I. numeri complessi, polinomi, vettori geometrici, matrici.

INSIEMI DI NUMERI

La *cardinalità* di un insieme è il numero dei suoi elementi. Un *insieme infinito* è un insieme che contiene un insieme proprio con la stessa cardinalità.
Esempio: L'insieme dei numeri naturali $\mathbf{N} = \{1, 2, 3, 4, \ldots\}$ è infinito, giacché l'insieme dei numeri naturali pari $\mathbf{P} = \{2, 4, \ldots\}$ è contenuto nell'insieme dei numeri naturali $\mathbf{N}$. Non tutti i naturali sono pari quindi $\mathbf{P}$ è strettamente contenuto in $\mathbf{N}$ inoltre la funzione $n \mapsto 2n$ è una biiezione tra $\mathbf{N}$ e $\mathbf{P}$ per cui hanno la stessa cardinalità.
Un insieme che ha la stessa cardinalità dei naturali si dice *numerabile.*

Gli insiemi dei numeri interi $\mathbf{Z} = \{\ldots, -2, -1, 0, 1, \ldots\}$ e dei numeri razionali $\mathbf{Q}$ (decimali periodici) sono numerabili, cioè: $|\mathbf{N}| = |\mathbf{Z}| = |\mathbf{Q}|$

Dimostrazione: $|\mathbf{N}| \leq |\mathbf{Z}| \leq |\mathbf{Q}| \leq |\mathbf{Z} \times \mathbf{Z}| = |\mathbf{N}|$, l'ultima uguaglianza si vede graficamente.

L'insieme dei numeri reali $\mathbf{R}$ (decimali) non è numerabile: $|\mathbf{R}| = |(-1, 1)| = |(0, 1)| > |\mathbf{N}|$.

Dimostrazione: la prima uguaglianza segue dalla biiezione $x \mapsto \frac{x}{1-|x|^2}$, la seconda dalla biiezione dall'intervallo (a, b) con l'intervallo $(0, 1)$ data da $x \mapsto (b - a)x + a$. La terza si dimostra per assurdo supponendo che l'intervallo $(0, 1)$ sia un insieme numerabile, cioé $(0, 1) = \{a_1, a_2, a_3, \ldots\}$ e si esibisce un $b \in (0, 1)$ tale che $b \neq a_i \forall i \in \mathbf{N}$.

NUMERI COMPLESSI:

$\mathbf{C}$ denota l'ambiente in cui l'equazione $x^2 + 1 = 0$ ha soluzione. Si denota con i tale numero (si ha che $i^2 = -1$, e quindi $i \notin \mathbf{R}$).
Un *numero complesso* z è una espressione del tipo $z = a + ib$ dove $a, b \in \mathbf{R}$
Uguaglianza $a + ib = c + id \Leftrightarrow a = c,\ b = d$
Operazioni di somma e prodotto:

La somma è definita da $(a + ib) + (c + id) = (a + c) + i(b + d)$
il prodotto è definito da $(a + ib).(c + id) = (ac - bd) + i(ad + bc)$
Se $z = a + ib \neq 0$ l'inverso è $z^{-1} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}$

Il coniugato di $z = a + ib$ è il numero complesso $\bar{z} = a - ib$.

Si osserva che $z + \bar{z} = 2a \in \mathbf{R}$ e che $z = \bar{z} \Leftrightarrow z \in \mathbf{R}$
Inoltre $z.\bar{z} = a^2 + b^2 \in \mathbf{R}$, e si definisce il *modulo* di $z = |z| = \sqrt{a^2 + b^2} = \sqrt{z\bar{z}}$
Si ha quindi $a = \text{Re } z = \frac{1}{2}(z + \bar{z})$; $b = \text{Im } z = \frac{1}{2i}(z - \bar{z})$ e $z^{-1} = \frac{\bar{z}}{|z|^2}$

Rappresentazione geometrica.

In un piano con un sistema di coordinate ortogonali cartesiane Oxy, dove O è l'origine, x è l'asse reale e y è l'asse immaginario, c'e una corrispondenza tra il numero complesso $z = a + ib$ e il punto P di coordinate (a, b).

Forma trigonometrica.

z è individuato dal modulo $\rho = |z|$ e dalla rotazione antioraria θ che l'asse x deve compiere per sovrapporsi al segmento OP. Si chiama argomento di z la rotazione θ. Se $z = 0$ l'argomento non è definito.

Se $z \neq 0$ si ricava $\sin\theta = \frac{b}{\rho}$ e $\cos\theta = \frac{a}{\rho}$ quindi $z = a + ib = \rho(\cos\theta + i \sin\theta)$.

Uguaglianza in forma trigonometrica

$z = z' \Leftrightarrow \rho = \rho'$ e $\theta = \theta' + 2\pi k$ per un opportuno $k \in \mathbf{Z}$

Prodotto in forma trigonometrica

$z.z' = \rho\rho'[\cos(\theta + \theta') + i \sin(\theta + \theta')]$,
in particolare $[\rho(\cos\theta + i \sin\theta)]^n = \rho^n(\cos n\theta + i \sin n\theta)$.

CAMPI:

Definizione: Un insieme $\mathbf{K}$ si chiama *campo* se in esso sono definite due operazioni: la prima, detta somma indicata comunemente con " + " soddisfa

1) $a+b=b+a\ \forall a,b\in\mathbf{K}$ (commutativa)
2) $a+(b+c)=(a+b)+c\ \forall a,b,c\in\mathbf{K}$ (associativa)
3) $\exists\, 0\in\mathbf{K} : 0+a=a\ \forall a\in\mathbf{K}$ (esistenza elemento neutro)
4) $\forall\, a\in\mathbf{K}\ \exists a'\in\mathbf{K} : a+a'=0$ e si denota $a'=-a$ (esistenza elemento opposto)

la seconda operazione, detta prodotto si indica con "." e soddisfa:

5) $a.b=b.a\ \forall a,b\in\mathbf{K}$ (commutativa)
6) $a.(b.c)=(a.b).c\ \forall a,b,c\in\mathbf{K}$ (associativa)
7) $\exists\, 1(\neq 0)\in\mathbf{K}\ : 1.a=a.1\forall\, a\in\mathbf{K}$ (esistenza elemento unitá)
8) $\forall\, a\in\mathbf{K} : a\neq 0\ \exists a^* :\ a.a^*=1$ e si denota $a^*=a^{-1}$ (esistenza elemento inverso)

in piú soddisfano le proprietá distributive

9) $a.(b+c)=a.b+a.c\ \forall a,b,c\in\mathbf{K}$
10) $(a+b).c=a.c+b.c\ \forall a,b,c\in\mathbf{K}$

Esempi: Gli insiemi dei numeri naturali $\mathbf{N}$ e interi $\mathbf{Z}$ non sono campi. Invece gli insiemi $\mathbf{Q}$ dei numeri razionali, $\mathbf{R}$ dei numeri reali, $\mathbf{C}$ dei numeri complessi sono campi, anche l'insieme $\mathbf{Z}_p=\{\bar{0},\bar{1},...,\bar{p}\}$ delle classi di resto modulo p, dove p è un numero primo è un campo.

POLINOMI

Un *polinomio* è una espressione del tipo $f(x)=a_0+a_1x+...a_nx^n$ dove x è la variabile indeterminata e $a_j\in\mathbf{K}$ sono i coefficienti di f, a_0 è detto il termine noto.
Se $a_1=\cdots=a_n=0$ si dice che il polinomio è costante. Se anche $a_0=0$ si dice che il polinomio é nullo.
Se gli $a_j\in\mathbf{C}, j=0,\ldots,n$ il polinomio si dice complesso, se gli $a_j\in\mathbf{R}, j=0,\ldots,n$ il polinomio si dice reale. L'insieme dei polinomi complessi si indica con $\mathbf{C}[x]$, l'insieme dei polinomi reali si indica con $\mathbf{R}[x]$

Definizione: Il *grado* di un polinomio non nullo è il numero intero k tale che $a_k\neq 0$ e $a_{k+j}=0,\ \forall j\in\mathbf{N}$. Il grado del polinomio nullo è $-\infty$.

Uguaglianza tra polinomi. Due polinomi $f(x)=a_0+a_1x+\cdots+a_nx^n$ e $g(x)=b_0+b_1x+\cdots+b_mx^m$ sono uguali se e solo se hanno lo stesso grado e i coefficienti sono uguali ($m=n,\ a_j=b_j\forall j$).

