & (L(w_{(r+2)})_n & \cdots & (L(w_n)_n \end{array}\right) =$$

$$Bx$$

Pertanto B è una matrice di rappresentazione per L e d il suo polinomio caratteristico è del tipo

$$p_B(\lambda) = (\lambda - \bar{\lambda})^r q(\lambda)$$

dove q è un polinomio di grado $n - r$. Ne segue che la molteplicità di $\bar{\lambda}$ è almeno r. □

Possiamo dimostrare che

Teorema 4.14 *Autovettori relativi ad autovalori distinti sono tra loro indipendenti.*

DIMOSTRAZIONE. Siano λ_1, λ_2 autovalori distinti di A, siano v_1, v_2 autovettori rispettivamente in $E_{\lambda_1}, E_{\lambda_2}$ e sia

$$\alpha v_1 + \beta v_2 = 0$$

Avremo, moltiplicando per λ_1

$$\lambda_1 \alpha v_1 + \lambda_1 \beta v_2 = 0$$

ed applicando A

$$A(\alpha v_1 + \beta v_2) = \lambda_1 \alpha v_1 + \lambda_2 \beta v_2 = 0$$

Ne segue che

$$(\lambda_1 - \lambda_1)\alpha v_1 + (\lambda_2 - \lambda_1)\beta v_2 = (\lambda_2 - \lambda_1)\beta v_2 = 0$$

e, visto che $(\lambda_2 - \lambda_1) \neq 0$ si ha $\beta = 0$.

Inoltre da $\alpha v_1 + \beta v_2 = 0$ segue $\alpha v_1 = 0$ e $\alpha = 0$.

Supponiamo ora che $\lambda_1, \lambda_2, \dots \lambda_m$ siano autovalori distinti di A, che $v_1, v_2, \dots v_m$ siano autovettori rispettivamente in $E_{\lambda_1}, E_{\lambda_2}, \dots E_{\lambda_m}$ e che $v_1, v_2 \cdots v_{m-1}$ siano linearmente indipendenti; allora se

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0 \tag{4.7}$$

moltiplicando per λ_m,

$$\lambda_m \alpha_1 v_1 + \lambda_m \alpha_2 v_2 + \cdots + \lambda_m \alpha_{m-1} v_{m-1} + \lambda_m \alpha_m v_m = 0$$

mentre applicando A

$$A(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m) = \lambda_1 \alpha_1 v_1 + \lambda_2 \alpha_2 v_2 + \cdots + \lambda_{m-1} \alpha_{m-1} v_{m-1} + \lambda_m \alpha_m v_m = 0$$

Ne segue che

$$(\lambda_1 - \lambda_m)\alpha_1 v_1 + (\lambda_2 - \lambda_m)\alpha_2 v_2 + \cdots + (\lambda_{m-1} - \lambda_m)\alpha_{m-1} v_{m-1} = 0$$

Da cui, essendo $v_1, v_2 + \cdots v_{m-1}$ linearmente indipendenti, segue

$$\alpha_1 = \alpha_2 = \cdots = \alpha_{m-1} = 0$$

Sostituendo in 4.7 si ottiene $\alpha_m = 0$ e la lineare indipendenza di $v_1, v_2 + \cdots v_{m-1}, v_m$

□

Corollario 4.5 *Se $\lambda_1, \lambda_2, \dots \lambda_m$ sono gli autovalori distinti di una matrice $A \in \mathcal{M}^n$, allora gli spazi $E_{\lambda_1}, E_{\lambda_2}, \dots E_{\lambda_m}$ sono indipendenti, cioè se*

$$v_1 + v_2 + \cdots + v_m = 0 \quad v_i \in E_{\lambda_i} \qquad \Longrightarrow \qquad v_i = 0 \quad \forall i$$

Ne segue che

In particolare, nel caso in cui la matrice A sia simmetrica si ha

Teorema 4.15 *Se $A \in \mathcal{M}^n$ è una matrice simmetrica autovettori relativi ad autovalori distinti sono tra loro ortogonali.*

DIMOSTRAZIONE. Siano λ e μ due autovalori distinti della matrice A e siano x ed y due autovettori ad essi rispettivamente relativi.

Poichè A è simmetrica avremo

$$\mu\langle x, y\rangle = \langle x, Ay\rangle = \langle Ax, y\rangle = \lambda\langle x, y\rangle$$

da cui, essendo $\lambda \neq \mu$ segue $\langle x, y\rangle = 0$. □

Sempre nel caso di una matrice simmetrica possiamo anche affermare che

Teorema 4.16 *Se $A \in \mathcal{M}^n$ è una matrice simmetrica a coefficienti reali, tutti i suoi autovalori sono reali.*

DIMOSTRAZIONE. Sia λ un autovalore di A. Avremo

$$Ax = \lambda x \qquad \text{e} \qquad \langle x, Ax\rangle = \lambda\|x\|^2$$

Tenendo conto che il coniugato del prodotto è il prodotto dei coniugati,

$$\bar{\lambda}\bar{x} = \bar{A}\bar{x} = A\bar{x} \qquad \text{e} \qquad \langle \bar{x}, A\bar{x}\rangle = \bar{\lambda}\|\bar{x}\|^2 = \bar{\lambda}\|x\|^2$$

Poichè $x \neq 0$ ne segue che $\lambda = \bar{\lambda}$ e che $\lambda \in \mathbb{R}$.

□

Definizione 4.10 *Siano V e U spazi vettoriali. Definiamo la somma*

$$W = V + U$$

mediante la

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

Diciamo che la somma è diretta e scriviamo

$$W = V \oplus U$$

se ogni elemento si W può essere espresso come somma di un elemento di U e di uno di V in maniera unica.

In altre parole la somma $U + V$ è diretta se da $v_1 + u_1 = v_2 + u_2$, $v_1, v_2 \in V$, $u_1, u_2 \in U$ si può dedurre che $v_1 = v_2$ e $u_1 = u_2$ oppure, equivalentemente, se da $v + u = 0$, $v \in V$, $u \in U$ si può dedurre che $v = u = 0$.

Teorema 4.17 *Sia $X = U \oplus V \oplus W$ uno spazio vettoriale di dimensione N somma diretta degli spazi U di dimensione p, V di dimensione q e W di dimensione r. Allora*

$$p + q + r = N$$

DIMOSTRAZIONE. Se infatti A, B, C sono, rispettivamente, basi di U,V, W allora $D = A \cup B \cup C$ è una base per X..

□

Il teorema 4.5 può quindi essere riformulato come segue

Teorema 4.18 *Se $\lambda_1, \lambda_2, \ldots \lambda_m$ sono gli autovalori distinti di una applicazione lineare L, allora*

$$E_{\lambda_1} + E_{\lambda_m} + \cdots + E_{\lambda_m} = E_{\lambda_1} \oplus E_{\lambda_m} \oplus \cdots \oplus E_{\lambda_m}$$

Possiamo dimostrare che

Teorema 4.19 *Sia $L : \mathbb{C}^n \to \mathbb{C}^n$ una applicazione lineare; allora*

$$\mathbb{C}^n = \ker(L) \oplus R(L)$$

DIMOSTRAZIONE. Sia $\{v_1, v_2, \ldots, v_r\}$ una base per $\ker(L)$ e sia

$$\{v_1, v_2, \ldots, v_r, w_{r+1}, w_{r+2}, \ldots, w_n\}$$

una base di $\mathbb{C}^n$.

Se $u \in R(L)$ possiamo trovare $v \in R^n$ tale che

$$u = L(v) = L(\sum a_i v_i + \sum b_j w_j) = \sum b_j L(w_j)$$

non appena si ricordi che $v_i \in \ker(L)$.

Pertanto $B = \{L(w_{r+1}), L(w_{r+2}), \ldots, L(w_n)\}$ è un insieme di generatori di $R(L)$.

Si verifica anche che B e' un insieme di vettori linearmente indipendenti, infatti, se $\sum b_j L(w_j) = 0$, si ha

$$L(\sum b_j w_j) = \sum b_j L(w_j) = 0$$

e

$$\sum_j b_j w_j \in \ker(L)$$

Pertanto è possibile affermare che

$$\sum_j b_j w_j = \sum_i a_i v_i$$

e

$$\sum b_j w_j - \sum_i a_i v_i = 0$$

che è una combinazione lineare nulla di elementi di una base di $\mathbb{C}^n$, per cui $a_i = b_j = 0$.

Pertanto B è una base per $R(L)$.

Sia $\hat{L} = L|_{R(L)}$; Verifichiamo che $\hat{L}$ è iniettiva.

Infatti se $u \in R(L)$ è tale che $L(u) = 0$, si ha

$$0 = \hat{L}(u) = L(u) = \sum b_j L(w_j) = L(\sum b_j w_j)$$

Quindi $\sum b_j w_j \in \ker(L)$ e si ha

$$\sum_j b_j w_j = \sum_i a_i v_i \qquad \text{da cui} \qquad \sum b_j w_j - \sum_i a_i v_i = 0$$

Possiamo concludere che $\hat{L} : R(L) \to R(L)$ è anche surgettiva.

Sia ora $v \in \mathbb{C}^n$, dal momento che B è una base di $R(L)$ si ha

$$L(v) = \sum b_j L(w_j) = L(\sum b_j w_j) = L(w)$$

quindi

$$w = \hat{L}^{-1}(\hat{L}(v)) \in R(L)$$

e, posto

$$u = v - w = v - \sum b_j w_j$$

avremo

$$L(u) = L(v - \sum b_j w_j) = L(v) - \sum b_j L(w_j) = 0$$

Quindi $u \in \ker(L)$ e $v = u + w$ con $u \in \ker(L)$ e $w \in R(L)$.

Concludiamo che $\mathbb{C}^n = \ker(L) + R(L)$, inoltre la somma è diretta in quanto se $u + w = 0$ con $u \in \ker(L)$ e $w \in R(L)$ si ha

$$\sum a_i v_i + \sum b_j w_j = 0$$

e $a_i = b_j = 0$ □

Sia λ un autovalore di $L : \mathbb{C}^n \to \mathbb{C}^n$ allora $\ker(L - \lambda I) \neq \varnothing$ cioè esiste $v \neq 0$ tale che

$$(L - \lambda I)(v) = L(v) - \lambda v = 0$$

In tal caso abbiamo anche che

$$(L - \lambda I)^2(v) = (L - \lambda I)(L - \lambda I)(v) = (L - \lambda I)(0) = 0$$

e quindi

$$v \in \ker(L - \lambda I)^2$$

Ne segue che, posto

$$N_\lambda^k = \ker(L - \lambda I)^k$$

si ha

$$N_\lambda^1 \subset N_\lambda^2 \subset \cdots \subset N_\lambda^k \subset \cdots$$

Teorema 4.20 *Se $v \neq 0$ e $v \in N_\lambda^k \setminus N_\lambda^{k-1}$ allora $k \leq n$.*

DIMOSTRAZIONE. Poichè $v \notin N_\lambda^{k-1}$ avremo che

$$v, (L - \lambda I)(v), (L - \lambda I)^2(v), \ldots, (L - \lambda I)^{k-1}(v)$$

sono tutti non nulli mentre

$$(L - \lambda I)^h(v) = 0 \quad \forall h \geq k$$

Consideriamo la combinazione lineare nulla

$$\alpha_0 v + \alpha_1 (L - \lambda I)(v) + \alpha_2 (L - \lambda I)^2(v) + \cdots + \alpha_{k-1}(L - \lambda I)^{k-1}(v) = 0$$

Applicando $(L - \lambda I)^{k-1}$ ad entrambi i membri otteniamo

$$\alpha_0 (L - \lambda I)^{k-1}(v) = 0$$

da cui $\alpha_0 = 0$

Similmente applicando $(L - \lambda I)^{k-2}$ ad entrambi i membri otteniamo

$$\alpha_1 (L - \lambda I)^{k-1}(v) = 0$$

da cui $\alpha_1 = 0$ e, procedendo, $\alpha_j = 0$ per ogni $j = 1, \ldots, k-1$

Pertanto i k vettori considerati sono linearmente indipendenti e ciò garantisce che $k \leq n$ □

Definizione 4.11 *Sia λ un autovalore di L; definiamo autospazio generalizzato l'insieme $\hat{N}_\lambda$ dei vettori $v \in \mathbb{C}^n$ per i quali esiste $k \in \mathbb{N}$, $k \leq n$ con $v \in N_\lambda^k$. Quanto abbiamo detto permette di affermare che*

$$\hat{N}_\lambda = \bigcup_h N_\lambda^h = N_\lambda^n = \ker((L - \lambda I)^n)$$

Teorema 4.21 *L trasforma ogni $\hat{N}_\lambda$ in se' stesso.*

DIMOSTRAZIONE. Se $x \in \hat{N}_\lambda$ allora esiste k tale che $(L - \lambda I)^k x = 0$. Si ha

$$0 = L\left((L - \lambda I)^k x\right) = L \sum_0^k \binom{k}{j} L^j I^{(k-j)} x =$$

$$= \sum_0^k \binom{k}{j} L^{j+1} x = \sum_0^k \binom{k}{j} L^j L x =$$

$$= (L - \lambda I)^k L(x)$$

e quindi $L(x) \in \hat{N}_\lambda$ □

Teorema 4.22 *Sia $L : \mathbb{C}^n \to \mathbb{C}^n$ un'applicazione lineare allora*

$$R^n = \ker((L - \lambda I)^n) \oplus R((L - \lambda I)^n)$$

DIMOSTRAZIONE.

Segue dal teorema 4.19 applicato a $(L - \lambda I)^n$.

□

Consideriamo ora due autovalori di L η e μ, distinti, e consideriamo l'applicazione lineare $(L - \mu I)$ ristretta allo spazio $\hat{N}_\eta$

$$(L - \mu I) : \hat{N}_\eta \to \hat{N}_\eta$$

$(L - \mu I)$ ristretta a $\hat{N}_\eta$ risulta iniettiva.

Se, infatti, non lo fosse potremmo trovare $x \in \hat{N}_\eta$, $x \neq 0$ tale che

$$(L - \mu I)(x) = 0$$

e si potrebbe affermare che

$$x \neq 0 \ , \ L(x) = \mu x$$

ed inoltre esiste $m \in \mathbb{N}$ tale che $(L - \eta I)^m(x) = 0$.

Ne segue che

$$(L - \eta I)(x) = L(x) - \eta x = \mu x - \eta x = (\mu - \eta)x$$

$$(L - \eta I)^2(x) = (L - \eta I)(\mu - \eta)x = L((\mu - \eta)x) - \eta(\mu - \eta)x = \\ = \mu(\mu - \eta)x - \eta(\mu - \eta)x = (\mu - \eta)^2 x$$

Fino a

$$(L - \eta I)^m(x) = (\mu - \eta)^m x$$

Poichè $(\mu - \eta)^m x \neq 0$ mentre $(L - \eta I)^m(x) = 0$, si trova una contraddizione.

$(L - \mu I)$ ristretta a $\hat{N}_\eta$ è iniettiva e , di conseguenza, anche surgettiva.

Pertanto

$$(L - \mu I)\hat{N}_\eta = \hat{N}_\eta$$

e iterando

$$(L - \mu I)^m \hat{N}_\eta = \hat{N}_\eta \qquad \forall m \in \mathbb{N}$$

da cui segue che

$$R\left((L - \mu I)^n\right) \supset \hat{N}_\eta$$

Consideriamo allora un'applicazione lineare

$$L : \mathbb{C}^n \to \mathbb{C}^n$$

e siano $\lambda_1, \lambda_2, \ldots, \lambda_p$ i suoi autovalori distinti, ciascuno con molteplicità algebrica $m_1, m_2, \ldots, m_p$.