Operazioni tra polinomi: se $n\leq m$

la somma è $(f+g)(x)=f(x)+g(x)=a_0+b_0+(a_1+b_1)x+\cdots+(a_n+b_n)x^n+\cdots+b_mx^m$,
il prodotto è $(f.g)(x)=f(x).g(x)=a_0b_0+(a_0b_1+a_1b_0)x+\cdots+(a_nb_m)x^{n+m}$
grado $(f+g)\leq\max($grado f, grado $g)\ =m$.
grado $(f.g)=$ grado f + grado $g=n+m$.

Principio di Euclide: Se f e g sono due polinomi e g è non nullo, allora esistono due polinomi q ed r tali che $f=gq+r$ con grado $r<$ grado g. Inoltre se f e g sono polinomi reali, allora q ed r sono reali. Il polinomio q è detto il quoziente e il polinomio r è detto il resto.

Definizione: Se il resto r è il polinomio nullo si dice che f è *divisibile* per g o che f è *multiplo* di g.

Definizione: Un numero z si dice *radice* di un polinomio f se soddisfa $f(z)=0$ (cioé se z è soluzione della equazione $f(x)=0$).

Teorema di Ruffini: Se f è un polinomio, z un numero e $g(x)=(x-z)$ allora il resto $r(x)=f(z)$.

Dimostrazione: Esistono q ed r tali che $f(x) = q(x)(x-z)+r(x)$, con grado $r <$ grado $(x-z) = 1$, questo implica che r è un polinomio di grado 0 o $-\infty$ e quindi é costante (dunque $r(x) = r(z)\forall x$). Calcolando f in z si ottiene $f(z) = q(z)(z-z) + r(z) = r(z)$ che implica $r(x) = f(z)\forall x$.

Corollario: Un numero z è radice del polinomio f se e solo se il polinomi $(x-z)$ divide f. Cioé se e solo se $f(x) = (x-z)q(x)$. Inoltre se f e z sono reali allora anche q è reale.

Definizione: Se z è una radice del polinomio f, si chiama *molteplicità* di z come radice del polinomio al maggiore numero intero m tale che $(x-z)^m$ divide f.

Teorema Fondamentale dell'Algebra:

Ogni polinomio complesso di grado positivo ha almeno una radice complessa.

Corollario: Ogni polinomio complesso di grado $n > 0$ ha esattamente n radici complesse.

Principio d'identità:

A) un polinomio di grado $n > 0$ può avere al più n radici,

B) un polinomio di grado n che ha m radici con $m > n$ è necessariamente il polinomio nullo,

C) due polinomi di grado $\leq n$ che assumono gli stessi valori per m valori di x con $m > n$ sono uguali.

Teorema (Radici di un polinomio reale): Se f è un polinomio reale e $z \in \mathbf{C}$ è una radice di f, allora anche il coniugato $\bar{z}$ è una radice di f. Inoltre z e $\bar{z}$ hanno la stessa molteplicitá.

Dimostrazione: $f(\bar{z}) = a_0 + a_1\bar{z} + ... + a_n\bar{z}^n = \bar{a}_0 + \bar{a}_1\bar{z} + ... + \bar{a}_n\overline{z^n} = \overline{f(z)} = 0$.

Definizione: Un polinomio $p \in \mathbf{K}[x]$ è *irreducibile* se gli unici divisori sono 1 e p stesso.

Corollario: Su $\mathbf{C}[x]$ gli unici polinomi irreducibili sono di primo grado. Su $\mathbf{R}[x]$ i polinomi irriducibili sono quelli di primo grado e quelli quadratici senza radici reali.

Corollario: Un polinomio complesso si fattorizza nel prodotto di polinomi complessi di primo grado come $\alpha(x-z_1)^{m_1}\dots(x-z_k)^{m_k},\ z_1,\dots z_k,\alpha \in \mathbf{C}$.

Un polinomio reale si fattorizza nel prodotto di polinomi reali di grado 1 e di polinomi reali di grado 2: $\alpha(x-x_1)^{m_1}\dots(x-x_k)^{m_k}(x^2+b_1x+a_1)^{n_1}\dots(x^2+b_hx+a_h)^{n_h},\ x_i; a_j, b_j, \alpha \in \mathbf{R}$.

Esempio: Il polinomio x^3-1 su $\mathbf{R}$ si fattorizza come $(x-1)(x^2+x+1)$, mentre su $\mathbf{C}[x]$ si fattorizza come $(x-1)(x-\frac{-1+i\sqrt{3}}{2})(x-\frac{-1-i\sqrt{3}}{2})$.

APPLICAZIONE: radice n-esima di un numero complesso z.

Per il corollario del Teorema Fondamentale dell'Algebra il polinomio $x^n - z$ ha esattamente n radici complesse, quindi l'equazione $x^n = z$ ha esattamente n soluzioni complesse.

Se $z = \rho(\cos\theta + i\sin\theta)$ e $x = \sigma(\cos\phi + i\sin\phi)$ si ha che $\sigma = \sqrt[n]{\rho}$ ed $n\phi = \theta + 2\pi k$ con $k \in \mathbf{Z}$.

Al variare di k ci sono n angoli $\phi_k = \frac{\theta+2\pi k}{n}$; $k = 0,\dots n-1$, che danno luogo a numeri distinti x_k dove $k = 0,\dots n-1$. Tutte le soluzioni si trovano nel cerchio di centro 0 e raggio σ. Inoltre dividono il cerchio in n parti uguali.

Esempi: a) $\sqrt[2]{2i} = \{\sqrt[2]{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}), \sqrt[2]{2}(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4})\}$

b) $\sqrt[2]{1+i} = \{\sqrt[4]{2}(\cos\frac{\pi}{8} + i\sin\frac{\pi}{8}), \sqrt[4]{2}(\cos\frac{9\pi}{8} + i\sin\frac{9\pi}{8})\}$

c) $\sqrt[3]{1} = \{1,\ \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3},\ \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3}\} = \{1,\ \frac{-1+i\sqrt{3}}{2},\ \frac{-1-i\sqrt{3}}{2}\}$

VETTORI GEOMETRICI

Su un piano con un sistema di coordinate ortogonali cartesiane Oxy, dove O è l'origine, x è l'asse delle ascisse e y è l'asse delle ordinate, si fissa una unità di misura e ad ogni punto P del piano si associa una coppia ordinata di numeri reali di coordinate (x, y).
Si denota con $\mathbf{R}^2 = \{(x, y) \ : x, y \in \mathbf{R}\}$. Il simbolo $\vec{0}$ rappresenta il vettore nullo di componenti $(0, 0)$ (di lunghezza 0).
I punti sulle ascisse soddisfano l'equazione $y = 0$, i punti sull'asse delle ordinate soddisfano l'equazione $x = 0$.
Il segmento orientato uscente dall'origine e di estremo il punto P di coordinate (x, y) rappresenta il vettore v di componenti (x, y).

Se $v_1 = (x_1, y_1)$ e $v_2 = (x_2, y_2)$ sono due vettori in $\mathbf{R}^2$ si definisce la somma:

$$v_1 + v_2 = (x_1, y_1) + (x_2, y_2) \ = \ (x_1 + x_2, y_1 + y_2)$$

se $\lambda \in \mathbf{R}$ il prodotto di $v = (x, y)$ per λ è

$$\lambda v = \lambda(x, y) = (\lambda x, \lambda y)$$

Se $\lambda > 0$ si ha che λv ha la stessa direzione e verso di v,
se $\lambda < 0$ λv ha la stessa direzione di v e verso opposto,
se $\lambda = 0$ si ha $\lambda v = \vec{0}$.
L'opposto di v è $-v = -(x, y) = (-x, -y)$.
Interpretazione geometrica (regola del parallelogramma): il vettore somma $v_1 + v_2$ coincide con la diagonale del parallelogramma costruito sui vettori v_1 e v_2. Il vettore differenza $v_1 - v_2$ è parallelo al segmento da (x_1, y_1) a (x_2, y_2).

Analogamente si definiscono su $\mathbf{R}^n$, lo spazio delle n-uple reali, e su $\mathbf{C}^n$, lo spazio delle n-uple complesse, l'addizione di vettori e moltiplicazione per uno scalare.

MATRICI

Una matrice è un insieme di numeri ordinati per righe e colonne

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

Si indica con $A_k = (a_{k1} \quad \dots \quad a_{kn})$ la k-esima riga.

Si indica con $A^k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}$ la k-esima colonna.

Si dice che A ha ordine m per n si ha m righe ed n colonne.

Se gli $a_{ij} \in \mathbf{R}$ la matrice A si dice reale. Se gli $a_{ij} \in \mathbf{C}$ la matrice A si dice complessa.
Si denota con $M(m, n; \mathbf{K})$ l'insieme delle matrici con coefficienti in $\mathbf{K}$ di ordine m per n.

Somma di matrici e moltiplicazione per uno scalare.

Se B è un'altra matrice di ordine m per n con coefficienti (b_{ij}) si definisce la somma

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}.$$

Se α è uno scalare si definisce il prodotto tra uno scalare e una matrice

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

Una matrice si dice:
nulla se $a_{ij} = 0 \ \forall i = 1, \dots, m; \ j = 1, \dots, n$, e si indica con $\mathbf{0}$,
trasposta di A se ha come $ij-$esimo coefficienti a_{ji} e si indica con A^t,
quadrata se $m = n$.

Una matrice quadrata è chiamata:
identità se $a_{ij} = 0 \ \forall i \neq j$, $a_{ii} = 1$ e si indica con I,
diagonale se $a_{ij} = 0 \ \forall i \neq j$,
scalare se $a_{ij} = 0 \ \forall i \neq j$ e $a_{ii} = a \forall i$.
triangolare superiore se $a_{ij} = 0 \ \forall i > j$,
triangolare inferiore se $a_{ij} = 0 \ \forall i < j$,
simmetrica $A = A^t$,
antisimmetrica $A = -A^t$.

Definizione: Si chiama *traccia* di A la somma degli elementi sulla diagonale

$$\text{traccia } A = a_{11} + \cdots + a_{nn}$$

PRODOTTO di MATRICI (riga per colonna).

Se $A \in M(m, n; \mathbf{K})$ e $B \in M(n, p; \mathbf{K})$ (il numero di colonne di A è uguale al numero di righe di B) il prodotto di $A = (a_{ij})$ per $B = (b_{ij})$ è la matrice $C = AB \in M(m, p; \mathbf{K})$ che ha come ij-esimo coefficiente

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

Proprietà del prodotto di matrici

1) $(AB)C = A(BC)$
2) $A(B+C) = AB + AC$
3) $(A+B)C = AC + BC$
4) $a(AB) = (aA)B = A(aB)$
5) $IA = AI = A$
6) $0A = A0 = 0$
7) $(AB)^t = B^t A^t$

verifica di 7): Siano $A = (a_{ij}),\ B = (b_{ij}),\ A^t := (\alpha_{ij} = a_{ji}),\ B^t := (\beta_{ij} = b_{ji})$
$B^t A^t = (\Sigma_k \beta_{ik}\alpha_{kj}) = (\Sigma_k b_{ki} a_{jk}) = (\Sigma_k a_{jk} b_{ki}) = (c_{ji}) = (AB)^t$.

8) Il prodotto non è commutativo:

Esempio: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; BA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

9) Non vale la legge dell'annulamento:

Il prodotto di due matrici non nulle può essere la matrice nulla:

Esempi: 1) $\begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; 2) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

DETERMINANTE.

Si dice *permutazione* di n elementi ogni ordinamento degli elementi stessi. Il numero di permutazioni di n elementi è $n!$. La permutazione fondamentale è $(1, 2, \cdots, n)$. Due elementi di una permutazione σ presentano una *inversione* rispetto alla permutazione naturale se si succedono in ordine opposto a quello naturale. La permutazione σ si dice *pari* o *dispari* a seconda che il numero totale delle inversioni sia pari o dispari.

Esempio: Le permutazioni dei tre primi naturali sono sei:
$(1\ \ 2\ \ 3), (2\ \ 3\ \ 1), (3\ \ 2\ \ 1), (1\ \ 3\ \ 2), (2\ \ 1\ \ 3), (3\ \ 1\ \ 2)$.
sono pari: $(1\ \ 2\ \ 3), (2\ \ 3\ \ 1), (3\ \ 1\ \ 2)$
sono dispari: $(3\ \ 2\ \ 1), (1\ \ 3\ \ 2), (2\ \ 1\ \ 3)$

Definizione:
Se $n = 1$ $\det(a) := a$.
Se $n = 2$, $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - cb$
Se $n = 3$,

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

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

Il determinante corrispondente alla matrice quadrata A di ordine n è il numero

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma} \epsilon(\sigma)\ a_{1\sigma_1} \cdot \ldots \cdot a_{n\sigma_n}$$

dove la somma è sulle permutazioni ed $\epsilon(\sigma)$ è la paritá della permutazione σ.

Segue dalla definizione che se A è una matrice di ordine n $\det(\alpha A) = \alpha^n \det A$.

Nota: in generale $\det(A+B) \neq \det A + \det B$

esempio: $\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Definizione assiomatica del determinante:

1) $\det(A^1, \dots, C + C', \dots, A^n) = \det(A^1, \dots, C, \dots, A^n) + \det(A^1, \dots, C', \dots, A^n)$
2) $\det(A^1, \dots, \alpha C, \dots, A^n) = \alpha \det(A^1, \dots, C, \dots, A^n)$
3) $\det(A^1, \dots, C, \dots, C, \dots, A^n) = 0$
4) $\det(I) = 1$

Inoltre il determinante ha le seguente proprietà:

5) $\det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) = -\det(A^1, \dots, A^j, \dots, A^i, \dots, A^n)$
6) $\det(A^1, \dots, A^k + \alpha A^j, \dots, A^n) = \det(A^1, \dots, A^k, \dots, A^n)$
7) $\det(A) = \det(A^t)$
8) $\det(AB) = \det(A)\det(B)$ (Teorema di Binet)
9) $\det(\mathbf{0}) = 0$

Teorema: Le proprietà 1,2,3 e 4 determinano univocamente il determinante come funzione delle colonne.

Esempio: Calcolo del determinante della matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ applicando gli assiomi di definizione

Notare che $\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ d \end{pmatrix}$ e $\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix}$

$$\begin{aligned}
\begin{vmatrix} a & b \\ c & d \end{vmatrix} &= \begin{vmatrix} a & b \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = a\begin{vmatrix} 1 & b \\ 0 & d \end{vmatrix} + c\begin{vmatrix} 0 & b \\ 1 & d \end{vmatrix} \\
&= a\begin{vmatrix} 1 & b \\ 0 & 0 \end{vmatrix} + a\begin{vmatrix} 1 & 0 \\ 0 & d \end{vmatrix} + c\begin{vmatrix} 0 & b \\ 1 & 0 \end{vmatrix} + c\begin{vmatrix} 0 & 0 \\ 1 & d \end{vmatrix} \\
&= ab\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + ad\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + cb\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + cd\begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} \\
&= ab\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + ad\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - cb\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + cd\begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} \\
&= ad - cb
\end{aligned}$$

dimostrazione di 5):

$$\begin{aligned}
0 &= \det(A^1, \dots, A^i + A^j, \dots, A^i + A^j, \dots, A^n) \\
&= \det(A^1, \dots, A^i, \dots, A^i, \dots, A^n) + \det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) \\
&\quad + \det(A^1, \dots, A^j, \dots, A^i, \dots, A^n) + \det(A^1, \dots, A^j, \dots, A^j, \dots, A^n)
\end{aligned}$$

$$\Rightarrow \; 0 = \det(A^1, \dots, A^i, \dots, A^j, \dots, A^n) + \det(A^1, \dots, A^j, \dots, A^i, \dots, A^n)$$

Corollario di Binet: Se A e B sono matrici quadrate di ordine n allora

a) $\det(AB) = \det(BA)$
b) $\det(A^k) = (\det A)^k$

Nota: Se A e B non sono quadrate la formula a) non è valida.