Per il teorema 4.22 possiamo affermare che

$$R^n = \ker((L - \lambda_1 I)^n) \oplus R((L - \lambda_1 I)^n) = \hat{N}_{\lambda_1} \oplus R((L - \lambda_1 I)^n)$$

ed inoltre abbiamo visto che

$$R(L - \lambda_1 I)^n) \supset \hat{N}_\lambda \qquad \forall \lambda = \lambda_2, \lambda_3, \ldots \lambda_p$$

Possiamo allora considerare la restrizione L_1 di L a $R(L - \lambda_1 I)^n)$

$$L_1 : R(L - \lambda_1 I)^n) \to R(L - \lambda_1 I)^n)$$

ed osservare che gli autovalori di L_1 sono $\lambda_2, \lambda_3, \ldots \lambda_p$ in modo da poter decomporre $R(L - \lambda_1 I)^n$ come

$$R(L - \lambda_1 I)^n = \hat{N}_{\lambda_2} \oplus R(L_1 - \lambda_2 I)^n$$

ed ottenere

$$R^n = \hat{N}_{\lambda_1} \oplus R((L - \lambda_1 I)^n) = \hat{N}_{\lambda_1} \oplus \hat{N}_{\lambda_2} \oplus R((L_1 - \lambda_2 I)^n)$$

Iterando si perviene ad una decomposizione del tipo

$$R^n = \hat{N}_{\lambda_1} \oplus \hat{N}_{\lambda_2} \oplus \cdots \oplus \hat{N}_{\lambda_p} \oplus R((L_{p-1} - \lambda_p I)^n)$$

e, dal momento che

- L_{p-1} puo' avere solo autovalori diversi da $\lambda_1, \lambda_2, \ldots, \lambda_p$;
- se $\dim(R((L_{p-1} - \lambda_p I)^n)) > 0$ deve avere almeno un autovalore.

Possiamo concludere che $R((L_{p-1} - \lambda_p I)^n) = \{0\}$, che

$$R^n = \hat{N}_{\lambda_1} \oplus \hat{N}_{\lambda_2} \oplus \cdots \oplus \hat{N}_{\lambda_p}$$

e che $\dim(\hat{N}_{\lambda_k}) = m_k$ in quanto certamente si ha $\dim(\hat{N}_{\lambda_k}) \leq m_k$ e $\sum \dim(\hat{N}_{\lambda_k}) = \sum m_k = n$.

Pertanto possiamo rappresentare L mediante una matrice del tipo:

$$J = \begin{pmatrix} \boxed{A_{\lambda_1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{A_{\lambda_2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{A_{\lambda_3}} & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \ddots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \boxed{A_{\lambda_{p-1}}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{A_{\lambda_p}} \end{pmatrix}$$

dove ogni blocco A_{λ_k} rappresenta L ristretta a $\hat{N}_{\lambda_k}$, ha dimensione m_k ed è relativo all'autovalore λ_k.

Per definire J sará pertanto sufficiente definire ogni singolo blocco; a questo scopo osserviamo che per ogni vettore non nullo $v \in \hat{N}_\lambda$ esiste $k \leq n$ tale che

$$(L - \lambda I)^k v = 0 \qquad \text{e} \qquad (L - \lambda I)^{k-1} v \neq 0$$

Osserviamo altresì che ne segue

$$(L - \lambda I)^h v \neq 0 \qquad \forall h = 1..k-1$$

Se $(L - \lambda I)^{\hat{h}} v = 0$ allora $(L - \lambda I)^h v = 0$, $\forall h \geq \hat{h}$

Sia pertanto $v_k \in \hat{N}_\lambda$ con $(L - \lambda I)^k v = 0$ e $(L - \lambda I)^h v \neq 0$ per ogni $h = 1, \ldots, k-1$

Posto

$$\begin{aligned} v_{k-1} &= (L-\lambda I)v_k \\ v_{k-2} &= (L-\lambda I)v_{k-1} \\ v_{k-3} &= (L-\lambda I)v_{k-2} \\ &\cdots\cdots\cdots\cdots \\ v_1 &= (L-\lambda I)v_2 \\ (L-\lambda I)v_1 &= 0 \end{aligned} \qquad (4.8)$$

$$\begin{aligned} v_{k-1} &= (L-\lambda I)v_k \\ v_{k-2} &= (L-\lambda I)^2 v_k \\ v_{k-3} &= (L-\lambda I)^3 v_k \\ &\cdots\cdots\cdots\cdots \\ v_1 &= (L-\lambda I)^{k-1} v_k \\ (L-\lambda I)v_1 &= (L-\lambda I)^k v_k = 0 \end{aligned}$$

E, per quanto visto in precedenza, i vettori $v_{k-1}, v_{k-2}, \ldots, v_1, v_0$ sono linearmente indipendenti e v_{k-1} è un autovettore di L relativo all'autovalore λ .

Possiamo quindi affermare che una base per $\hat{N}_\lambda$ è costituita da uno o più serie di vettori che si ottengono partendo da un autovettore di L, risolvendo le 4.8

Le 4.8 possono essere riscritte come:

$$\begin{aligned} L(v_1) &= \lambda v_1 \\ L(v_2) &= v_1 + \lambda v_2 \\ L(v_3) &= v_2 + \lambda v_3 \\ &\cdots\cdots\cdots\cdots \\ L(v_k) &= v_{k-1} + \lambda v_k \end{aligned}$$

e quindi, considerando la matrice

$$V = \begin{pmatrix} v_1 & v_2 & v_3 & \cdots & v_k \end{pmatrix}$$

se A è la matrice di rappresentazione di L , ($L(v) = Av$), si ha

$$AV = V \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{pmatrix} = AJ$$

Poichè V ha per colonne vettori linearmente indipendenti, è invertibile e si ha

$$A = VJV^{-1}$$

per cui J è una matrice equivalente ad A.

Naturalmente servono diversi blocchi del tipo di quello appena costruito per rappresentare L ristretta ad $\hat{N}_\lambda$ ottenuta tale rappresentazione si ripete il lavoro per ogni autovalore e si costruisce J.

5. Geometria Analitica nel piano

L'idea di adottare un sistema di riferimento per introdurre il formalismo algebrico nella trattazione di problemi geometrici risale al lavoro contemporaneo ed indipendente di Pierre de Fermat (1601-1665) e di René Descartes (1596-1650) e si colloca tra il 1629 anno in cui Fermat cominciò a lavorare in tal senso fino al 1679 anno in cui il suo lavoro fu pubblicato postumo. Fermat secondo gli storici fu quindi il primo ad occuparsi della questione studiando problemi di massimo e minimo ma non comunicò ad altri il suo lavoro fino al 1636; nel frattempo Descartes sviluppò i suoi studi e pubblicò nel 1637 il suo Discours de la Methode, in un'appendice del quale intitolata La Géométrie, è contenuta una applicazione dell'algebra alla geometria.

Per introdurre un sistema di riferimento nel piano occorre considerare due rette orientate, perpendicolari dette assi, che si intersecano in un punto detto origine. Per retta orientata si intende una retta su cui è fissato un punto che la divide in due semirette una delle quali è definita come la semiretta degli elementi positivi. Le due rette che costituiscono il riferimento sono orientate scegliendo tale punto coincidente con il punto di intersezione, cioè l'origine.

È d'uso scegliere un sistema di riferimento, che si definisce destrorso, in cui per sovrapporre la prima semiretta positiva alla seconda occorre ruotare in senso antiorario; in questo modo si definisce anche un verso positivo per gli angoli : diremo che si ruota di un angolo positivo nel caso si compia una rotazione in senso antiorario, negativo in caso contrario.

Su ciascuna retta sono fissati due punti che corrispondono ad 1 e definiscono l'unità di misura. Possiamo in tal modo individuare su ognuna delle due rette un punto da associare ad ogni numero reale, usando criteri di proporzionalità.

Un punto P nel piano è individuato univocamente da una coppia ordinata di numeri reali (x, y) dove x è la proiezione di P sul primo asse, che solitamente è rappresentato orizzontale e viene chiamato asse delle ascisse, mentre y è la proiezione di P sul secondo asse, che solitamente è rappresentato verticale e si chiama asse delle ordinate.

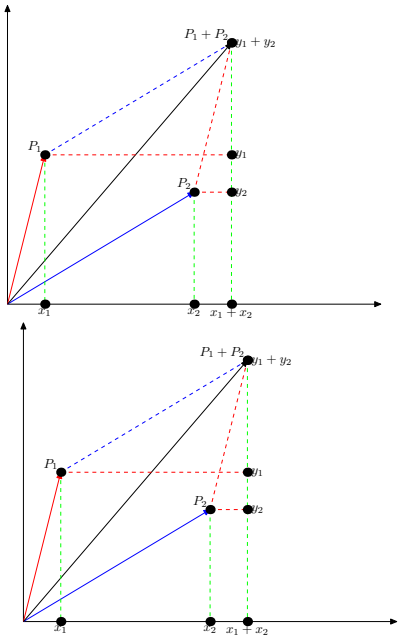
È naturale introdurre due operazioni: la somma di due elementi

$P_1 = (x_1, y_1)$ e $P_2 = (x_2, , y_2)$ definita da

$$P_1 + P_2 = (x_1 + x_2, y_1 + y_2)$$

ed il prodotto di un elemento $P = (x, y)$ per uno scalare $\alpha \in \mathbb{R}$ definito da

$$\alpha P = (\alpha x, \alpha y)$$



Tali operazioni sono interpretabili graficamente nel piano mediante la regola del parallelogrammo, per quanto riguarda la somma, e la modifica della lunghezza del segmento OP per quel che riguarda il prodotto per uno scalare.

Più precisamente se sommiamo mediante la regola del parallelogramma i segmenti orientati OP_1 ed OP_2 otterremo il segmento orientato $O(P_1 + P_2)$ mentre se moltiplichiamo P per α, il segmento orientato $O(\alpha P)$ giace sulla stessa semiretta sui cui giace P a distanza dall'origine pari alla distanza dall'origine di P moltiplicata per α.

Il piano dotato di un tale sistema di riferimento viene solitamente indicato come Piano Cartesiano ed è evidentemente un modello dello spazio vettoriale $\mathbb{R}^2$ in cui è naturale considerare come base l'insieme

$$B = \{e_1 = (1, 0), e_2 = (0, 1)\}$$

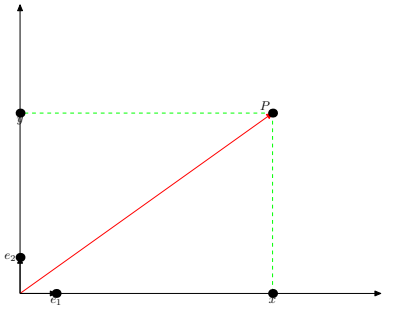
Che B sia una base per R^2 è evidente dal fatto che

$$(x, y) = x(1, 0) + y(0, 1) = xe_1 + ye_2$$

essendo tale decomposizione univoca.

È immediato verificare, usando il teorema di Pitagora, che la distanza di un punto $P = (x, y)$ dall'origine $O = (0, 0)$ è data da

$$\rho = d(P, O) = \sqrt{x^2 + y^2}$$



mentre la distanza tra due punti $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ è

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Si può dotare $\mathbb{R}^2$ di una norma definendo

$$\|P\| = \|(x, y)\| = d(P, O) = \sqrt{x^2 + y^2}$$

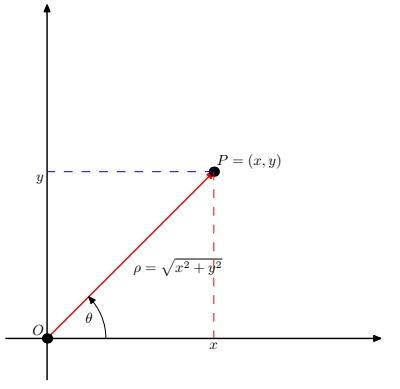
ed inoltre possiamo definire un prodotto scalare mediante la

$$\langle P_1, P_2 \rangle = x_1 x_2 + y_1 y_2$$

e possiamo verificare che

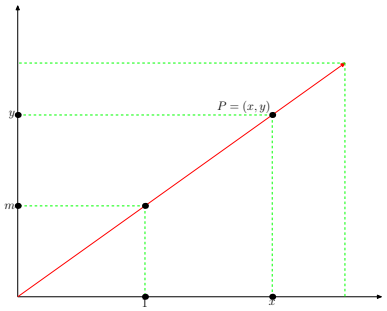
$$\langle P_1, P_2 \rangle = \|P_1\| \|P_2\| \cos(\theta)$$

essendo θ l'angolo compreso tra $\overline{OP_1}$ e $\overline{OP_2}$.



Nel piano cartesiano possono essere individuate mediante relazioni algebriche tra le variabili x ed y diverse figure geometriche e possono essere studiate le loro interazioni.

5.1 *Equazione della Retta*



Una retta nel piano può essere rappresentata a partire dalle proprietà che la caratterizzano.

Cominciamo con il considerare una retta r passante per l'origine. Sia $(1, m)$ il punto di r avente ascissa 1

Un punto $P = (x, y)$ appartiene alla retta se e solo se

$$\frac{y}{x} = \frac{m}{1} \text{ cioè se } y = mx$$

pertanto possiamo affermare che il luogo dei punti del piano cartesiano tali che $y = mx$ costituisce una retta di pendenza m passante per l'origine.

Ovviamente sommando ad mx una quantità fissa n si ottengono i punti di una retta generica, possiamo quindi dire che

$$y = mx + n$$

è l'equazione di una retta nel piano cartesiano.

Dalla precedente equazione ricaviamo che, se (x_0, y_0) è un punto della retta, si ha

$$y_0 = mx_0 + n \text{ da cui } y = y_0 + m(x - x_0)$$

dove sono in evidenza il punto (x_0, y_0) ed il coefficiente m. Si vede quindi come una retta sia individuata da un suo punto e dalla sua pendenza.

Possiamo d'altro canto osservare che si descrive parimenti una retta considerando un punto del piano P_0 che possiamo identificare con un vettore (x_0, y_0) ed una direzione D che possiamo identificare a sua volta con un vettore (α, β).

In questo contesto i punti della retta che passa per il punto P_0 ed è parallela alla direzione D sono identificati, al variare di $t \in \mathbb{R}$, come

$$P = P_0 + tD$$

Se, come al solito, $P = (x, y)$, $P_0 = (x_0, y_0)$,$D = (\alpha, \beta)$ avremo che

$$(x, y) = (x_0, y_0) + t(\alpha, \beta) \text{ ovvero } \begin{cases} x = x_0 + t\alpha \\ y = t_0 + t\beta \end{cases}$$

In questo modo individuiamo la retta per P_0 parallela alla direzione D mediante quelle che si chiamano equazioni parametriche.

Si vede che questa rappresentazione consente di identificare anche l'asse delle y ponendo $\alpha = 0$; inoltre ricavando t ed uguagliando si ottiene

$$\frac{x - x_0}{\alpha} = \frac{y - y_0}{\beta} \text{ da cui } y = y_0 + \frac{\beta}{\alpha}(x - x_0)$$

Le equazioni parametriche mettono in evidenza il modo con cui il punto P si muove sulla retta infatti se consideriamo i punti $P(t_1) = (x(t_1), y(t_1))$ e $P(t_2) = (x(t_2), y(t_2))$ vediamo che

$$\frac{P(t_2) - P(t_1)}{t_2 - t_1} = \left(\frac{x(t_2) - x(t_1)}{t_2 - t_1}, \frac{y(t_2) - y(t_1)}{t_2 - t_1} \right) = (\alpha, \beta)$$

Quindi il vettore $D = (\alpha, \beta)$ identifica, la velocità con cui $P(t)$ si muove sulla retta.

È spesso comodo supporre che il vettore direzione D sia di norma 1 infatti, in tal caso la lunghezza dello spostamento del punto $P(t)$ sulla retta non è influenzato dal vettore scelto.

Un modo standard di considerare vettori unitari nel piano consiste nello scegliere

$$D = (\cos\theta, \sin\theta)$$

infatti

$$\sin^2\theta + \cos^2\theta = 1$$

e viceversa se $|D|^2 = \alpha^2 + \beta^2 = 1$ è possibile trovare $\theta \in [0, 2\pi]$ in modo che $\alpha = \cos\theta$ e $\beta = \sin\theta$ semplicemente definendo *theta* come l'angolo che la semiretta per D e l'origine forma con il semiasse positivo delle x.

È utile osservare che, posto $N = (a, b)$ e $P = (x, y)$

$$ax + by = \langle (a, b), (x, y) \rangle = \langle N, p \rangle = 0$$

e quindi possiamo caratterizzare la retta come il luogo dei punti del piano $\mathbb{R}^2$ rappresentati dai vettori che sono perpendicolari la vettore $N = (a, b)$, che quindi possiamo chiamare vettore normale alla retta.