Esempio: $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$,

$$AB = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}; \; BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

$\det(AB) = 3$

$$\det(BA) = 1 \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} - 1 \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} + 0 \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} 1(1) - 1(1) + 0(1) = 0$$

Definizione: Il complemento algebrico o cofattore A_{ij} di un elemento a_{ij} è il determinante del minore complementare di a_{ij} moltiplicato per $(-1)^{i+j}$.

Teorema di Laplace (I):
La somma dei prodotti degli elementi di una riga per i suoi complementi algebrici è $\det A$.

$\det A = a_{r1}A_{r1} + a_{r2}A_{r2} + \cdots + a_{rn}A_{rn}$ [sviluppo secondo la riga r (r fisso)]

$\det A = a_{1s}A_{1s} + a_{2s}A_{2s} + \cdots + a_{ns}A_{ns}$ [sviluppo secondo la colonna s (s fisso)]

Teorema di Laplace (II):
La somma dei prodotti degli elementi di una riga per i complementi algebrici di un'altra riga è zero

$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0$ per $i \neq j$

$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = 0$ per $i \neq j$

MATRICE INVERSA. Sia A una matrice quadrata di ordine n. Se esiste una matrice B tale che $AB = BA = I$ si dice che A è invertibile e che B è l'inversa di A.

Se A e B sono matrici invertibili allora anche AB lo è e la sua inversa è $(AB)^{-1} = B^{-1}A^{-1}$

Teorema:

1) la matrice A è invertibile se e solo se $\det A \neq 0$.
2) Se $\det A \neq 0$ allora $A^{-1} = \frac{1}{\det A}(A_{ij})^t = \frac{1}{\det A}(A_{ji})$
3) Se A è invertibile allora $\det A^{-1} = \frac{1}{\det A}$

Dimostrazione:

Se $\exists A^{-1}$ allora $\det(AA^{-1}) = \det I = 1$,
per Binet $\det(AA^{-1}) = \det(A)\det(A^{-1})$
quindi $\det(A)\det(A^{-1}) = 1 \Rightarrow \det A \neq 0$, $\det(A^{-1}) \neq 0$ e $\det(A^{-1}) = (\det A)^{-1}$.

Per i Teoremi di Laplace

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & \dots & A_{n1} \\ \vdots & \vdots & \vdots \\ A_{1n} & \dots & A_{nn} \end{pmatrix} = \begin{pmatrix} \det A & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \det A \end{pmatrix} = (\det A) \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

Se $\det A \neq 0$ si ha che A è invertibile e $\frac{1}{\det A}(A_{ij})^t$ è la sua inversa.

RIDUZIONE
Operazione elementari sulle righe:

1) moltiplicazione della riga j-esima R_j di A per uno scalare α non nullo: $R_j \longrightarrow \alpha R_j$,
2) sostituzione della riga R_j per la riga R_j più α volte la riga R_i $j \neq i, \alpha \neq 0$: $R_j \longrightarrow R_j + \alpha R_i$,
3) Intercambio fra le righe j-esima e la riga i-esima: $R_j \longleftrightarrow R_i$.

Una matrice è *ridotta per righe* se vale: in ogni riga non nulla di A c'e un elemento non nullo al di sotto del quale ci sono soltanto zeri. Il primo elemento di ogni riga con questa proprietà si chiama elemento speciale. In una riga nulla non ci sono elementi speciali.
Una matrice A si dice *equivalente per righe a* una matrice B se e solo se B è ottenuta da A per una successione finita di operazioni elementari sulle righe.

Si definiscono analoghe operazioni elementari sulle colonne e si dice che una matrice è *ridotta per colonne* se vale: in ogni colonna non nulla di A c'e un elemento non nullo a destra del quale ci sono soltanto zeri. il primo elemento di ogni colonna. con questa proprietà si chiama elemento speciale. In una colonna nulla non ci sono elementi speciali.
Una matrice A si dice *equivalente per colonne a* una matrice B se e solo se B è ottenuta da A per una successione finita di operazioni elementari sulle colonne.

RIDUZIONE A SCALA (Algoritmo di Gauss)
La riduzione a scala (l'algoritmo di Gauss) trasforma una matrice (quadrata) in una matrice a scala (triangolare superiore).
Si considera la prima colonna. Se contiene solo zeri si pone $p_1 = 0$ e si passa alla seconda colonna. Se invece la colonna ha qualche coefficiente non nullo, si scambiano le righe in modo che sia
$a_{11} \neq 0$ e si pone $p_1 = a_{11}$,
sulle altre righe si effettuano le operazioni elementari che annullano i coefficienti sotto $p_1 = a_{11}$:
per $j = 2, \dots, n$ si sostituisce R_j per $\frac{-a_{j1}}{p_1} R_1 + R_j$
Si procede in modo analogo con la seguente colonna fino ad ottenere una matrice a scala (triangolare superiore nel caso che la matrice iniziale sia quadrata) equivalente per righe alla matrice iniziale.

Esempio: calcolare il determinante (Sanini)

$$\begin{vmatrix} 5 & 3 & 1 & 4 \\ 4 & 2 & 1 & 3 \\ 6 & 0 & 3 & 2 \\ 23 & 8 & 7 & 9 \end{vmatrix} = \begin{vmatrix} 5 & 3 & 1 & 4 \\ -1 & -1 & 0 & -1 \\ -9 & -9 & 0 & -10 \\ -12 & -13 & 0 & -19 \end{vmatrix} = \begin{vmatrix} -1 & -1 & -1 \\ -9 & -9 & -10 \\ -12 & -13 & -19 \end{vmatrix} = \begin{vmatrix} -1 & 0 & 0 \\ -9 & 0 & -1 \\ -12 & -1 & -7 \end{vmatrix} = 1$$

dove la prima uguaglianza segue da: $R_2 \to R_2 - R_1,\ R_3 \to R_3 - 3R_1,\ R_4 \to R_4 - 7R_1$;
la seconda usando lo sviluppo di Laplace per la terza colonna;
la terza segue da $C^2 \to C^2 - C^1,\ C^3 \to C^3 - C^1$;
la quarta usando lo sviluppo di Laplace per la prima riga.

APPLICAZIONE Calcolo del rango di una matrice usando la riduzione:

Il rango di una matrice ridotta per righe coincide col numero di righe non nulle. Il rango di una matrice A (non ridotta) è uguale al rango di una matrice ridotta B equivalente per righe ad A.

Il rango di una matrice ridotta per colonna coincide col numero di colonne non nulle. Il rango di una matrice A è uguale al rango di una matrice ridotta B equivalente per colonne ad A.

Il rango è un invariante della matrice, sia essa ridotta per colonne o ridotta per righe. Se A è una matrice di ordine $m \times n$ allora rango $A \leq \min\{m, n\}$ (Il rango è minore o uguale al minimo tra il numero di righe e il numero di colonne).

Esempi:

$$\begin{pmatrix} 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 0 \\ 5 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 5 & 3 & 1 & 0 \\ 4 & 2 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 2 & 8 & -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{pmatrix}.$$

La prima matrice è ridotta per righe, la seconda per righe e colonne, la terza per colonne, la quarta non è ridotta ne per righe ne per colonne.
Le matrici hanno rango 4, tranne l'ultima che ha rango 2.

Per ridurre l'ultima matrice per righe si possono fare le seguenti operazioni elementari sulle righe: prima $R_2 \to R_2 - R_1,\ R_3 \to R_3 - 2R_1,\ R_4 \to R_4 + R_1$ e poi $R_3 \to R_3 - R_2,\ R_4 \to R_4 + \frac{1}{2}R_2$.

Per ridurre per colonne si possono fare la seguenti operazioni sulle colonne: $C_2 \to C_2 - C_1,\ C_3 \to C_3 - C_1$ e successivamente $C_3 \to C_3 - C_2$. Quindi la matrice è equivalente per righe a
$\begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ e per colonne a $\begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & 0 \\ 2 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ quindi il rango è 2.

RANGO
Definizione: Il *rango* di A è l'ordine massimo dei minori che si possono estrarre da A con determinante diverso da zero.

Si verifica che rango $A \leq min(m,n)$.
Se A è quadrata rango $A = n$ se e solo se $\det A \neq 0$
(equivalentemente $\det A = 0$ se e solo se rango $A < n$)

Esempio: Sia $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$

A ha quattro righe e tre colonne quindi rango $A \leq 3$. I minori di ordine 3 sono:
$m_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$, $m_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix}$, $m_3 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$, $m_4 = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$.