Evidentemente N caratterizza la direzione della retta e quindi se definiamo $N^\perp$ in modo che $\langle N, N^\perp \rangle = 0$, possiamo scrivere l'equazione della retta perpendicolare passante per l'origine nella forma

$$\langle N^\perp, P \rangle = 0$$

e quella di una generica retta perpendicolare nella forma

$$\langle N^\perp, P \rangle = c$$

Poichè, ad esempio, possiamo scegliere $N^\perp = (b, -a)$ le equazioni della retta perpendicolare è:

$$bx - ay = c$$

c può essere determinato imponendo che la retta passi per un punto (x_0, y_0), chiedendo cioè che

$$bx_0 - ay_0 = c$$

ed in tal caso la retta diventa

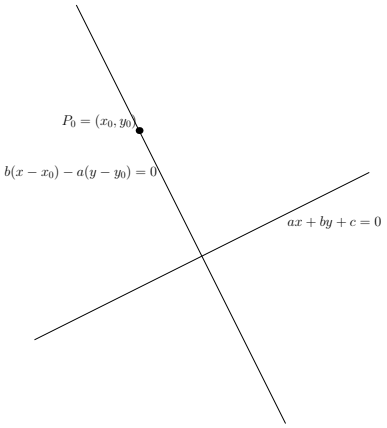
$$b(x - x_0) - a(y - y_0) = 0$$

che rappresenta quindi la retta perpendicolare alla retta assegnata, passante per (x_0, y_0).

5.1.1 *Distanza di un punto da una retta*

Se intersechiamo la retta data $ax + by = c$ con la retta ad essa perpendicolare passante per il punto (x_0, y_0), (la cui equazione è $b(x - x_0) - a(y - y_0) = 0$), cioè se cerchiamo la soluzione del sistema

$$\begin{cases} ax + by + c = 0 \\ b(x - x_0) - a(y - y_0) = 0 \end{cases}$$



otteniamo il punto della retta $ax + by = c$ avente distanza minima dal punto(x_0, y_0).

Possiamo scrivere che

$$\begin{cases} a(x - x_0) + b(y - y_0) = -ax_0 - by_0 - c \\ b(x - x_0) - a(y - y_0) = 0 \end{cases}$$

ed anche

$$\begin{cases} a^2(x - x_0) + ab(y - y_0) = -a(c + ax_0 + by_0) \\ b^2(x - x_0) - ab(y - y_0) = 0 \end{cases} \quad \begin{cases} ab(x - x_0) + b^2(y - y_0) = -b(c + ax_0 + by_0) \\ ab(x - x_0) - a^2(y - y_0) = 0 \end{cases}$$

da cui, sommando e sottraendo membro a membro,

$$\begin{cases} x - x_0 = -a\frac{ax_0 + by_0 + c}{a^2 + b^2} \\ y - y_0 = -b\frac{ax_0 + by_0 + c}{a^2 + b^2} \end{cases}$$

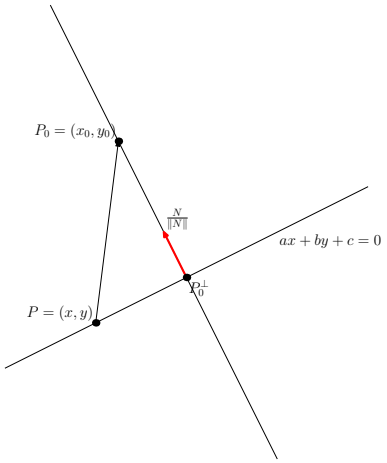
e quindi la distanza tra il punto P_0 e la retta $ax + by + c = 0$ si calcola come

$$(x - x_0)^2 + (y - y_0)^2 = \frac{(ax_0 + by_0 + c)^2}{a^2 + b^2}$$

Pertanto

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

Possiamo anche determinare la distanza di P_0 dalla retta osservando che essa è pari alla lunghezza della proiezione del vettore $P_0 - P$, essendo $P = (x, y)$ un qualunque punto sulla retta, sul versore normale alla retta stessa.



Sia ha

$$P_0 - P^\perp = \left\langle P_0 - P, \frac{N}{\|N\|} \right\rangle \frac{N}{\|N\|}$$

per cui, essendo $ax + by = -c$ dal momento che P è un punto della retta,

$$\begin{aligned}\|P_0 - P^\perp\| &= \left|\left\langle P_0 - P, \frac{N}{\|N\|}\right\rangle\right| = \left|\left\langle P_0 - P, \frac{(a,b)}{\sqrt{a^2+b^2}}\right\rangle\right| = \\ &= \left|\left\langle (x_0,y_0), \frac{(a,b)}{\sqrt{a^2+b^2}}\right\rangle - \left\langle (x,y), \frac{(a,b)}{\sqrt{a^2+b^2}}\right\rangle\right| = \\ &= \left|\frac{ax_0+by_0-ax-ay}{\sqrt{a^2+b^2}}\right| = \frac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}\end{aligned}$$

5.1.2 *Angolo tra due rette*

Siano r ed s due rette nel piano

$$r: \quad ax+by+c=0 \qquad \text{e} \qquad s: \quad \alpha x+\beta y+\gamma = 0$$

Nel caso r ed s siano parallele tali risultano anche i vettori ad esse normali che sono, rispettivamente (a,b) e (α,β); deve pertanto risultare

$$(a,b) = k(\alpha,\beta)$$

cioè i vettori normali devono essere proporzionali.

Nel caso r ed s siano incidenti in un punto (x_0,y_0) possiamo riscrivere le equazioni delle rette come

$$a(x-x_0)+b(y-y_0)=0 \quad \text{e} \quad \alpha(x-x_0)+\beta(y-y_0)=0$$

In tal caso r ed s formano due angoli θ e $\pi-\theta$ che sono uguali agli angoli formato dai rispettivi vettori normali e dai loro opposti.

Poichè sappiamo che

$$\langle (a,b),(\alpha,\beta)\rangle = \|(a,b)\|\ \|(\alpha,\beta)\|\cos(\theta)$$

siamo anche in grado di determinare gli angoli θ e $\pi-\theta$ formati dalla due rette.

Osserviamo inoltre che l'angolo tra i due vettori, in quanto orientati, è univocamente determinato mentre due rette formano nel piano due angoli, tra loro supplementari.

5.1.3 *Circonferenza e cerchio*

Nel piano una circonferenza è definita come il luogo dei punti equidistanti da un punto assegnato, detto centro. La distanza comune di tutti i punti di una circonferenza dal suo centro è detta raggio.

Pertanto, se $P_0 = (x_0,y_0) \in \mathbb{R}^2$ e d $r \in \mathbb{R}_+$ la circonferenza di centro P_0 e raggio r è individuata come l'insieme

$$C = \{(x,y) \in \mathbb{R}^2 \ : \ (x-x_0)^2+(y-y_0)^2 = r^2\}$$

Diciamo quindi, intendendo quanto sopra, che

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$

è l'equazione di una circonferenza di centro P_0 e raggio r.

È immediato verificare che l'equazione della circonferenza può essere scritta nella forma

$$x^2 + y^2 + ax + by + c = 0$$

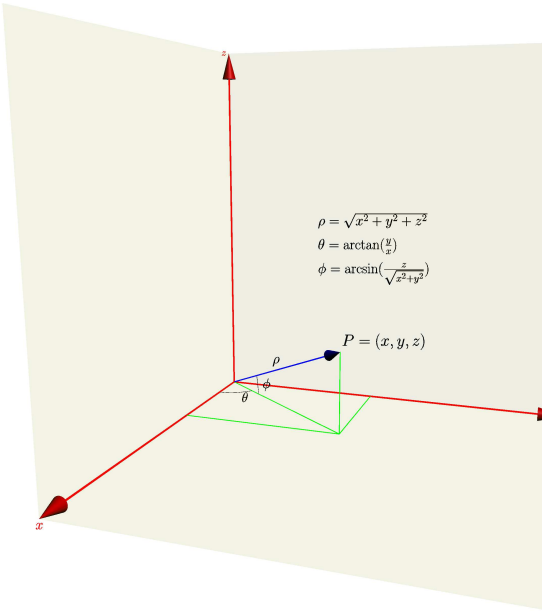
ed un confronto tra le sue equazioni permette di affermare che il suo centro è nel punto $(-\frac{a}{2}, -\frac{b}{2})$ e che il suo raggio è $r = \sqrt{c - \frac{a^2}{4} - \frac{b^2}{4}}$.

In maniera del tutto simile definiamo cerchio l'insieme dei punti del piano interni ad una circonferenza, pertanto

$$D = \{(x, y) \in \mathbb{R}^2 \ : \ (x - x_0)^2 + (y - y_0)^2 \leq r^2\}$$

individua un cerchio di centro P_0 e raggio r.

6. Geometria Analitica nello spazio



Possiamo estendere facilmente l'idea di sistema di riferimento anche alla geometria nello spazio. A questo scopo è sufficiente considerare tre rette orientate e mutuamente perpendicolari che si intersecano in un punto detto origine.

Su ciascuna retta fissiamo un punto che definisce l'unità di misura e la identifichiamo con una copia di $\mathbb{R}$.

Un punto P nello spazio è quindi univocamente determinato da una terna ordinata di numeri reali (x, y, z) dove x ,y, z sono le proiezioni di P sul ciascuno dei tre assi.

La terna di assi è orientata in modo che seguendo con le dita della mano destra la rotazione antioraria per sovrapporre l'asse delle x all'asse delle y, il pollice è rivolto nel senso positivo per l'asse delle z. Una terna siffatta si dice destrorsa.

Si definiscono anche nello spazio le operazioni: la somma di due elementi $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ definita da

$$P_1 + P_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

ed il prodotto di un elemento $P = (x, y, z)$ per uno scalare $\alpha \in \mathbb{R}$ definito da

$$\alpha P = (\alpha x, \alpha y, \alpha z)$$

Anche queste operazioni sono interpretabili graficamente mediante la regola del parallelogrammo, per quanto riguarda la somma, e la modifica della lunghezza del segmento OP per quel che riguarda il prodotto.

Lo spazio dotato di un tale sistema di riferimento viene solitamente indicato come Spazio Cartesiano ed è un modello dello spazio vettoriale $\mathbb{R}^3$, che di solito si chiama Spazio Euclideo, ed in cui è naturale considerare come base l'insieme

$$B = \{e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)\}$$

Ovviamente

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) = xe_1 + ye_2 + ze_3$$

essendo tale decomposizione univoca.

È immediato verificare, usando il teorema di Pitagora, che la distanza di un punto $P = (x, y, z)$ dall'origine $O = (0, 0, 0)$ è data da

$$\rho = d(P, O) = \sqrt{x^2 + y^2 + z^2}$$

mentre la distanza tra due punti $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ è

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Si può dotare $\mathbb{R}^3$ di una norma definendo

$$\|P\| = \|(x, y)\| = d(P, O) = \sqrt{x^2 + y^2 + z^2}$$

ed inoltre possiamo definire un prodotto scalare mediante la

$$\langle P_1, P_2 \rangle = x_1 x_2 + y_1 y_2 + z_1 z_2$$

e possiamo anche qui verificare che

$$\langle P_1, P_2 \rangle = \|P_1\| \|P_2\| \cos(\theta)$$

essendo θ l'angolo compreso tra $\overline{OP_1}$ e $\overline{OP_2}$ nel piano da essi individuato nello spazio.

6.1 *Equazione di un piano*

Sia $N = (a, b, c) \in \mathbb{R}^3$, $P_0 \in \mathbb{R}^3$, chiamiamo piano nello spazio l'insieme

$$\Pi = \{P \in \mathbb{R}^3 \ : \ \langle P - P_0, N \rangle = 0\}$$

In altre parole un piano è individuato da un suo punto P_0 e da un vettore N cui i vettori $P - P_0$ sono ortogonali $\forall P \in \Pi$.

Evidentemente quindi la relazione che definisce un piano, che chiameremo equazione del piano, può essere scritta come

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

ed anche

$$ax + by + cz + d = 0$$

non appena si sia definito $d = -(ax_0 + by_0 + cz_0)$: un punto $P = (x, y, z) \in \Pi$ se e solo se $ax + by + cz + d = 0$.

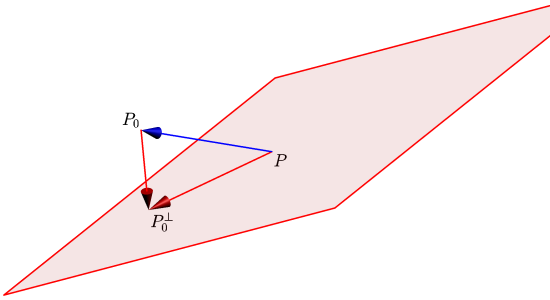
6.2 *Distanza di un punto da un piano*

Sia $P_0 = (x_0, y_0, z_0)$ un punto dello spazio e sia Π un piano di equazione $ax + by + cz + d = 0$ (il cui vettore normale è $N = (a, b, c)$) ; per

calcolare la distanza di P_0 da Π possiamo procedere in maniera analoga a quella seguita per calcolare la distanza di un punto da una retta nel piano.

Ad esempio se $P_0^\perp$ è la proiezione di P_0 su Π e se $P = (x, y, z) \in \Pi$ possiamo affermare che

$$P_0^\perp - P_0 = \langle P - P_0, \frac{N}{\|N\|} \rangle \frac{N}{\|N\|}$$



Ne segue che

$$\|P_0^\perp - P_0\| = |\langle P - P_0, \frac{N}{\|N\|} \rangle| = \frac{|\langle p, N \rangle - \langle P_0, N \rangle|}{\|N\|}$$

per cui

$$d(P_0, \Pi) = \|P_0^\perp - P_0\| = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

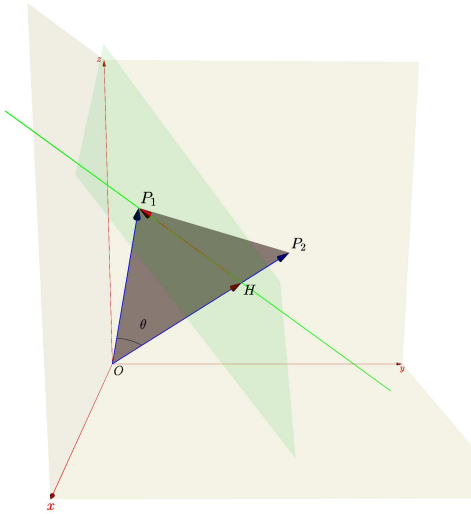
6.3 *Interpretazione Geometrica del prodotto scalare in* $\mathbb{R}^3$

Siano $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$ il prodotto scalare tra P_1 e P_2 è dato da:

$$\langle P_1, P_2 \rangle = x_1x_2 + y_1y_2 + z_1z_2$$

Consideriamo la retta che passa per l'origine e per il punto P_2 che è identificata dalle equazioni parametriche

$$\begin{cases} x = tx_2 \\ y = ty_2 \\ x = tz_2 \end{cases}$$



e proponiamoci di calcolare la distanza del punto P_1 da tale retta.

A questo scopo calcoliamo l'intersezione di tale retta con il piano, ad essa perpendicolare, passante per P_1

L'equazione di tale piano è $(x - x_1)x2 + (y - y_1)y2 + (z - z_1)z2 = \langle P - P_1, P_2 \rangle = 0$.