$\det m_1 = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} - 1 \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} + 0 = 1 - 1 = 0$

$\det m_2 = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{vmatrix} = 0$

$\det m_3 = \begin{vmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{vmatrix} = 1 \begin{vmatrix} 0 & 1 \\ 2 & 2 \end{vmatrix} - 1 \begin{vmatrix} -1 & 1 \\ 0 & 2 \end{vmatrix} + 0 = -2 + 2 = 0$

$\det m_4 = \begin{vmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{vmatrix} = -1(-1) \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$

$\det m_1 = \det m_2 = \det m_3 = \det m_4 = 0$ quindi rango $A < 3$

C'è un minore di ordine due con determinante non nullo: $\det \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq 0$ quindi rango $A = 2$

Calcolo del rango della matrice A usando la riduzione

$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow A$ ha rango 2

Assiomi di definizione:

Uno *spazio vettoriale su un campo* $\mathbf{K}$ è un insieme V dove sono definite due operazioni: addizione tra elementi di V e moltiplicazione per elementi di $\mathbf{K}$ indicati con "+" e "." rispettivamente che soddisfano le seguente proprietà

1) $v+w = w+v \ \forall v, w \in V$ (commutativa)

2) $u+(v+w) = (u+v)+w \ \forall u, v, w \in V$ (associativa)

3) $\exists\ 0_V \in V / 0_V+v = v \ \forall v \in V$ (esistenza elemento neutro)

4) $\forall\ v \in V \exists u \in V / v+u = 0_V$ e si denota $u = -v$ (esistenza elemento opposto)

5) $a.(u+v) = a.u+a.v \ \forall a \in \mathbf{K} \forall v \in V$ (distributiva)

6) $(a+b)v = a.v+b.v \ \forall a, b \in \mathbf{K} \forall v \in V$ (distributiva)

7) $(ab)v = a.(b.v) \forall a, b \in \mathbf{K} \forall v \in V$

8) $1.v = v$ dove $1 \in \mathbf{K} \forall v \in V$

Proprietà di $(V,\ \mathbf{K},\ +,\ \cdot)$

a) l'elemento neutro $\vec{0}$ è unico,
b) l'opposto di un dato vettore è unico,
c) $0 \cdot v = \vec{0}$,
d) $a \cdot \vec{0} = \vec{0}$,
e) se $a \cdot v = \vec{0}$ allora $a = 0$ oppure $v = \vec{0}$ (legge dell'annullamento)
f) l'opposto di $a \cdot v$ è $(-a) \cdot v \ = a \cdot (-v) = \ -a \cdot v$.

Dimostrazione a) Supporre che esistono due elementi neutri $\vec{0}$ e $\vec{0}'$
allora $\vec{0}' = \vec{0} + \vec{0}' = \vec{0}$
la prima uguaglianza si verifica per che $\vec{0}$ è neutro e la seconda per che anche $\vec{0}'$ lo è.

b) Supporre che esistono due elementi opposti w e w' per v
allora $\vec{0} = w + v$
sommando w' si ottiene $w' = \vec{0} + w' = (w+v) + w' = w + (v + w') = w + \vec{0} = w$.

c) $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$ quindi $0 \cdot v = 0 \cdot v + 0 \cdot v$
sommando l'opposto $-(0 \cdot v)$ si ha
$\vec{0} = -(0 \cdot v) + 0 \cdot v = -(0 \cdot v) + [0 \cdot v + 0 \cdot v]$.
ma $-(0 \cdot v) + [0 \cdot v + 0 \cdot v] = [-(0 \cdot v) + 0 \cdot v] + 0 \cdot v = \vec{0} + 0 \cdot v = 0 \cdot v$.
si ha quindi che $\vec{0} = 0 \cdot v$.

d) $a \cdot \vec{0} = a \cdot (\vec{0} + \vec{0}) = a \cdot \vec{0} + a \cdot \vec{0}$
sommando l'opposto $-(a\vec{0})$ si ha che
$\vec{0} = -(a \cdot \vec{0}) + a \cdot \vec{0} = -(a \cdot \vec{0}) + [a \cdot \vec{0} + a \cdot \vec{0}] = [-(a \cdot \vec{0}) + a \cdot \vec{0}] + a \cdot \vec{0} = a \cdot \vec{0}$.

e) Se $a \neq 0$ allora esiste l'inverso $a^{-1} \in \mathbf{K}$, si ha dunque che
$v = 1 \cdot v = (a^{-1}a) \cdot v = a^{-1} \cdot (a \cdot v) = a^{-1} \cdot \vec{0} = \vec{0}$.

Verifica di f)
$a \cdot v + (-a) \cdot v = (a-a) \cdot v = 0 \cdot v = \vec{0}$.
$a \cdot v + a \cdot (-v) = a \cdot (v - v) = a\vec{0} = \vec{0}$.
Per unicità dell'opposto $-(a \cdot v)$ si ha che
$(-a) \cdot v = -(a \cdot v)$ e che $a \cdot (-v) = -(a \cdot v)$.

Esempi: $\mathbf{R}^n$, $M(n,m,\mathbf{R})$, $\mathbf{R}[x]$, $\mathbf{R}_n[x]$, con le operazioni di somma e prodotto per uno scalare definite prima, sono spazi vettoriali sul campo $\mathbf{R}$.

SOTTOSPAZI Un insieme non vuoto W dello spazio vettoriale $(V,\mathbf{K},+,.)$ è un sottospazio vettoriale di V se con le stesse leggi di composizione di V è uno spazio vettoriale.

Teorema: $W \subset V$ è sottospazio vettoriale di V se e solo se a) $\forall v, w \in W \Rightarrow v+w \in W$ e b) $\forall a \in \mathbf{K}\ \forall v \in W \Rightarrow av \in W$.

COMBINAZIONE LINEARE Un vettore $v \in V$ è combinazione lineare dei vettori $v_1, \ldots, v_k \in V$ se esistono $a_1, \ldots, a_k \in \mathbf{K}$ tale che si scrive $v = a_1v_1 + \cdots + a_kv_k$.

GENERATORI L'insieme di tutte le combinazioni lineari di $v_1, \ldots, v_k$ si chiama spazio generato da $v_1, \ldots, v_k$ e si indica con $\text{Span}\{v_1, \ldots, v_k\}$. ($v \in \text{Span}\ \{v_1, \ldots, v_k\} \Longleftrightarrow v = a_1v_1 + \cdots + a_kv_k$)

Esempio: Il sottospazio generato dalle righe $R_1, \ldots, R_i, \ldots, R_j, \ldots, R_k$ di una matrice coincide col sottospazio generato da $R_1, \ldots, R_i + aR_j, \ldots, R_j, \ldots, R_k$, cioè
$\text{Span}\{R_1, \ldots, R_i, \ldots, R_j, \ldots, R_k\} = \text{Span}\{R_1, \ldots, R_i + aR_j, \ldots, R_j, \ldots, R_k\}$.

Teorema: $\text{Span}\{v_1, \ldots, v_k\}$ è il più piccolo sottospazio di V che contiene $v_1, \ldots, v_k$.
cioè: 1) $v_i \in \text{Span}\{v_1, \ldots, v_k\}$, 2) $\text{Span}\{v_1, \ldots, v_k\}$ è un sottospazio vettoriale di V, 3) se $W \subset V$ è un sottospazio tale che $v_i \in W$ $i = 1, \ldots, k$ allora $\text{Span}\{v_1, \ldots, v_k\} \subset W$.

Definizione: Sia $W \subset V$. Si dice che W è generato da $v_1, \ldots, v_k$ (o che $v_1, \ldots, v_k$ è un insieme di generatori di W) se $W = \text{Span}\{v_1, \ldots, v_k\}$.

Definizione: Sia V uno spazio vettoriale sul campo $\mathbf{K}$, se esistono $v_1, \ldots, v_k \in V$ che generano V si dice che V è finitamente generato.

INDIPENDENZA LINEARE La collezione di vettori $\{v_1, \ldots, v_k\}$ è detta linearmente indipendente se e solo se $\vec{0} = a_1v_1 + \cdots + a_kv_k \Longrightarrow a_i = 0\ \forall i = 1, \ldots, k$. La collezione $\{v_1, \ldots, v_k\}$ è detta linearmente dipendente se esistono $a_1, \ldots, a_k \in \mathbf{K}$ non tutti nulli tale che $a_1v_1 + \cdots + a_kv_k = \vec{0}$.

Lemma: La collezione $\{v_1, \ldots, v_k\}$ è linearmente indipendenti se e solo se ogni loro combinazione lineare si scrive in modo unico.

Dimostrazione: $\Rightarrow$) Se $a_1v_1 + \cdots + a_kv_k = b_1v_1 + \cdots + b_kv_k$ quindi $(a_1 - b_1)v_1 + \cdots + (a_k - b_k)v_k = \vec{0}$, siccome la collezione di vettori è linearmente indipendenti allora $a_1 = b_1, \ldots, a_k = b_k$.

$\Leftarrow$) Se $a_1v_1 + \cdots + a_kv_k = \vec{0} = 0v_1 + \cdots + 0v_k$, siccome il modo di scrivere una combinazione lineare è unico allora $a_1 = 0, \ldots, a_k = 0$.

Esempio: $\{v\}$ è linearmente indipendente se e solo se $v \neq \vec{0}$.
Esempio: le righe non nulle di una matrice ridotta per righe sono linearmente indipendenti.

Verifica: Possiamo supporre che la matrice A sia ridotta a scala

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\rho} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2\rho} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{\rho\rho} & \cdots & a_{\rho n} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$ dove i coefficienti $a_{ii} \neq 0\ \forall i = 1, \ldots, \rho$.

Se $\alpha_1R_1 + \alpha_2R_2 + \cdots + \alpha_\rho R_\rho = \vec{0}$ allora si ha che

$\alpha_1 a_{11} = 0$
$\alpha_2 a_{22} + \alpha_1 a_{12} = 0$
$\dots$
$\alpha_\rho a_{\rho\rho} + \dots + \alpha_2 a_{2\rho} + \alpha_1 a_{1\rho} = 0.$
Quindi $a_{11} \neq 0 \Rightarrow \alpha_1 = 0,\ a_{22} \neq 0 \Rightarrow \alpha_2 = 0,\ \dots, a_{\rho\rho} \neq 0 \Rightarrow \alpha_\rho = 0$

BASE e DIMENSIONE Sia V uno spazio vettoriale finitamente generato.
Definizione: Una *base* $\mathcal{B}$ di V è un insieme ordinato $\{v_1, \dots, v_n\}$ di generatori di V linearmente indipendente.

Teorema: $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V se ogni elemento di V si scrive in modo unico nella forma $v = a_1 v_1 + \dots + a_n v_n$.

Esempio: le righe non nulle di una matrice ridotta per righe formano una base dello spazio delle righe di A e il loro numero è il rango A.

Definizione: Le *componenti* di un vettore v rispetto a una base $\mathcal{B} = \{v_1, \dots, v_n\}$ sono gli n numeri $a_1, \dots, a_n$ tali che $v = a_1 v_1 + \dots + a_n v_n$ presi nel ordine.

Quindi fissata una base $\mathcal{B}$ si fa corrispondere al vettore $v \in V$ la n-upla $v_{\mathcal{B}} = (a_1, \dots, a_n) \in \mathbf{K}^n$.

Teorema (di esistenza di una base): Ogni insieme finito di generatori di V contiene una base di V.

Dimostrazione: metodo degli scarti successivi: 1) si scartano tutti i vettori nulli, 2) si scarta il primo dei v_i che è combinazione lineare dei precedenti, 3) si ripete il procedimento fino ad ottenere un insieme linearmente indipendente.

Lemma di Steinitz: Se $\{v_1, \dots, v_k\}$ è un insieme di generatori di V e $\{w_1, \dots, w_h\}$ è un insieme linearmente indipendente allora $h \leq k$.

Dimostrazione: per induzione sul numero di vettori LI.

Se $h = 1$, per ipotesi esistono $a_1, \dots, a_k$ tali che $w_1 = a_1 v_1 + \dots + a_k v_k$, siccome $\{w_1\}$ è LI si ha che $w_1 \neq \vec{0}$, allora esiste almeno un i per cui $a_i v_i \neq \vec{0}$, quindi $\exists i|\ v_i \neq \vec{0}$ il che implica che $k \geq 1 = h$. Quindi per $h = 1$ il lemma è verificato.

Si può supporre che $a_1 v_1 \neq \vec{0}$, in questo caso si ha che $v_1 = \frac{1}{a_1}(w_1 - a_2 v_2 - \dots - a_k v_k) \in \mathrm{Span}\{w_1, \dots, v_k\}$. Quindi $v_1 \in \mathrm{Span}\{w_1, v_2, \dots, v_k\} \supset \mathrm{Span}\{v_1, v_2, \dots, v_k\} = V$. Si ha dunque che $V = \mathrm{Span}\{w_1, \dots, v_k\}$ (cioè, v_1 è stato sostituito da w_1).

Possiamo suporre il lemma valido per $h - 1$, in particolare che $v_1, \dots, v_{h-1}$ possano essere sostituiti da $w_1, \dots, w_{h-1}$, vale a dire che $V = \mathrm{Span}\{w_1, \dots, w_{h-1}, v_h, \dots, v_k\}$. Per ipotesi induttiva esistono dei coefficienti $b_1, \dots, b_k$ tali che $w_h = b_1 w_1 + \dots + b_{h-1} w_{h-1} + b_h v_h + \dots + b_k v_k$. Nel caso in cui i vettori $v_h, \dots, v_k$ fossero tutti uguali a $\vec{0}$ il vettore w_h sarebbe combinazione lineare di $w_1, \dots, w_{h-1}$, il che contradice l'ipotesi di indipendenza lineare, quindi si ha che $k \geq h$.

Corollario: Sia $\mathcal{B} = \{u_1, \dots, u_n\}$ una base di V e siano $w_1, \dots, w_m$ m-vettori in V. Se $m > n$ allora $\{w_1, \dots, w_m\}$ è linearmente dipendente (LD).

Teorema: Due basi qualsiasi di uno spazio vettoriale hanno la stessa cardinalità.

Dimostrazione: Se $\mathcal{B} = \{u_1, \dots, u_n\}$ e $\mathcal{B}' = \{u'_1, \dots, u'_m\}$ sono due basi di V, si applica due volte il lemma di Steinitz. Siccome $V = \mathrm{Span}\{u_1, \dots, u_n\}$ e $\{u'_1, \dots, u'_m\}$ è LI allora $n \geq m$ e siccome $V = \mathrm{Span}\{u'_1, \dots, u'_m\}$ e $\{u_1, \dots, u_n\}$ è LI allora $n \leq m$, quindi $m = n$.

Definizione: La *dimensione* di uno spazio vettoriale V è la cardinalità di una sua (qualsiasi) base. Se $V = \{\vec{0}\}$ si definisce $\dim V = 0$.

Esempio: Sia $M(2,\mathbf{R})$ lo spazio vettoriale delle matrici quadrate reali di 2^o ordine.
Le matrici $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ generano $M(2,\mathbf{R})$ e inoltre sono linearmente indipendenti. Quindi la dimensione di $M(2,\mathbf{R})$ è 4.

Esempio: Sia $\mathbf{R}_n[x]$ lo spazio dei polinomi a coefficienti reali di grado minore o uguale a 2.
I polinomi $1,\ x, \ldots,\ x^n$ generano $\mathbf{R}_n[x]$ e inoltre sono linearmente indipendenti. Quindi la dimensione di $\mathbf{R}_n[x]$ è $n+1$.

Teorema (del completamento a una base): Ogni insieme di vettori linearmente indipendente è contenuto in una base. Vale a dire, se $\{w_1, \ldots, w_h\}$ è linearmente indipendente, allora è possibile trovare vettori $v_{h+1}, \ldots, v_n \in V$ in modo che $\{w_1, \ldots, w_h, v_{h+1}, \ldots, v_k\}$ sia una base di V.

Dimostrazione: se $V = \text{Span}\{v_1, \ldots, v_k\}$ allora si può estrarre una base col metodo degli scarti successivi dall'insieme di generatori $\{w_1, \ldots, w_h, v_1, \ldots, v_k\}$ senza scartare alcun w_i.