Sostituendo le equazioni parametriche della retta nel piano otteniamo che il punto di intersezione si trova per t che soddisfa la seguente equazione:

$$(tx_2 - x_1)x2 + (ty_2 - y_1)y2 + (tz_2 - z_1)z2 = 0$$

da cui $t = \frac{\langle P_2, P_1 \rangle}{\|P_2\|^2}$ ed il punto H di intersezione è dato da

$$H = \left(x_2 \frac{\langle P_2, P_1 \rangle}{\|P_2\|^2}, y_2 \frac{\langle P_2, P_1 \rangle}{\|P_2\|^2}, z_2 \frac{\langle P_2, P_1 \rangle}{\|P_2\|^2} \right) = P_2 \frac{\langle P_2, P_1 \rangle}{\|P_2\|^2}$$

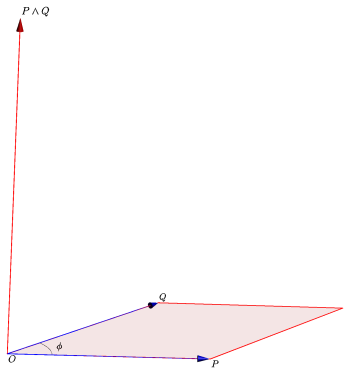
Poichè $\cos(\theta) = \frac{\|H\|}{\|P_1\|}$ possiamo concludere che

$$\cos(\theta) = \frac{\|P_2\| \langle P_2, P_1 \rangle}{\|P_2\|^2 \|P_1\|} = \frac{\langle P_2, P_1 \rangle}{\|P_2\| \|P_1\|}$$

6.4 *Prodotto vettoriale*

Siano

$$A = (a_1, a_2, a_3) = a_1e_1 + a_2e_2 + a_3e_3 \quad , \quad B = (b_1, b_2, b_3) = b_1e_1 + b_2e_2 + b_3e_3$$



dove e_1, e_2, e_3 sono i vettori ortonormali canonici della base euclidea di $\mathbb{R}^3$.

Il piano generato dai vettori A e B può essere descritto parametricamente mediante le combinazioni lineari dei due vettori:

$$\begin{cases} x = \lambda a_1 + \mu b_1 \\ y = \lambda a_2 + \mu b_2 \\ z = \lambda a_3 + \mu b_3 \end{cases}$$

Ricavando λ e μ dalle prime due equazioni e sostituendo nella terza si ottiene l'equazione cartesiana del piano che è:

$$\det \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} x + \det \begin{pmatrix} a_3 & a_1 \\ b_3 & b_1 \end{pmatrix} y + \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} z = 0$$

Ora, se consideriamo la matrice che ha per righe A e B

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

risulta evidente che il vettore le cui componenti sono i determinanti con segno dei suoi minori principali è ortogonale al piano generato da A e B.

Definiamo il *Prodotto Vettoriale* $A \wedge B$

$$A \wedge B = \left(det \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix}, \det \begin{pmatrix} a_3 & a_1 \\ b_3 & b_1 \end{pmatrix}, \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \right) =$$

$$= det \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} e_1 + \det \begin{pmatrix} a_3 & a_1 \\ b_3 & b_1 \end{pmatrix} e_2 + \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} e_3 =$$

$$= \det \begin{pmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

essendo l'ultimo determinante solo un ausilio mnemonico.

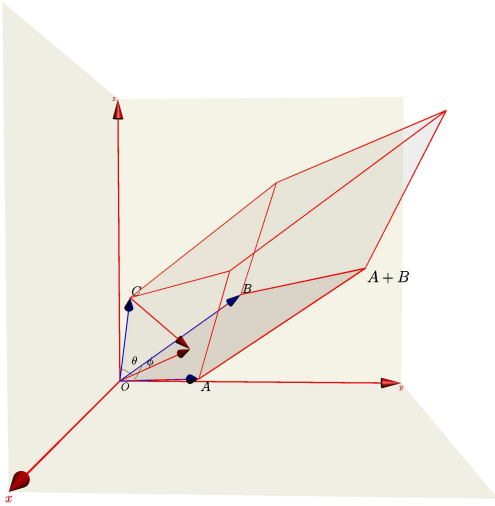
Evidentemente, quindi, $A \wedge B$ è definito in modo che sia ortogonale al piano generato da A e B.

Inoltre si ha che

La norma del vettore $||A||||B||$ è uguale all'area del parallelogramma individuato da$A$ e B

Infatti

$$||A \wedge B||^2 = (a_1^2b_2^2 + a_2^2b_1^2) + (a_2^2b_3^2 + a_3^2b_2^2) + (a_3^2b_1^2 + a_1^2b_3^2) - 2(a_1b_2a_2b_1 + a_2b_2a_3b_3 + a_3b_1a_1b_3) =$$
$$= a_1^2(b_3^2 + b_2^2) + a_2^2(b_3^2 + b_1^2) + a_3^2(b_2^2 + b_1^2) - 2(a_1b_2a_2b_1 + a_2b_2a_3b_3 + a_3b_1a_1b_3)$$
$$= a_1^2(b_3^2 + b_2^2 + b_1^2) + a_2^2(b_3^2 + b_1^2 + b_2^2) + a_3^2(b_2^2 + b_1^2 + b_3^2) - (a_1^2b_1^2 + a_2^2b_2^2 + a_3^2b_3^2 -$$
$$= ||A||^2||B||^2 - \langle A, b\rangle^2 = ||A||^2||B||^2 - ||A||^2||B||^2 cos^2\theta = ||A||^2||B||^2 \sin^2\theta$$



dove θ è l'angolo compreso tra i vettori A e B

Pertanto

$$||A \wedge B|| = ||A||||B|| \sin\theta$$

è l'area del parallelogramma di lati A e B.

Ora, posto $N = A \wedge B$ e consideriamo un terzo vettore $C \in \mathbb{R}^3$, avremo che

$$\frac{\langle C, N\rangle}{||N||} = ||C|| \cos\phi$$

dove ϕ è l'angolo compreso tra i vettori C ed N, rappresenta è la lunghezza della proiezione di C su N.

Poichè $||A \wedge B|| = ||N|| = ||A||||B|| \sin\theta$ è l'area del parallelogrammo di lati A e B, possiamo concludere che

$$||A \wedge B||||C|| \cos\phi = ||N||||C|| \cos\phi$$

fornisce il volume del parallelepipedo individuato dal volume di A, B e C

Ma

$$||A \wedge B||||C|| \cos\phi = ||N|| \frac{\langle C, N\rangle}{||N||} = \langle C, N\rangle = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Si può infine verificare, tenendo presente unicamente le definizioni, che

- $A \wedge B = -B \wedge A$
- $A \wedge B = 0$ se e solo se $A \parallel B$
- $A \wedge (B + C) = A \wedge B + A \wedge C$

6.4.1 *Sfera e superficie sferica*

Nello spazio una superficie sferica è definita come il luogo dei punti equidistanti da un punto assegnato, detto centro. La distanza comune di tutti i punti di una superficie sferica dal suo centro è detta raggio.

Pertanto, se $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ e d $r \in \mathbb{R}_+$ la superficie sferica di centro P_0 e raggio r è individuata come l'insieme

$$S = \{(x, y, z) \in \mathbb{R}^3 \ : \ (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2\}$$

Diciamo quindi, intendendo quanto sopra, che

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$$

è l'equazione di una superficie sferica di centro P_0 e raggio r.

È immediato verificare che possiamo scrivere l'equazione di una superficie sferica nella forma

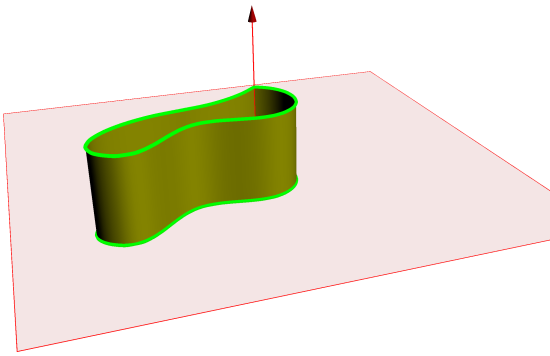
$$x^2 + y^2 + z^2 + ax + by + cz + d = 0$$

ed un confronto tra le sue equazioni permette di affermare che il centro è nel punto $(-\frac{a}{2}, -\frac{b}{2}, -\frac{d}{2})$ e che il suo raggio è $r = \sqrt{d - \frac{a^2}{4} - \frac{b^2}{4} - \frac{c^2}{4}}$.

In maniera del tutto simile definiamo sfera l'insieme dei punti

$$D = \{(x, y) \in \mathbb{R}^2 \ : \ (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < r^2\}$$

6.5 *Cilindri*



Consideriamo una curva, che possiamo supporre piana e che chiamiamo generatrice del cilindro, ed una retta non complanare con la curva, che chiamiamo asse del cilindro. Chiamiamo cilindro la superficie ottenuta considerando tutte le rette parallele all'asse che passano per un punto della generatrice.

Se consideriamo un sistema di riferimento in cui il piano xy coincide con il piano in cui giace la curva generatrice e il cui asse z coincide con l'asse del cilindro, supposto che la generatrice G sia identificata come

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

è immediato che il clindro C è individuato in $\mathbb{R}^3$ come

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

È opportuno sottolineare l'unica differenza formale nelle definizioni di G e di C risiede nel fatto che G è fatto di punti di $\mathbb{R}^2$ mentre i punti in C sono in $\mathbb{R}^3$ l'equazione $f(x, y) = 0$ essendo la stessa in entrambi i casi. come conseguenza si vede che C contiene i punti del tipo

$$r_{(x_0,y_0)} = \{(x_0, y_0, z) \ : \ z \in \mathbb{R}$$

e per ogni $(x_0, y_0) \in G$, fissato, $r_{(x_0,y_0)}$ definisce, evidentemente , una retta che passa per (x_0, y_0) ed è parallela all'asse z

6.6 *Le coniche*

Si definisce cono una superficie nello spazio generata dalla rotazione di una retta, detta generatrice, attorno ad una seconda retta ad essa incidente, detta asse.

Qualora sia necessario trattare analiticamente un cono possiamo considerare un sistema di riferimento in cui l'asse del cono coincida con uno degli assi, ad esempio l'asse z, del sistema di riferimento e il punto di incidenza coincida con l'origine del sistema.

In queste condizioni, nel piano $y = 0$, l'equazione della retta generatrice del cono è $z = p\rho$ e la superficie ottenuta facendo ruotare tale retta è caratterizzata dal fatto che

$$z = p\sqrt{x^2 + y^2}$$

Infatti ogni punto della retta generatrice ha, nello spazio, coordinate (ρ, o, z), con $z = p\rho$; ruotando,la quota si mantiene costante per i punti del piano (x, y) che hanno la stessa distanza dall'asse z di $(\rho, 0)$; questi punti sono, nel piano xy sulla circonferenza di raggio

$$\rho = \sqrt{x^2 + y^2}$$

Ad essi corrisponde quindi una quota z tale che

$$z^2 = p^2\rho^2 = p^2(x^2 + y^2)$$

Ne segue che i punti dello spazio che appartengono al cono sono identificati dall'uguaglianza precedente. Ovviamente il cono avra' due falde: la prima, nel semispazio $z \geq 0$, di equazione

$$z = p\rho = p\sqrt{x^2 + y^2}$$

e la seconda, nel semispazio $z \leq 0$, di equazione

$$z = -p\rho = -p\sqrt{x^2 + y^2}$$

Consideriamo ora un piano; a meno di scegliere opportunamente il sistema di riferimento nel piano xy possiamo supporre che il piano sia individuato da una delle seguenti equazioni.

- $y = a$
- $z = ay + b$

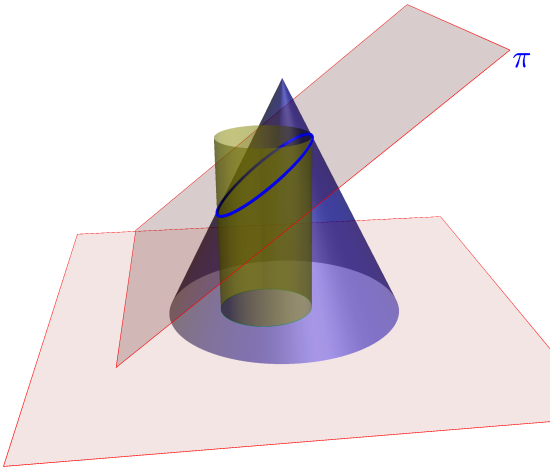
Nel primo caso quindi, i punti comuni a cono e piano sono definiti dal sistema di equazioni

$$\begin{cases} y = a \\ z^2 = p^2(x^2 + y^2) \end{cases}$$

cioè

$$\begin{cases} y = a \\ z^2 = p^2(x^2 + a^2) \end{cases}$$

La seconda equazione dell'ultimo sistema non contiene la variabile x e pertanto rappresenta un cilindro con asse parallelo all'asse x il sistema definisce quindi una curva nel piano $y = a$ che possiamo proiettare sul piano xz ottenendo il luogo dei punti del piano tali che $z^2 = p^2(x^2 + a^2)$.



Evidentemente esistono punti che soddisfano l'equazione per ogni a e per $a = 0$ si ottiene una coppia di rette nel piano $y = 0$ di equazioni $z = p|x|$

Nel secondo caso, i punti comuni a cono e piano sono definiti dal sistema di equazioni

$$\begin{cases} z = ay + b \\ z^2 = p^2(x^2 + y^2) \end{cases}$$

cioè

$$\begin{cases} z = ay + b \\ (ay + b)^2 = p^2(x^2 + y^2) \end{cases}$$

da cui

$$\begin{cases} z = ay + b \\ a^2y^2 + b^2 + 2aby = p^2x^2 + py^2 \end{cases} \qquad \begin{cases} z = ay + b \\ p^2x^2 - b^2 + (p^2 - a^2)y^2 - 2aby = 0 \end{cases}$$

Risulta pertanto che l'intersezione tra cono e piano coincide con l'intersezione tra il piano ed un cilindro la cui generatrice G è definita nel piano xy come luogo dei punti (x, y) tali che

$$(p^2 - a^2)x^2 + p^2y^2 - b^2 - 2abx = 0$$

Intanto se $a = 0$ G è il luogo dei punti del piano tali che

$$p^2x^2 + p^2y^2 - b^2 = 0$$

e quindi è una circonferenza centrata in $(0, 0)$ di raggio $\sqrt{|b/p|}$.

Nel caso in cui $p = a$ G è individuata dall'equazione

$$a^2x^2 - b^2 - 2aby = 0$$

ed identifica nel piano una curva che si chiama parabola.

Infine se $a \neq 0$ ed $a \neq p$, G è individuata come luogo dei punti del piano tali che

$$\begin{aligned} &(p^2 - a^2)y^2 + p^2x^2 - b^2 - 2aby = \\ &\quad \pm\left(\sqrt{|p^2 - a^2|}y \mp \frac{ab}{\sqrt{|p^2 - a^2|}}\right)^2 + p^2x^2 \mp \frac{a^2b^2}{|p^2 - a^2|} - b^2 = \\ &\pm\left(\sqrt{|p^2 - a^2|}y \mp \frac{ab}{|\sqrt{|p^2 - a^2|}}\right)^2 + p^2x^2 \mp b^2\left(1 \pm \frac{a^2}{|p^2 - a^2|}\right) = 0 \end{aligned}$$

Ora se $p^2 > a^2$, risulta $p^2 - a^2 = |p^2 - a^2|$ e si deve intendere $+$ in luogo di $\pm$ e $-$ in luogo di $\mp$ per cui l'equazione di G è

$$\left(\sqrt{p^2 - a^2}y - \frac{ab}{\sqrt{p^2 - a^2}}\right)^2 + p^2x^2 - b^2\left(1 + \frac{a^2}{p^2 - a^2}\right) = 0$$

da cui

$$\frac{1}{p^2 - a^2}\left((p^2 - a^2)y - ab\right)^2 + p^2x^2 = b^2\left(\frac{p^2}{p^2 - a^2}\right)$$

e

$$\left((p^2 - a^2)y - ab\right)^2 + p^2(p^2 - a^2)x^2 = b^2p^2$$

Evidentemente esistono punti del piano che soddisfano tale uguaglianza . G in questo caso di definisce ellisse.

poichè deve risultare

$$\left((p^2 - a^2)y - ab\right)^2 = b^2p^2 - p^2(p^2 - a^2)x^2 \geq 0$$

e

$$p^2(p^2 - a^2)x^2 = b^2p^2 - \left((p^2 - a^2)y - ab\right)^2 \geq 0$$

ed entrambe le precedenti equazioni sono soddisfatte per valori interni, si vede che una ellisse è limitata

Se infine $p^2 < a^2$, risulta $p^2 - a^2 = -|p^2 - a^2|$ e si deve intendere $-$ in luogo di $\pm$ e $+$ in luogo di $\mp$ per cui l'equazione di G è

$$-\left(\sqrt{a^2 - p^2}y + \frac{ab}{\sqrt{a^2 - p^2}}\right)^2 + p^2y^2 + b^2\left(1 - \frac{a^2}{a^2 - p^2}\right) = 0$$

da cui

$$-\frac{1}{a^2 - p^2}\left((a^2 - p^2)y - ab\right)^2 + p^2x^2 = -b^2\left(\frac{-p^2}{a^2 - p^2}\right)$$

e

$$-\left((a^2 - p^2)y - ab\right)^2 + p^2(a^2 - p^2)y^2 = b^2p^2$$

Evidentemente esistono punti del piano che soddisfano tale uguaglianza . G in questo caso di definisce iperbole.

poichè deve risultare

$$\left((p^2 - a^2)y - ab\right)^2 = b^2p^2 - p^2(p^2 - a^2)x^2 \geq 0$$

e

$$p^2(p^2 - a^2)x^2 = b^2p^2 + \left((p^2 - a^2)y - ab\right)^2 \geq 0$$

essendo la prima disequazione soddisfatta per valori esterni e la seconda per ogni valore,si vede che un'iperbole non è limitata e che non ha punti vicini all'origine.