Alternativamente, si procede come nella dimostrazione del Lemma di Steinitz, dove i primi $v_1, \ldots, v_h$ sono stati sostituiti dai $w_1, \ldots, w_h$. Se la collezione $\{v_1, v_2, \ldots, v_k\}$ è linearmente indipendente, per dimostrare che $\{w_1, \ldots, w_h, v_{h+1}, \ldots, v_k\}$ è linearmente indipendente si fa per induzione su h, dome segue.

Per $h = 1$, se $\vec{0} = \alpha_1 w_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = \alpha_1(a_1 v_1 + \ldots + a_k v_k) + \alpha_2 v_2 + \ldots + \alpha_k v_k = \alpha_1 a_1 v_1 + (\alpha_1 a_2 + \alpha_2) v_2 + \ldots + (\alpha_1 a_k + \alpha_k) v_k$ implica che $\alpha_1 a_1 = \alpha_1 a_2 + \alpha_2 = \cdots = \alpha_1 a_k + \alpha_k = 0$. Se $a_1 \neq 0$ allora $\alpha_1 = 0$ e quindi $\alpha_2 = \cdots = \alpha_k = 0$.

Il passo induttivo si fa in modo analogo: supponendo che $\{w_1, \ldots, w_{h-1}, v_h, \ldots, v_k\}$ sia linearmente indipendente si deve dimostrare che $\{w_1, \ldots, w_h, v_{h+1}, \ldots, v_k\}$ è linearmente indipendente.

Siccome $V =$ Span $\{w_1, \ldots, w_{h-1}, v_h, \ldots, v_k\}$ allora
$w_h = a_1 w_1 + \ldots + a_{h-1} w_{h-1} + a_h v_h + \ldots + a_k v_k$. Si può suporre che $a_h \neq 0$.
$\vec{0} = \alpha_1 w_1 + \ldots + \alpha_{h-1} w_{h-1} + \alpha_h w_h + \ldots + \alpha_k v_k$
$= \alpha_1 w_1 + \ldots + \alpha_{h-1} w_{h-1} + \alpha_h(a_1 w_1 + \ldots + a_{h-1} w_{h-1} + a_h v_h + \ldots + a_k v_k) + \alpha_{k+1} v_{k+1} + \ldots + \alpha_k v_k$
$= (\alpha_1 a_h + \alpha_1) w_1 + \ldots + (\alpha_{h-1} a_h + \alpha_{h-1}) w_{h-1} + + \alpha_h a_h v_h + \ldots + (\alpha_h a_k + \alpha_k) v_k$
l'ipotesi indutiva implica che $\alpha_h a_1 + \alpha_1 = \cdots = \alpha_h a_h = \ldots = \alpha_h a_k + \alpha_k = 0$.
Siccome $a_h \neq 0$ allora $\alpha_h = 0$ e quindi $\alpha_1 = \cdots = \alpha_k = 0$.

Nota: La dimensione di V è pari al
a) numero minimo di generatori di V.
b) numero massimo di vettori linearmente indipendenti di V.

Lemma: Se W è un sottospazio di V allora $\dim W \leq \dim V$.

Teorema: Sia A una matrice con m righe e n colonne. Sia $\mathcal{R} \subset \mathbf{K}^n$ lo spazio generato dalle m righe della matrice A e sia $\mathcal{C} \subset \mathbf{K}^m$ lo spazio generato dalle n colonne della matrice A, allora $\dim \mathcal{R} = \dim \mathcal{C} =$ Rango A.

Dimostrazione: Sia $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} A^1 & \cdots & A^n \end{pmatrix}$

Siccome $\forall i \ A_i = \begin{pmatrix} a_{i1} & \ldots & a_{in} \end{pmatrix} \in \mathbf{R}^n$ si ha che $\mathcal{R} = \text{Span}\{A_1, \ldots, A_m\} \subset \mathbf{R}^n$
e $\forall j \ A^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbf{R}^m$ si ha che $\mathcal{C} = \text{Span}\{A^1, \ldots, A^n\} \subset \mathbf{R}^m$.

Sia $k = \dim \mathcal{R}$ e sia $\mathcal{B}_{\mathcal{R}} = \{R_1, \ldots, R_k\}$ una base de $\mathcal{R}$, dove
$R_1 = (b_{11}, \ldots, b_{1n})^t, \cdots, R_k = (b_{k1}, \ldots, b_{kn})^t$.

In questa base le righe si scrivono:

$$A_1 = c_{11}R_1 + \ldots + c_{1k}R_k = c_{11}\begin{pmatrix} b_{11} \\ \vdots \\ b_{1n} \end{pmatrix} + \ldots + c_{1k}\begin{pmatrix} b_{k1} \\ \vdots \\ b_{kn} \end{pmatrix}$$

$$\vdots$$

$$A_m = c_{m1}R_1 + \ldots + c_{mk}R_k = c_{m1}\begin{pmatrix} b_{11} \\ \vdots \\ b_{1n} \end{pmatrix} + \ldots + c_{mk}\begin{pmatrix} b_{k1} \\ \vdots \\ b_{kn} \end{pmatrix}$$

Da queste espressioni si ottiene che le colonne di A si scrivono come segue: la prima colonna è

$$A^1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} = \begin{pmatrix} c_{11}b_{11} + \ldots + c_{1k}b_{k1} \\ \vdots \\ c_{m1}b_{11} + \ldots + c_{mk}b_{k1} \end{pmatrix}$$

la n-esima colonna è

$$A^n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} c_{11}b_{1n} + \ldots + c_{1k}b_{kn} \\ \vdots \\ c_{m1}b_{1n} + \ldots + c_{mk}b_{kn} \end{pmatrix}$$

che si possono riscrivere come

$$A^1 = b_{11}\begin{pmatrix} c_{11} \\ \vdots \\ c_{m1} \end{pmatrix} + \ldots + b_{k1}\begin{pmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{pmatrix}$$

$$\vdots$$

$$A^n = b_{1n}\begin{pmatrix} c_{11} \\ \vdots \\ c_{m1} \end{pmatrix} + \ldots + b_{kn}\begin{pmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{pmatrix}$$

Ponendo $C^1 = \begin{pmatrix} c_{11} \\ \vdots \\ c_{m1} \end{pmatrix}, \ldots, C^k = \begin{pmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{pmatrix}$ si ha che $A^1, \ldots, A^n \in \mathrm{Span}\{C^1, \ldots, C^k\}$.

Quindi $\mathcal{C} = \mathrm{Span}\{C^1, \ldots, C^k\}$ di conseguenza $\dim \mathcal{C} \le k = \dim \mathcal{R}$.

Invertendo i ruoli di $\mathcal{C}$ e $\mathcal{R}$ in modo analogo si dimostra che $\dim \mathcal{R} \le \dim \mathcal{C}$.
Quindi $\dim \mathcal{R} = \dim \mathcal{C}$.

APPLICAZIONI

I) La dimensione di un sottospazio vettoriale $W = \mathrm{Span}\{v_1, \ldots, v_k\}$ di uno spazio vettoriale V è il rango della matrice che ha per righe (o per colonne) le componenti dei vettori $v_1, \ldots, v_k$ in una base di V.

II) $w \in \mathrm{Span}\{v_1, \ldots, v_k\}$ se e solo se $\dim \mathrm{Span}\{v_1, \ldots, v_k\} = \dim \mathrm{Span}\{v_1, \ldots, v_k, w\}$.

INTERSEZIONE, UNIONE, SOMMA, SOMMA DIRETTA di SOTTOSPAZI

Dati due insiemi U e W contenuti in V, si definiscono gli insiemi
intersezione $U \cap W := \{v \in V \text{ tali che } v \in U \ \text{ e } \ v \in W\}$,
unione $U \cup W := \{v \in V \text{ tali che } v \in U \ \text{ oppure } \ v \in W\}$ e
somma $U + W := \{v \in V \text{ tali che } \exists u \in U, w \in W : v = u + w\}$.

Lemma: Se U e W sono sottospazi dello spazio vettoriale V allora gli insiemi $U \cap W$ e $U + W$ sono sottospazi vettoriali di V.

Definizione: La somma $U + W$ è *diretta* se ogni vettore del sottospazio somma si puó scrivere in modo unico nella forma $u + w$ dove $u \in U, w \in W$. In tal caso si scrive $U + W = U \oplus W$.