6.7 Le proprieta' geometriche delle coniche.

Le curve definite dall'intersezione di un cono con un piano, che abbiamo chiamato coniche, godono di numerose interessanti proprietà che cercheremo di illustrare nel seguito

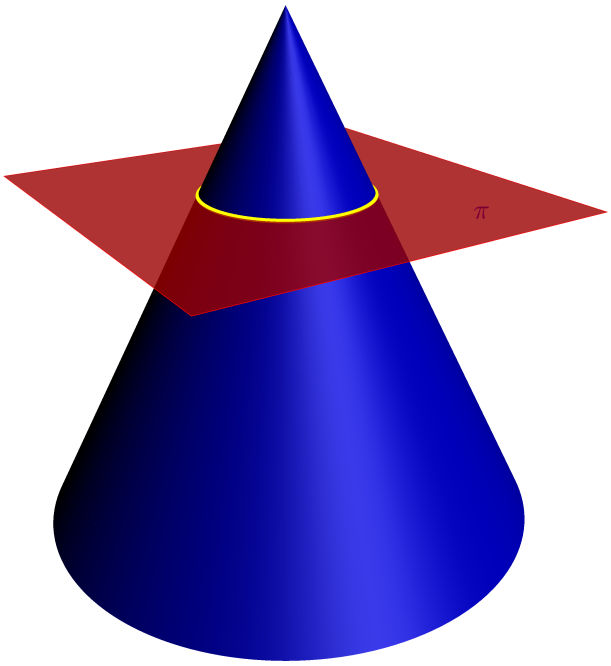
Consideriamo innanzi tutto un cono, che a meno di scegliere opportunamente il sistema di riferimento, ha equazione

$$z = p\sqrt{x^2 + y^2}$$

ed un piano, che sempre per opportuna scelta del sistema di riferimento, avrà equazione

$$z = ay + b \qquad \text{oppure} \qquad y = b$$

Cominciamo a considerare il caso in cui $a = 0$ e il piano si riduce a $z = b$.



In tal caso abbiamo già mostrato che la curva intersezione può essere definita da

$$\begin{cases} z = b \\ z = p\sqrt{x^2+y^2} \end{cases} \quad \text{ovvero} \quad \begin{cases} z = b \\ p^2x^2 + p^2y^2 = b^2 \end{cases}$$

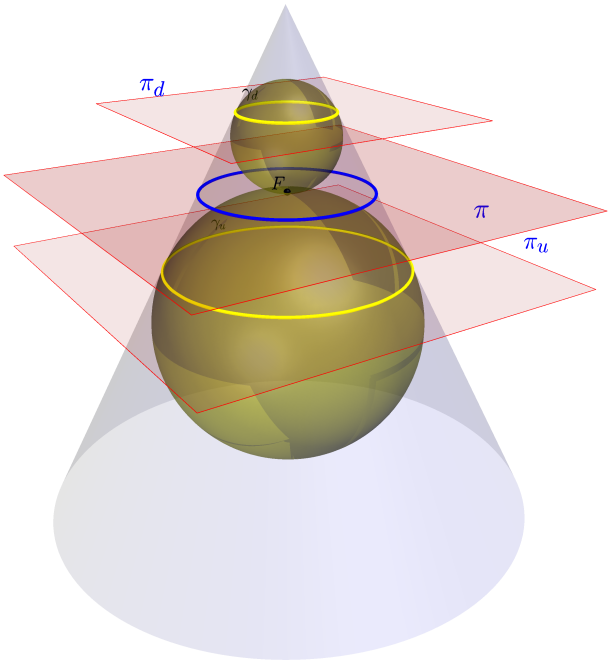
e che quindi si può caratterizzare come l'intersezione del piano $z = b$ con il cilindro con asse parallelo all'asse z generato dalla curva di equazione

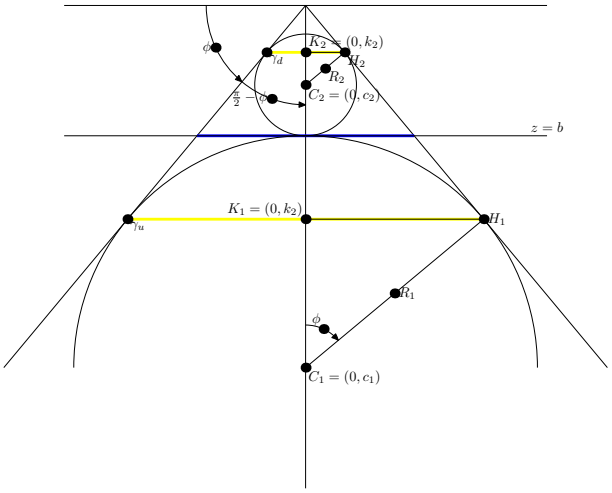
$$x^2 + y^2 = \left(\frac{b}{p}\right)^2$$

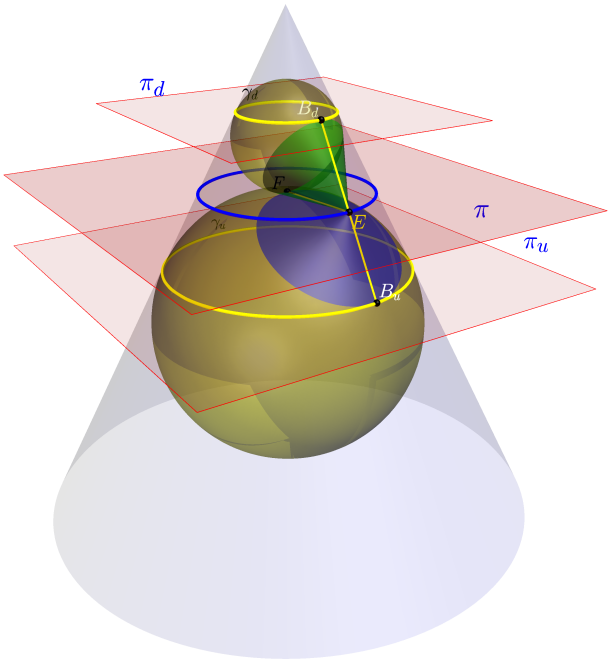
Evidentemente i punti della curva appartengono al piano $z = b$ e sono equidistanti dall'asse z o, equivalentemente, dall'intersezione dell'asse z con il piano $z = b$.

Si ha $p = \tan(\phi)$, $\cos(\phi) = \frac{1}{\sqrt{1+p^2}}$, $\sin(\phi) = \frac{p}{\sqrt{1+p^2}}$ ed esistono due sfere tangenti al cono ed al piano. Se $C_i = (0, c_i)$ è il centro ed R_i il raggio delle due sfere, si ha $(b - c_i)^2 = (c_i\cos(\phi))^2$. Ne segue che $b - c_i = \pm c_i\cos(\phi)$ e $b = c_i(1 \pm \cos(\phi))$ da cui $c = \frac{b}{1\pm\cos(\phi)} = \frac{b\sqrt{1+p^2}}{\sqrt{1+p^2}\pm 1}$

Inoltre $R_i = c_i\cos(\phi) = \frac{b\sqrt{1+p^2}}{\sqrt{1+p^2}\pm 1}\frac{1}{\sqrt{1+p^2}} = \frac{b}{\sqrt{1+p^2}\pm 1}$. Le sfere sono tangenti al cono nei punti di due circonferenzeγ_i giacenti nei piani $z = k_i$ e $c_i - k_i = R_i\cos(\phi)$, da cui $k_i = c_i - R_i\cos(\phi) = c_i - c_i\cos^2(\phi) = c_i\sin^2(\phi) = \frac{c_i p}{1+p^2}$







La circonferenza è ovviamente il luogo dei punti del piano, in cui giace, equidistanti da un punto che e' detto centro. Questa proprietà si può dedurre dal fatto che il cono è una superficie di rotazione attorno all'asse z e che il piano $z = b$ è ortogonale all'asse; si evince anche facilmente dai calcoli algebrici dai quali risulta che che i punti di una circonferenza hanno distanza costante dall'origine.

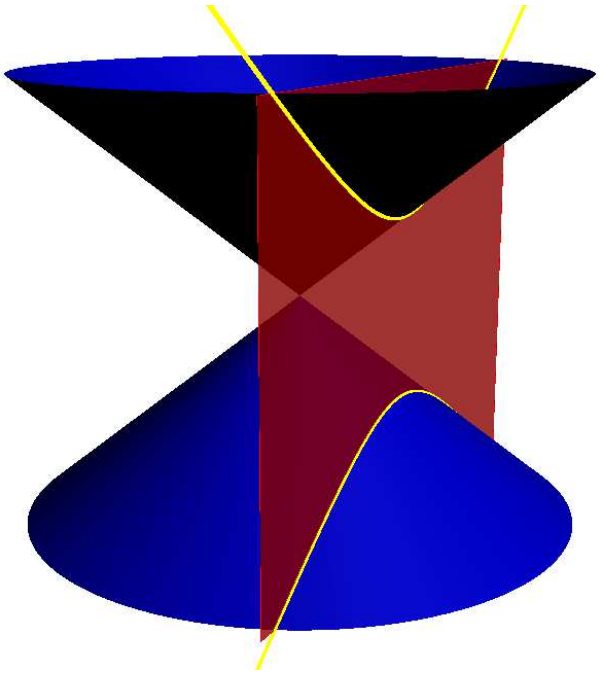
È anche interessante notare che si puo' dedurre questa proprietà dalla costruzione riportata precedentemente. Infatti , con riferimento alla figura possiamo vedere che, se E è un punto della circonferenza, B_d e B_u sono i punti in cui la generatrice del cono per E interseca le circonferenze γ_d e γ_u ed F è il punto in cui le sfere sono tangenti al piano $z = b$, per le proprietà delle rette per un punto tangenti ad una sfera, si ha

$$B_dE = B_uE = EF$$

e quindi EF è costante.

Osserviamo che il centro della circonferenza è il punto in cui il piano $z = b$ è tangente alle due sfere che abbiamo introdotto.

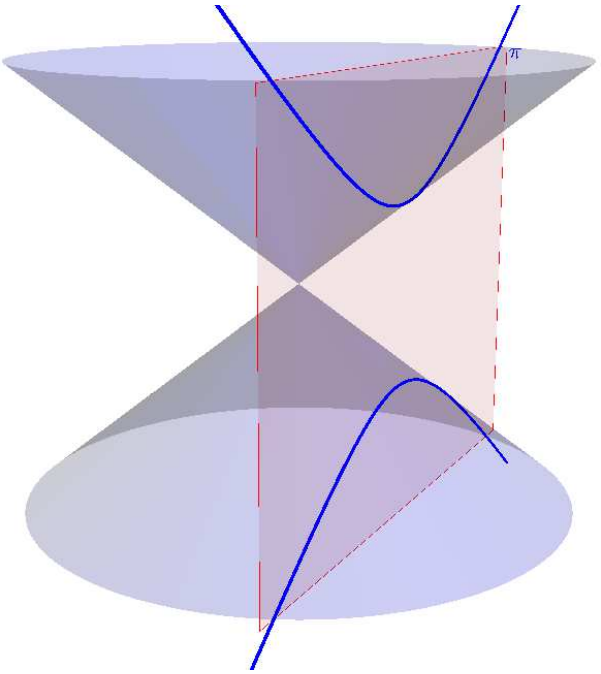
Consideriamo ora il caso in cui il cono sia intersecato da un piano di equazione $y = b$

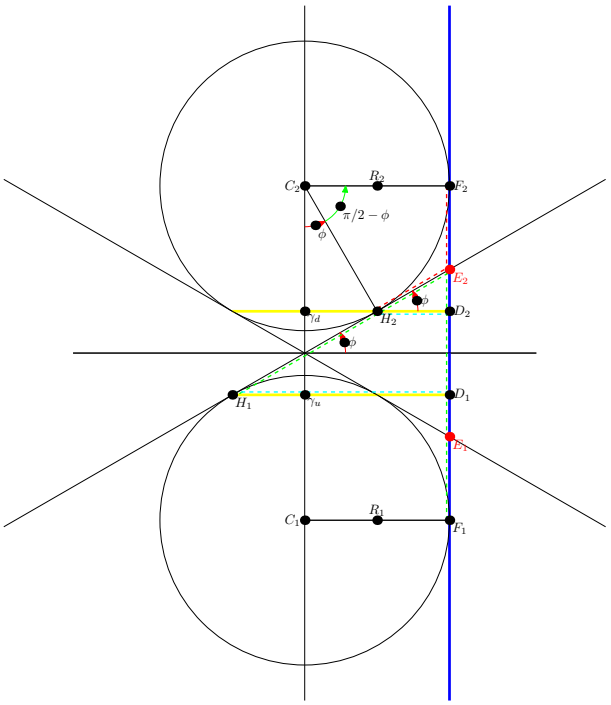


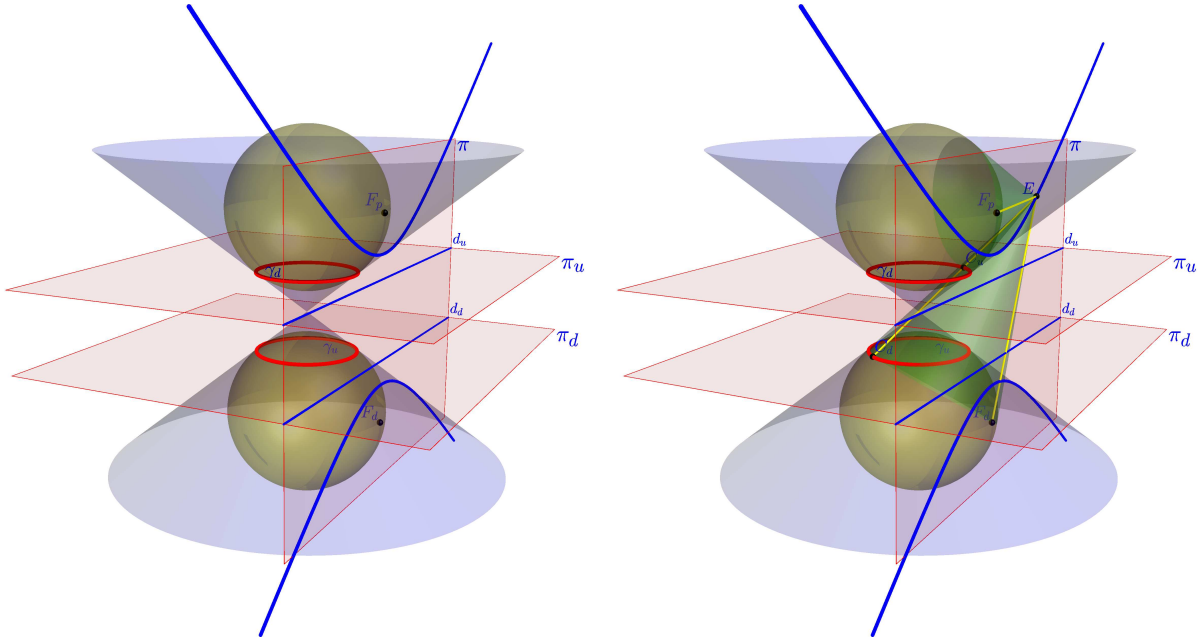
In tal caso abbiamo già mostrato che la curva intersezione può essere definita da

$$\begin{cases} y = b \\ z = p\sqrt{x^2+y^2} \end{cases} \text{ ovvero } \begin{cases} y = b \\ p^2x^2 + p^2b^2 = z^2 \end{cases}$$

Anche qui $p = \tan(\phi)$, $\cos(\phi) = \frac{1}{\sqrt{1+p^2}}$, $\sin(\phi) = \frac{p}{\sqrt{1+p^2}}$ ed esistono due sfere tangenti al cono ed al piano. $C_i = (0, c_i)$ è il centro ed R_i il raggio delle due sfere, si ha $\pm b = c_i \cos(\phi)$, da cui $c_i = \pm\frac{b}{\cos(\phi)} = b\sqrt{1+p^2}$. Inoltre $R_i = b$. Le sfere sono tangenti al cono nei punti di due circonferenze γ_i giacenti nei piani $z = \pm k_i$ e $k_i = \frac{b}{\cos(\phi)} - b\cos(\phi) = b\frac{\sin^2(\phi)}{\cos(\phi)}$, da cui $k_i = \frac{bp^2}{\sqrt{1+p^2}}$ ed aventi raggio $r_i = \frac{b}{\sin(\phi)} = \frac{bp}{\sqrt{1+p^2}}$. Le sfere sono tangenti al piano nei punti $F_i = (0, b, b/\cos(\phi))$ che prendono il nome di fuochi. Su ogni generatrice del cono ci sono un punto di ciascuna delle circonferenze γ_i ed un punto del piano considerato che descrive, al variare delle generatrici, la curva intersezione.







Ogni punto della circonferenza γ_2 può essere rappresentato mediante le $\begin{cases} x = b\,\sin(\phi)\cos(\theta) \\ y = b\,\sin(\phi)\sin(\theta) \\ z = k_i = \frac{b}{\cos(\phi)} \end{cases}$ e $\begin{cases} x = t\,b\,\sin(\phi)\cos(\theta) \\ y = t\,b\,\sin(\phi)\sin(\theta) \\ z = t\frac{b}{\cos(\phi)} \end{cases}$ descrive, per $t \in \mathbb{R}$, la retta individuata da un generico punto di γ_2 e dall'origine; poichè il punto scelto su γ_2 sta sul cono, così come l'origine, tale retta è una generatrice del cono ed interseca tanto la circonferenza γ_2 quanto il piano $y = b$; tale retta individua in tal modo un punto della curva intersezione tra piano e cono, le cui coordinate si possono calcolare scegliendo t in modo che $y = tb\sin(\phi)\sin(\theta) = b$, cioè $t = \frac{1}{\sin(\phi)\sin(\theta)}$. Ne segue che la conica in oggetto può essere rappresentata parametricamente mediante le: $\begin{cases} x = b\frac{\sin(\phi)}{\sin(\phi)\sin(\theta)}\cos(\theta) \\ y = b\frac{\sin(\phi)}{\sin(\phi)\sin(\theta)}\sin(\theta) \\ z = \frac{bp^2}{\sqrt{1+p^2}}\frac{1}{\sin(\phi)\sin(\theta)} \end{cases}$

$\begin{cases} x = \frac{b\cos(\theta)}{\sin(\theta)} \\ y = b \\ z = \frac{bp}{\sin(\theta)} \end{cases}$

Ognuno dei piani Π_u e Π_d su cui giacciono le circonferenze di tangenza individuano sul piano $z = b$ una retta, d_u e d_d,che si chiama direttrice.

Dalle precedenti figure si verifica che la conica, in questo caso, soddisfa due notevoli proprietà:

- La differenza delle distanze dei punti della conica dai fuochi è costante.
- Il rapporto tra le distanze dei punti della conica da un fuoco e dalla corrispondente direttrice è costante.

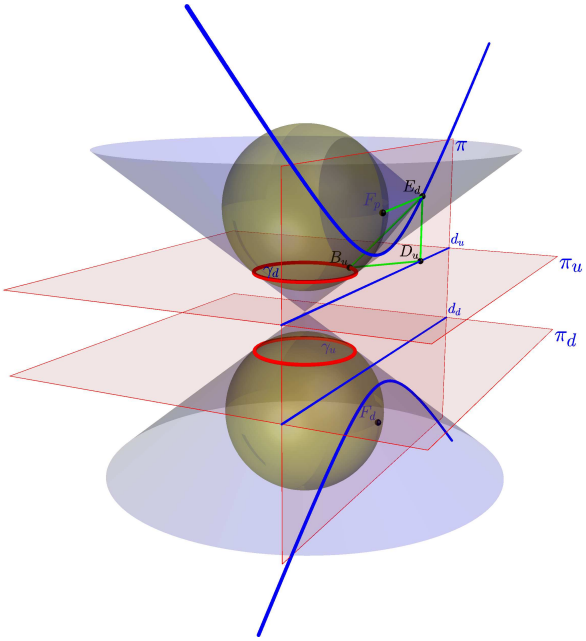
Possiamo convincerci che la prima affermazione sussiste dalle figure precedenti, osservando che ogni generatrice della conica passa per l'origine e contiene un punto di ciascuna delle circonferenze di tangenza, simmetrici rispetto all'origine; inoltre un punto della conica è individuato dall'intersezione di una generatrice col piano$y = b$.

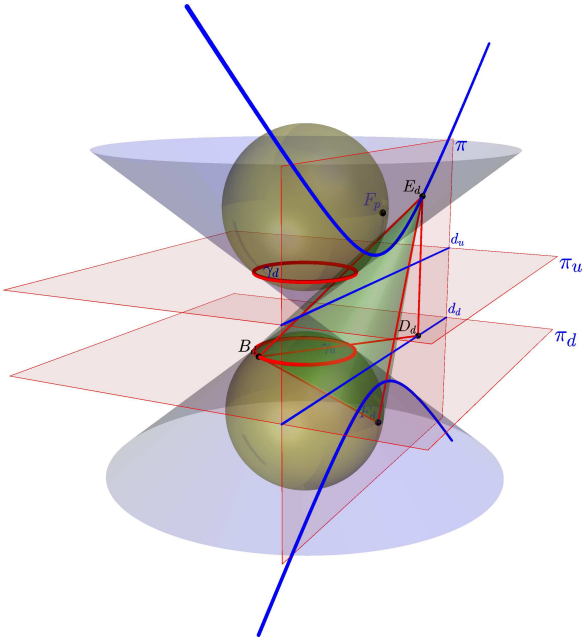
Nella figura in sezione, ad esempio, il punto della conica è indicato con E_2, la generatrice è la retta per e_2 e per l'origine e taglia le circonferenza di tangenza in H_1 e H_2. Si ha $H_2H_1 = E_2H_1 - E_2H - 2 = E_2F_1 - E_2F_2$ inoltre H_2H_1 è costante.

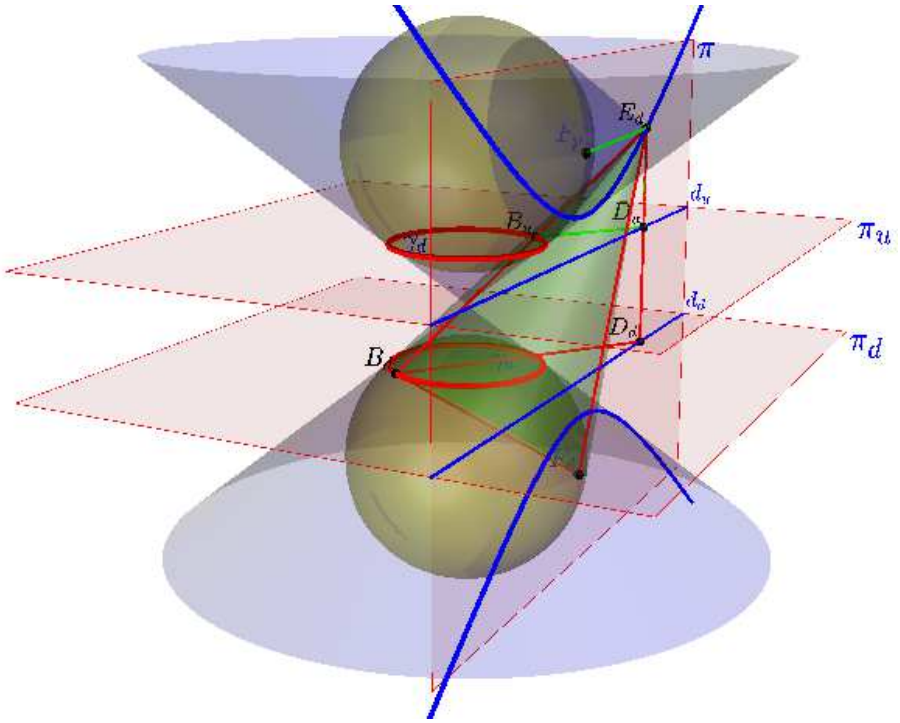
Simili considerazioni possono essere fatte osservando le figure nello spazio che immediatamente precedono.

Per quanto riguarda la seconda affermazione, sempre dalla figura in sezione si vede che il rapporto $\frac{E_2D_2}{E_2H_2} = \frac{E_2D_2}{E_2F_2} = \sin(\phi) = \frac{p}{\sqrt{1+p^2}}$ è costante e dipende unicamente dall'inclinazione delle generatrici del cono.

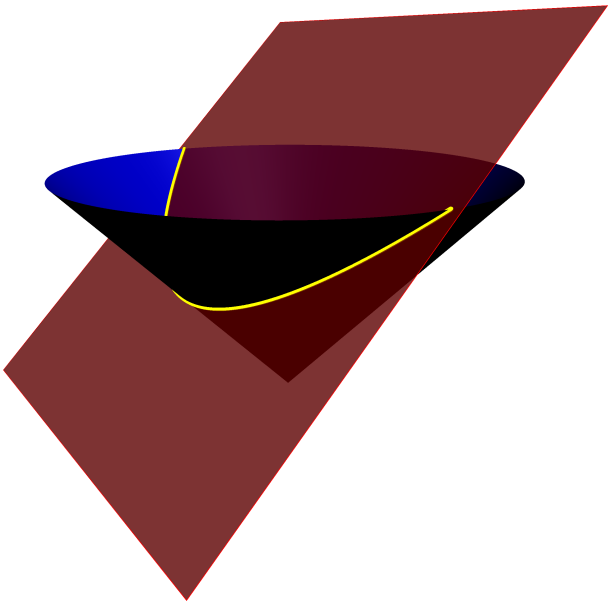
Le figure nella pagina successiva mostrano come lo stesso argomento si applichi i generale.


π
E_d
F_u
d_u
D_u
B_u
γ_d
π_u
d_d
π_d
γ_u
F_d


π
E_d
F_u
d_u
γ_d
π_u
D_d
d_d
B_u
π_d


π
E_d
F_u
d_u
D_u
B_u
γ_d
π_u
d_d
D_d
B_d
π_d
γ_u
F_d

Consideriamo ora il caso in cui $a = p$.



In tal caso abbiamo già mostrato che la curva intersezione può essere definita da

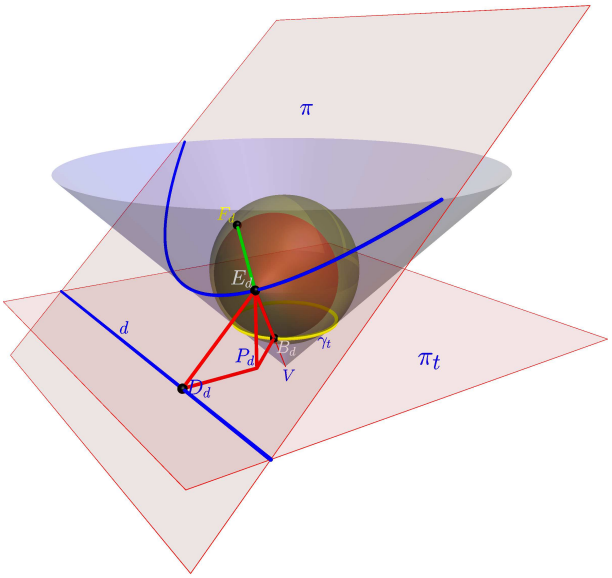
$$\begin{cases} z = py + b \\ z = p\sqrt{x^2 + y^2} \end{cases} \quad \text{ovvero} \quad \begin{cases} z = py + b \\ y = \frac{p}{2b}x^2 - \frac{b}{2p} \end{cases}$$

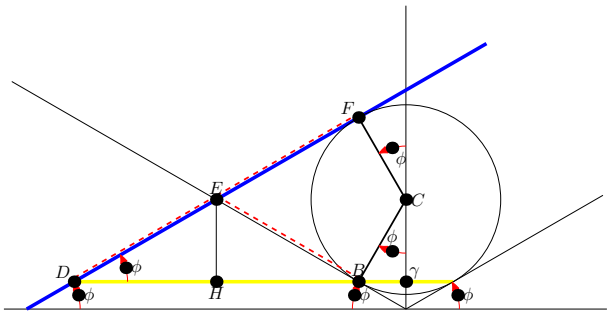
Sia Π il piano secante il cono e sia $p = \tan(\phi)$ l'apertura del cono. Π interseca solamente una delle due falde del cono ed esiste una sola sfera tangente al cono ed al piano Π ; $C = (0, 0, c)$ è il centro ed R il raggio della sfera e si ha $c = \frac{b}{2}$, e $R = \frac{b}{2}\cos(\phi) = \frac{b}{2\sqrt{1+p^2}}$. La sfera è tangente al cono nei punti di una circonferenza γ giacente nel piano $z = k$, che chiamiamo Π_t, con $k = \frac{b}{2}\sin^2(\phi) = \frac{b}{2\sin^2(\phi)} = \frac{bp^2}{2(1+p^2)}$ ed avente raggio $r = \frac{b}{2}\sin(\phi)\cos(\phi) = \frac{bp}{2(1+p^2)}$. Il piano $z = k$ su cui giace la circonferenza γ interseca il piano Π in una retta d che si chiama direttrice. La sfera è tangente al piano nel punto $F = (0, \frac{b}{2}\cos(\phi)\sin(\phi), \frac{b}{2}\cos^2(\phi))$ che prende il nome di fuoco. Ogni generatrice del cono contiene un punto della circonferenza γ e determina un punto E del piano considerato che descrive, la curva intersezione cioè la conica. d è la proiezione du d di E per cui ED è la distanza di un punto della conica dalla retta direttrice; EF è la distanza di un punto della conica dal fuoco, P è la proiezione di E su Π_t e B è la proiezione di P su γ.

Dalle figure si vede che $\frac{EP}{ED} = \sin(\phi) = \frac{EP}{EB} = \frac{EP}{EF}$ e si ottiene che $\frac{EP}{EF} = 1$ cioè $EF = ED$.

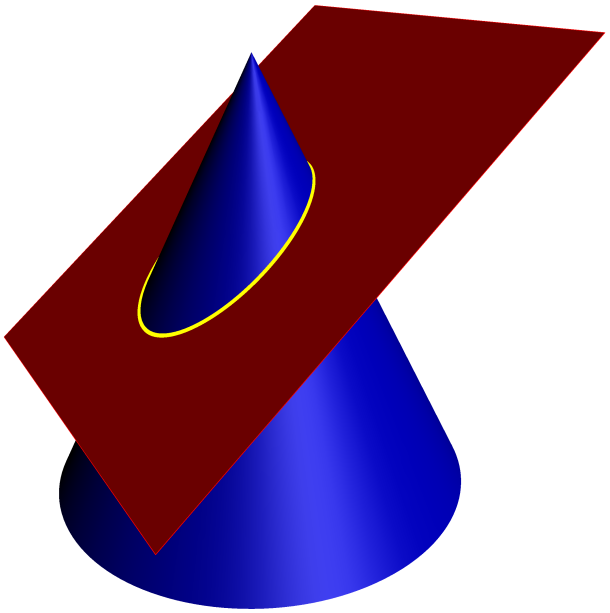
Possiamo quindi concludere che i punti della conica sono equidistanti da fuoco e direttrice e anche che il rapporto tra le distanze di un punto da fuoco e direttrice è costante.

I punti della circonferenza γ possono essere rappresentati mediante da $\begin{cases} x = r\cos(\theta) \\ y = r\sin(\theta) \\ z = k = pr \end{cases}$ e la generatrice pe tale punto si ottiene, per θ fissato, dalle $\begin{cases} x = tr\cos(\theta) \\ y = tr\sin(\theta) \\ z = tk = tpr \end{cases}$ al variare di t. Dal momento che al variare di θ la direttrice descrive il cono, la sua intersezione col piano Π descriverà la conica; tale intersezione di ottiene per t tale che $z = py + b$. Ne viene che $tpr = tpr\sin(\theta) + b$ da cui $t = \frac{b}{pr - pr\sin(\theta)}$. Avremo quindi che $\begin{cases} x = \frac{b}{p - p\sin(\theta)}\cos(\theta) \\ y = \frac{b}{p - p\sin(\theta)}\sin(\theta) \\ z = \frac{b}{1 - \sin(\theta)} \end{cases}$ che quindi forniscono una rappresentazione parametrica della conica.





Consideriamo ora il caso in cui $a < p$.

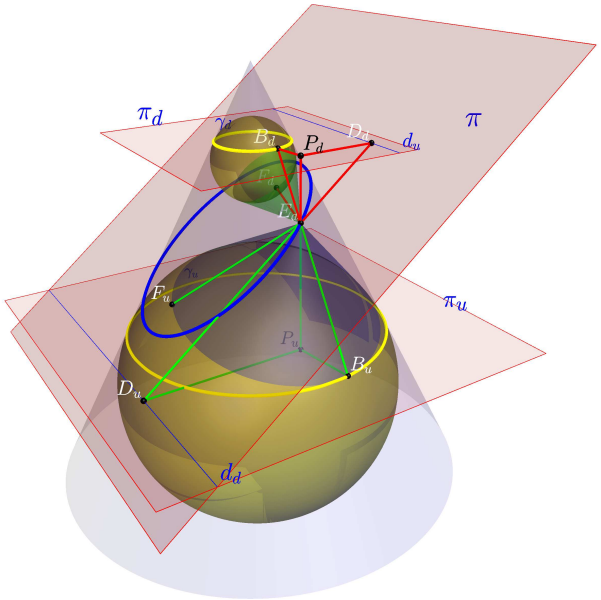


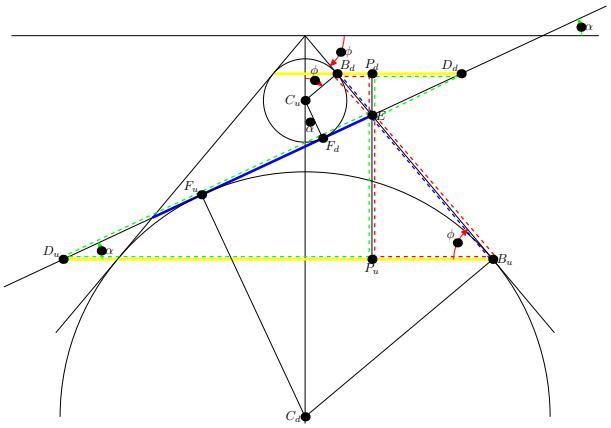
In tal caso la curva intersezione può essere definita da

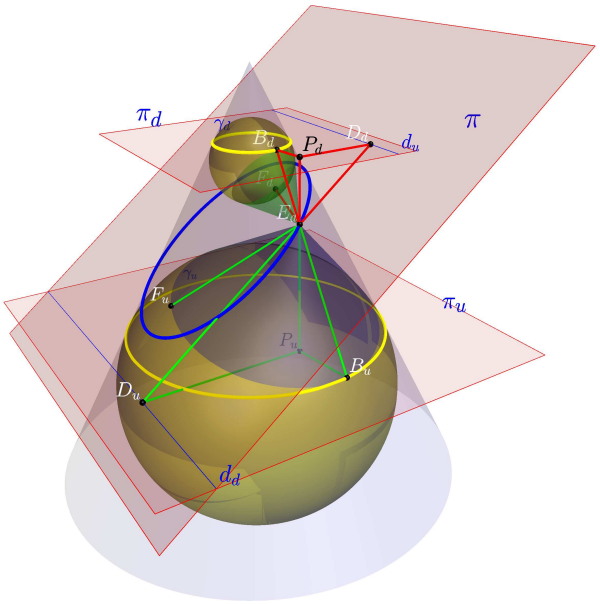
$$\begin{cases} z = ay + b \\ z = p\sqrt{x^2 + y^2} \end{cases} \quad \text{ovvero} \quad \begin{cases} z = py + b \\ \left((p^2 - a^2)y - ab\right)^2 + p^2(p^2 - a^2)x^2 = b^2p^2 \end{cases}$$

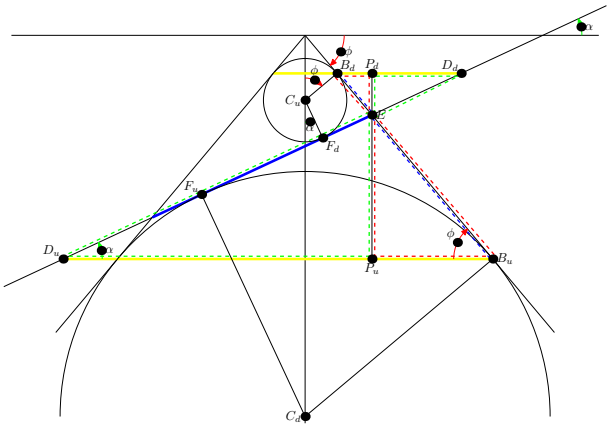
Sia Π il piano secante il cono e sia $p = \tan(\phi)$ l'apertura del cono, sia $p = \tan(\phi)$ l'apertura del cono e sia $a = \tan(\alpha)$ l'inclinazione deli piano. Esistono due sfere tangenti al cono ed al piano Π ; $C_{d,u)} = (0, 0, c_{d,u})$ sono i centri ed $R_{d,u)}$ i raggi di tali sfere e si ha $c_{d,u} = \frac{b\sqrt{1+p^2}}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$, e $R_{d,u} = \frac{b}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$. Ogni sfera è tangente al cono nei punti di una circonferenza $\gamma_{d,u}$ giacente nel piano $z = k_{d,u}$, che chiamiamo $\Pi_{d,u}$, con $k_{d,u} = \frac{bp^2}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$ ed avente raggio $r_{d,u} = \frac{bp}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$. Il piano $z = k_{d,u}$ su cui giace la circonferenza $\gamma_{d,u}$ interseca il piano $\Pi_{d,u}$ in una retta $d_{d,u)}$ che si chiama direttrice. La sfera è tangente al piano nel punto $F_{d,u} = (0, \frac{a(c_{d,u}-b)}{1+a^2}, b + \frac{a^2(c_{d,u}-b)}{1+a^2})$ che prendono il nome di fuochi. Ogni generatrice del cono contiene un punto di ognuna delle circonferenze $\gamma_{d,u}$ e determina un punto E del piano Π che descrive, la curva intersezione cioè la conica. D_d, D_u sono le proiezioni su $d_{d,u}$ di E per cui $ED_{d,u}$ è la distanza di un punto della conica dalla retta direttrice $d_{d,u}$; $EF_{d,u}$ è la distanza di un punto della conica dai fuochi, $P_{d,u}$ è la proiezione di E su $\Pi_{d,u}$ mentre $B_{d,u}$ è la proiezione di E su $\Pi_{d,u}$.

Dalle figure si vede che $B_dB_u = B_dE + EB_u = EF_d + EF_u$ e si deduce che i punti della curva intersezione hanno somma delle distanze dai fuochi costante. Inoltre $\frac{EB_{d,u}}{ED_{d,u}} = \frac{EB_{d,u}}{EP_{d,u}} \frac{EP_{d,u}}{ED_{d,u}} = \frac{\sin(\phi)}{\sin(\alpha)}$ per cui risulta costante il rapporto tra le distanze dai fuochi e dalle direttrici di un punto della curva intersezione



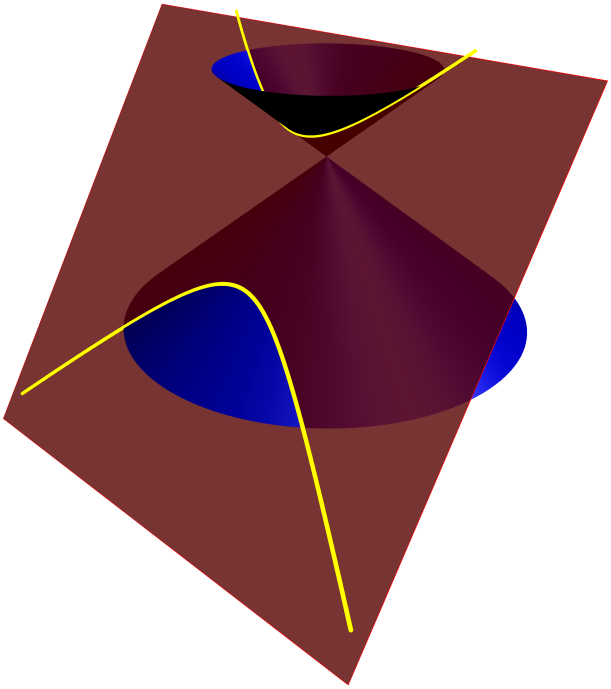


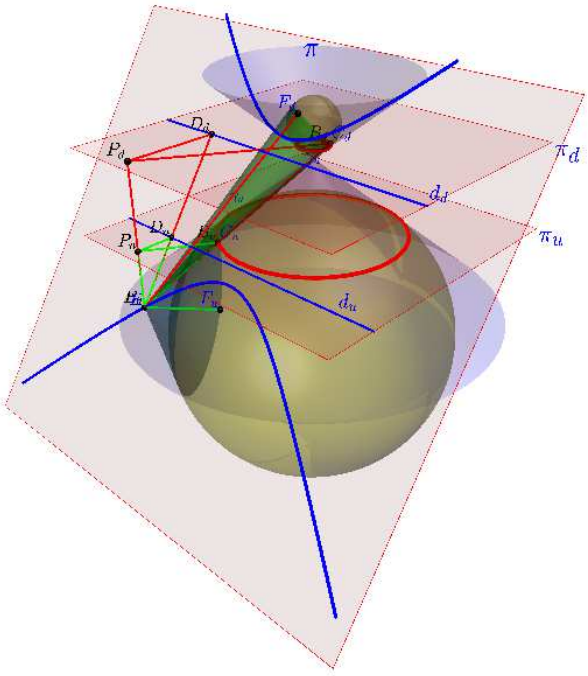




I punti delle circonferenza $\gamma_{d,u}$ possono essere rappresentati mediante da $\begin{cases} x_{d,u} = r_{d,u}\cos(\theta) \\ y_{d,u} = r_{d,u}\sin(\theta) \\ z_{d,u} = k_{d,u} = pr_{d,u} \end{cases}$ e la generatrice per tale punto si ottiene, per θ fissato, da una delle $\begin{cases} x_{d,u} = tr_{d,u}\cos(\theta) \\ y_{d,u} = tr_{d,u}\sin(\theta) \\ z_{d,u} = tpr_{d,u} \end{cases}$ al variare di t. Dal momento che al variare di θ la direttrice descrive il cono, la sua intersezione col piano Π descriverà la conica; tale intersezione di ottiene per t tale che $z = ay + b$. Ne viene che $tpr = atr\sin(\theta) + b$ da cui $tr = \frac{b}{p - a\sin(\theta)}$. Avremo quindi che le $\begin{cases} x = \frac{b}{p-a\sin(\theta)}\cos(\theta) \\ y = \frac{b}{p-a\sin(\theta)}\sin(\theta) \\ z = \frac{b}{1-a\sin(\theta)} \end{cases}$ forniscono una rappresentazione parametrica della conica.

Consideriamo ora il caso in cui $a > p$.





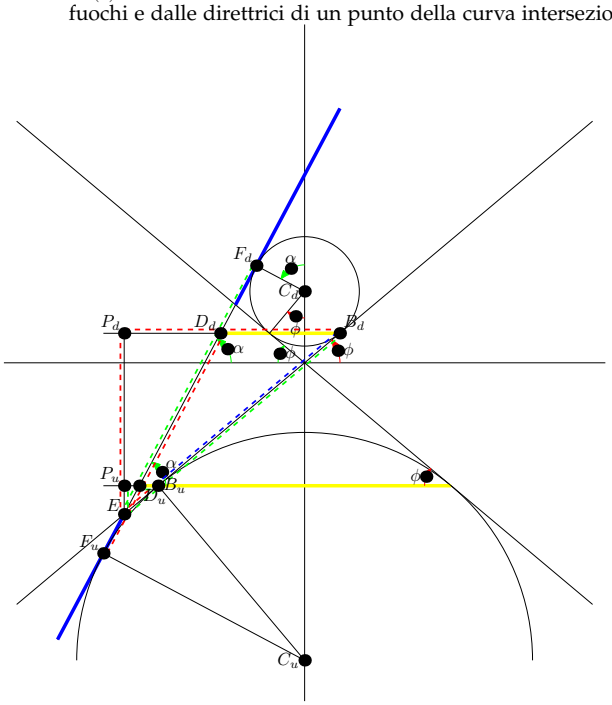
Anche in tal caso la curva intersezione può essere definita da

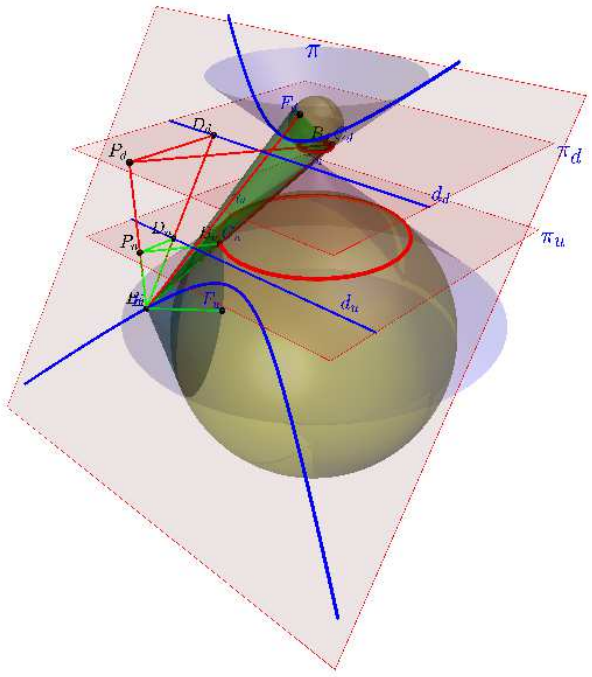
$$\begin{cases} z = ay + b \\ z = p\sqrt{x^2 + y^2} \end{cases} \quad \text{ovvero} \quad \begin{cases} z = py + b \\ \left((p^2 - a^2)y - ab\right)^2 + p^2(p^2 - a^2)x^2 = b^2p^2 \end{cases}$$

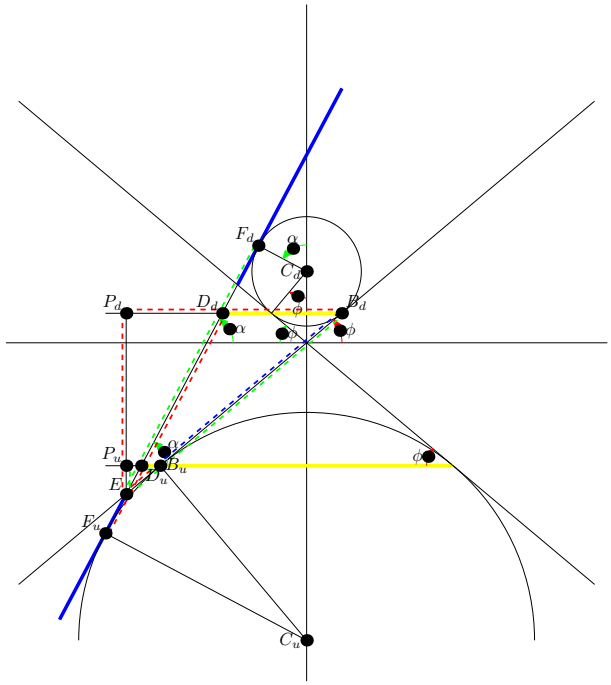
dove ovviamente $p^2 - a^2 < 0$

Sia Π il piano secante il cono, sia $p = \tan(\phi)$ l'apertura del cono e sia $a = \tan(\alpha)$ l'inclinazione deli piano. Esistono due sfere tangenti al cono ed al piano Π ; $C_{d,u)} = (0, 0, c_{d,u})$ sono i centri ed $R_{d,u)}$ i raggi di tali sfere e si ha $c_{d,u} = \frac{b\sqrt{1+p^2}}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$, e $R_{d,u} = \frac{b}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$. Ogni sfera è tangente al cono nei punti di una circonferenza $\gamma_{d,u}$ giacente nel piano $z = k_{d,u}$, che chiamiamo $\Pi_{d,u}$, con $k_{d,u} = \frac{bp^2}{\sqrt{1+p^2} \pm \sqrt{1+a^2}}$ ed avente raggio $r_{d,u} = \left| \frac{bp}{\sqrt{1+p^2} \pm \sqrt{1+a^2}} \right|$. Il piano $z = k_{d,u}$ su cui giace la circonferenza $\gamma_{d,u}$ interseca il piano $\Pi_{d,u}$ in una retta $d_{d,u)}$ che si chiama direttrice. La sfera è tangente al piano nel punto $F_{d,u} = (0, \frac{a(c_{d,u}-b)}{1+a^2}, b + \frac{a^2(c_{d,u}-b)}{1+a^2})$ che prendono il nome di fuochi. Ogni generatrice del cono contiene un punto di ognuna delle circonferenze $\gamma_{d,u}$ e determina un punto E del piano Π che descrive, la curva intersezione cioè la conica. D_d, D_u sono le proiezioni su $d_{d,u}$ di E per cui $ED_{d,u}$ è la distanza di un punto della conica dalla retta direttrice $d_{d,u}$; $EF_{d,u}$ è la distanza di un punto della conica dai fuochi, $P_{d,u}$ è la proiezione di E su $\Pi_{d,u}$ mentre $B_{d,u}$ è la proiezione di E su $\Pi_{d,u}$.

Dalle figure si vede che $B_dB_u = B_dE - EB_u = EF_d - EF_u$ e si deduce che per i punti della curva intersezione è costante la differenza delle distanze dai fuochi. Inoltre $\frac{EB_{d,u}}{ED_{d,u}} = \frac{EB_{d,u}}{EP_{d,u}} \frac{EP_{d,u}}{ED_{d,u}} = \frac{\sin(\phi)}{\sin(\alpha)}$ per cui risulta costante il rapporto tra le distanze dai fuochi e dalle direttrici di un punto della curva intersezione





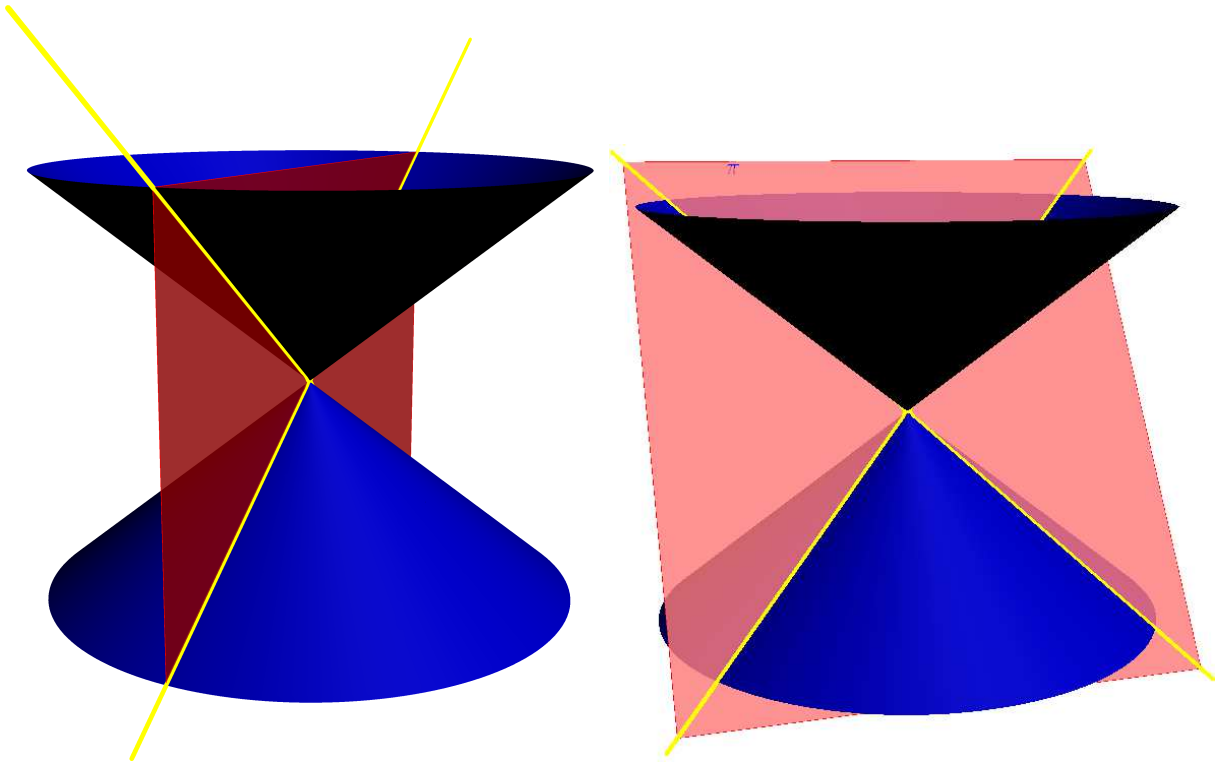


I punti delle circonferenza $\gamma_{d,u}$ possono essere rappresentati mediante da $\begin{cases} x_{d,u} = r_{d,u}\cos(\theta) \\ y_{d,u} = r_{d,u}\sin(\theta) \\ z_{d,u} = k_{d,u} = pr_{d,u} \end{cases}$ e la generatrice per tale punto si ottiene, per θ fissato, da una delle $\begin{cases} x_{d,u} = tr_{d,u}\cos(\theta) \\ y_{d,u} = tr_{d,u}\sin(\theta) \\ z_{d,u} = tpr_{d,u} \end{cases}$ al variare di t. Dal momento che al variare di θ la direttrice descrive il cono, la sua intersezione col piano Π descriverà la conica; tale intersezione di ottiene per t tale che $z = ay + b$. Ne viene che $tpr = atr\sin(\theta) + b$ da cui $tr = \frac{b}{p - a\sin(\theta)}$. Avremo quindi che le $\begin{cases} x = \frac{b}{p-a\sin(\theta)}\cos(\theta) \\ y = \frac{b}{p-a\sin(\theta)}\sin(\theta) \\ z = \frac{b}{1-a\sin(\theta)} \end{cases}$ forniscono una rappresentazione parametrica della conica.

Le due figure seguenti si riferiscono infine al caso in cui il piano secante il cono passa per il vertice del cono stesso.

La curva intersezione si riduce allora ad una coppia di rette per l'origine e le sfere di Dandelin si riducono all'origine.

Nel caso in cui il piano oltre che a passare per l'origine sia anche perpendicolare all'asse del cono, la conica si riduce ad un solo punto, l'origine stessa.



6.8 *Le equazioni canoniche per una conica*

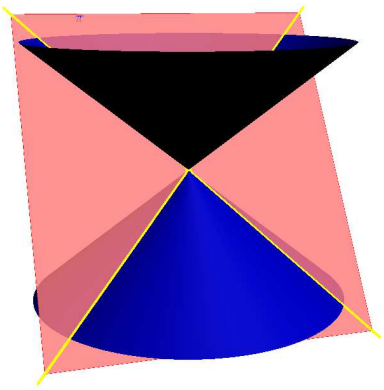
Abbiamo visto che i punti di una sezione conica sono caratterizzati, a seconda dell'inclinazione del piano secante, da proprieà che coinvolgono punti detti fuochi e rette dette direttrici.

A parte tre casi, possiamo sempre definire almeno un fuoco ed una direttrice per la sezione conica. I tre casi in questione si verificano quando

- il piano secante sia perpendicolare all'asse del cono e passi per il suo vertice. (con la scelta del riferimento che abbiamo fatta si tratta del caso in cui $a = b = 0$)
- il piano secante passi per il vertice ($b = 0$ oppure $y = 0$)

- il piano secante sia perpendicolare all'asse del cono ma non passi per l'origine ($a = 0$, $b \neq 0$)

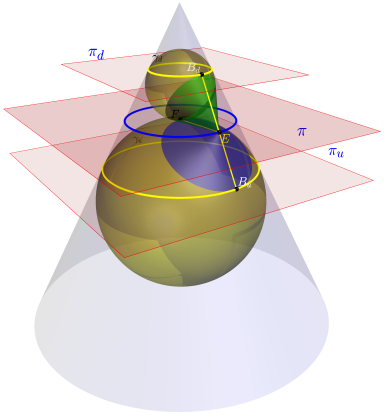
Otteniamo in ciascuno dei casi



- un punto (l'origine stessa)
- una coppia di rette
- una circonferenza

Nei primi due casi le sfere di Dandelin si riducono alla sola origine e quindi non è possibile parlare di fuochi (definiti come i punti di tangenza di ciascuna sfera al piano secante) nè di direttrici (definite come le rette intersezione tra il piano secante ed il piano che contiene la circonferenza di tangenza tra cono e sfera).

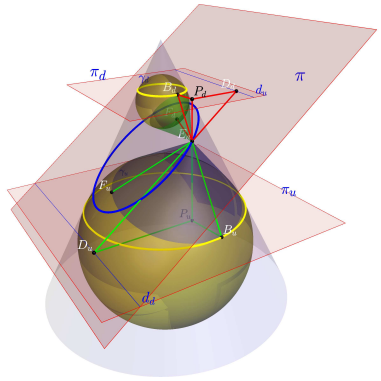
Nel terzo caso otteniamo due fuochi, coincidenti ma, poichè il piano secante ed i piani che contengono le circonferenze di tangenza sono paralleli, non è possibile parlare di direttrici; possiamo al più, dire che le direttrici si collocano all'infinito.



I casi restanti, quando cioè $a \neq 0$ e $b \neq 0$, si differenziano in base all'inclinazione del piano secante.

Osserviamo che il cono $z^2 = p^2(x^2 + y^2)$ è generato dalla rotazione attorno all'asse z della retta $z = py$ del piano (y, z) , scelto opportunamente il sistema di riferimento, il piano secante ha equazione $z = ay$

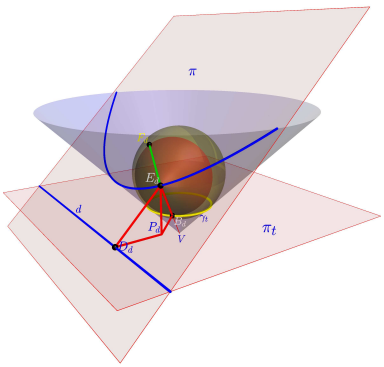
Nel caso in cui $p > a$ il piano interseca una sola falda del cono, si trovano due sfere di Dandelin tutte e due nella stessa falda del cono, ci sono due fuochi F_1 ed F_2(punti di tangenza piano sfere), e due direttrici d_1 e d_2 (rette intersezione tra il piano secante ed i piani su cui giacciono le circonferenze di tangenza tra cono e sfere).



I punti P della curva intersezione tra cono e piano soddisfano le seguenti proprietà:

1. la somma delle distanze di P dai fuochi F_1 ed F_2 è costante.
2. il rapporto tra de distanze di P da un fuoco F_i e da una direttrice d_i è costante

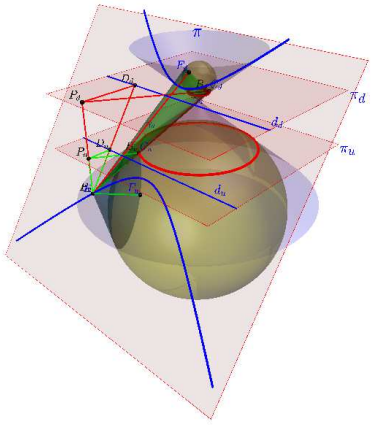
Nel caso in cui $p = a$ il piano interseca una sola falda del cono, si trova una sola sfera di Dandelin troviamo un fuoco F (punto di tangenza piano sfera), ed una direttrice d (rette intersezione tra il piano secante ed il piano su cui giace la circonferenza di tangenza tra cono e sfera).



I punti P della curva intersezione tra cono e piano soddisfano le seguenti proprietà:

1. la distanza di P dai F è uguale alla distanza di p da d.

2. il rapporto tra de distanze di P da un fuoco F_i e da una direttrice d_i è costante



Nel caso in cui $p < a$ il piano interseca entrambe le falde del cono, si trovano due sfere di Dandelin una in ciascuna falda del cono, ci sono due fuochi F_1 ed F_2(punti di tangenza piano sfere), e due direttrici d_1 e d_2 (rette intersezione tra il piano secante ed i piani su cui giacciono le circonferenze di tangenza tra cono e sfere).

I punti P della curva intersezione tra cono e piano soddisfano le seguenti proprietà:

1. la differenza delle distanze di P dai fuochi F_1 ed F_2 è costante.

2. il rapporto tra le distanze di P da un fuoco F_i e da una direttrice d_i è costante

Troviamo un fuoco F (punto di tangenza piano sfera), ed una direttrice d (retta intersezione tra il piano secante ed il piano su cui giace la circonferenza di tangenza tra cono e sfera).

I punti P della curva intersezione tra cono e piano soddisfano le seguenti proprietà:

1. la distanza di P dai F è uguale alla distanza di p da d.

2. il rapporto tra de distanze di P da un fuoco F_i e da una direttrice d_i è costante

Mediante le proprietà enunciate è possibile ricavare mediante calcoli quelle che sono le equazioni canoniche di ellisse parabola ed iperbole, trascurando le ben più semplici equazioni di circonferenza e coppie di rette, va tuttavia osservato che occorre, in ogni caso seguire strade leggermente diverse.

Esiste però una proprietà comune a tutti i casi: il rapporto tra la distanza di un punto della conica da un fuoco e da una direttrice è costante .

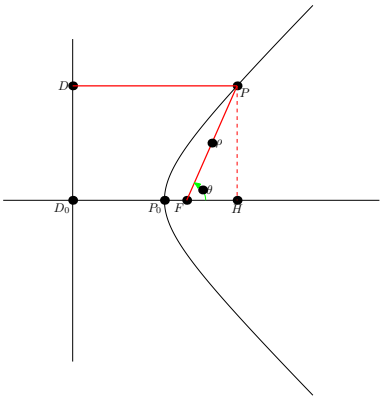
Questa proprietà consente di trovare una equazione in grado di descrivere tutti e tre i tipi di coniche al variare di un parametro che ne identifica il tipo.

Allo scopo basta considerare un riferimento polare nel piano centrato in un fuoco F e scegliere come semiasse positivo delle x la semiretta passante per F e perpendicolare alla direttrice d.

Siano P_0 il punto della conica che giace sull'asse x eD_0 il punto di intersezione tra asse x e direttrice. Siano inoltre P e D un generico punto della conica e la sua proiezione sulla direttrice d rispettivamente. Poniamo inoltre $FD_0 = f$. Avremo che, detto e il valore del rapporto tra la distanza di un generico punto della conica da fuoco e direttrice

ed f la distanza tra fuoco e direttrice,

$$\frac{P_0F}{P_0D} = \frac{PF}{PD} = e$$



Indicando, come d'uso, con ρ e θ le usuali coordinate polari del punto P, si ha che

$$\frac{\rho}{f + \rho\cos(\theta)} = e$$

da cui

$$\rho = ef + e\rho\cos(\theta)$$

e

$$\rho = \frac{ef}{1 - e\cos(\theta)}$$

Elenco delle figure

Indice

1 *I numeri Complessi* 3

2 *Introduzione al concetto di vettore* 19

3 *Spazi vettoriali* 23

4 *Applicazioni Lineari e matrici* 37

5 *Geometria Analitica nel piano* 79

6 *Geometria Analitica nello spazio* 87