Teorema (Caratterizzazione): $U + W = U \oplus W \Leftrightarrow U \cap W = \{\vec{0}\}$.

Dimostrazione: $\Leftarrow$) $v = u + w = u' + w'$ implica $U \ni u - u' = w' - w \in W$ quindi $u - u' = w' - w = \vec{0} \in U \cap W$.

$\Rightarrow$) Se $\exists v \neq \vec{0} \in U \cap W$ allora il modo di scrivere v non è unico ad esempio $v = 1v + \alpha\vec{0} = \beta\vec{0} + 1v = 2v - v$.

Lemma: Se $\mathcal{B}$ è un insieme di generatori di U e $\mathcal{C}$ è un sistema di generatori di W allora $\mathcal{B} \cup \mathcal{C}$ è un insieme di generatori di $U + W$.

Dimostrazione: se $u = a_1u_1 + \cdots + a_su_s$ e $w = b_1w_1 + \cdots + b_rw_r$ allora
$u + w = a_1u_1 + \cdots + a_su_s + b_1w_1 + \cdots + b_rw_r$

Teorema (Formula di Grassman): $\dim(U + W) + \dim(U \cap W) = \dim U + \dim W$.

Dimostrazione: Sia $k = \dim(U \cap W)$ e sia$\{v_1, \ldots, v_k\}$ una base di $U \cap W$,
completiamo la base di $U \cap W$ a una base $\{v_1, \ldots, v_k, u_{k+1}, \ldots, u_p\}$ di U, dove $p = \dim U$,
completiamo la base di $U \cap W$ a una base $\{v_1, \ldots, v_k, w_{k+1}, \ldots, w_q\}$ di W, dove $q = \dim W$.
Basta dimostrare che $\{v_1, \ldots, v_k, u_{k+1}, \ldots, u_p, w_{k+1}, \ldots, w_q\}$ è una base di $U + W$ (per il lemma precedente è un insieme di generatori di $U + W$, rimane da verificare che è LI).

Corollario: $U + W = U \oplus W \Leftrightarrow \dim(U + W) = \dim U + \dim W$.

ESEMPIO 1: In $V = \mathbf{R}^3$, si considerino i sottospazi vettoriali $U = \{(x, y, z) \in V \ : x = z = 0\}$ e $W = \{(x, y, z) \in V \ : x = y = 0\}$. $U \cap W = \{(x, y, z) \in V \ : x = y = z = 0\} = \{\vec{0}\}$, $U + W = \{(x, y, z) \in V \ : x = 0\}$. Mentre $U \cup W = \{(x, y, z) \in V \ : x = 0 \text{ e } yz = 0\}$ non è sottospazio vettoriale di V. Siccome $\dim(U \cap W) = 0$ si ha che $U + W = U \oplus W$, si osservi inoltre che $U + W \neq V$.

Notare che $U = \text{Span}\left\{\begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ e $W = \text{Span}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$, quindi $U + W = \text{Span}\left\{\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$.

Risoluzione dei sistemi lineari col metodo di riduzione

Una equazione lineare è una espressione del tipo $a_1x_1 + \cdots + a_nx_n = b$. Gli a_i sono i coefficienti della equazione, gli x_i sono le incognite e b è il termine noto. Una *soluzione* della equazione è una n-upla $(\alpha_1, \ldots, \alpha_n) \in \mathbf{K}^n$ che la verifica.

Una collezione di equazioni $\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{cases}$

è chiamata sistema di m-equazioni lineari ed n-incognite.

Una *soluzione* del sistema lineare è una n-upla $(\alpha_1, \ldots, \alpha_n) \in \mathbf{K}^n$ che verifica tutte le m-equazioni.
Definizioni: Un sistema di equazioni lineari è *risolubile o compatibile* se e solo se ammette soluzioni.

RAPPRESENTAZIONE MATRICIALE

Se $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ è la matrice dei coefficienti del sistema, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ e $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sono la colonna dei termini noti e la colonna delle incognite, rispettivamente, utilizzando il prodotto di matrici il sistema di equazioni lineari si scrive come $AX = B$.

La matrice $(A|B) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$ è detta la matrice completa del sistema.

Se si denota con A^j la colonna j-esima della matrice A il sistema $AX = B$ può scriversi come una equazione lineare a coefficienti vettoriali

$$A^1x_1 + \cdots + A^nx_n = B$$

Lemma: Il sistema $AX = B$ è risolubile (o compatibile) se e solo se B è combinazione lineare di $A^1, \ldots, A^n$. (cioé $B \in \mathrm{Span}\{A^1, \ldots, A^n\}$).
Due sistemi lineari $AX = B$ e $A'X = B'$ si dicono *equivalenti* se hanno esattamente le stesse soluzioni.
Un sistema lineare $AX = B$ si dice *ridotto per righe* se la matrice dei coefficienti A è ridotta per righe. Il seguente lemma mostra che ogni sistema lineare è equivalente a uno ridotto per righe

Lemma: Se $k \neq 0$ il sistema lineare $\begin{cases} a_1x_1 + \cdots + a_nx_n = a \\ b_1x_1 + \cdots + b_nx_n = b \end{cases}$ è equivalente al sistema lineare

$$\begin{cases} a_1x_1 + \cdots + a_nx_n & = a \\ h(a_1x_1 + \cdots + a_nx_n) + k(b_1x_1 + \cdots + b_nx_n) & = ha + kb. \end{cases}$$

Teorema (Rouché-Capelli sulla risolubilità):
Il sistema $AX = B$ è risolubile se e solo se rango A = rango $(A|B)$.

Dimostrazione: $\Rightarrow$) Se il sistema $AX = B$ è risolubile allora B è combinazione lineare delle colonne $A^1, \ldots, A^n$ di A. Quindi la matrice $(A|B)$ che si ottiene aggiungendo ad A la colonna B ha lo stesso rango di A.

$\Leftarrow$) Viceversa, si osservi che $\mathrm{Span}\{A^1, \ldots, A^n\} \subset \mathrm{Span}\{A^1, \ldots, A^n, B\}$. Inoltre rango $A = \dim \mathrm{Span}\ \{A^1, \ldots, A^n\}$ e rango $(A|B) = \dim \mathrm{Span}\ \{A^1, \ldots, A^n, B\}$. Quindi per ipotesi questi sottospazi hanno la stessa dimensione, siccome l'uno è contenuto nell'altro allora coincidono. Si ha dunque che B è combinazione lineare di $A^1, \ldots, A^n$. Cioè, il sistema è risolubile.

Teorema (Rouché-Capelli sul numero di soluzioni dei sistemi risolubili):

Sia $AX = B$ un sistema risolubile e sia ρ il rango di A ($=$ rango $(A|B)$), allora il sistema ha $\infty^{n-\rho}$ soluzioni che dipendono di $n-\rho$ incognite libere.
Se $n = \rho$ c'e una e solo una soluzione,
se $n > \rho$ c'e ne sono infinite soluzioni che dipendono di $n-\rho$ parametri liberi.

Dimostrazione: Possiamo supporre che la matrice $(A|B)$ sia ridotta a scala

$$\left(\begin{array}{ccccccc|c} a_{11} & a_{12} & \cdots & a_{1\rho} & a_{1\rho+1} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & \cdots & a_{2\rho} & a_{2\rho+1} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{\rho\rho} & a_{\rho\rho+1} & \cdots & a_{\rho n} & b_\rho \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right),$$

dove i coefficienti $a_{ii} \neq 0 \; \forall i = 1, \ldots, \rho$.
Da l'ultima equazione del sistema (ultima riga non nulla) $a_{\rho\rho}x_\rho + a_{\rho\rho+1} + x_{\rho+1} + \cdots + a_{\rho n}x_n = b_\rho$ si ricava $x_\rho = \frac{1}{a_{\rho\rho}}(b_\rho - a_{\rho\rho+1}x_{\rho+1} + \cdots - a_{\rho n}x_n)$. Quindi la variabile x_ρ dipende (è funzione) dalle variabile $x_{\rho+1}, \cdots, x_n$, cioè $x_\rho = f_\rho(x_{\rho+1}, \cdots, x_n)$.
Analogamente dalla penultima equazione si ricava $x_{\rho-1} = \frac{1}{a_{\rho-1\rho-1}}(b_{\rho-1} - a_{\rho-1\rho}x_\rho + \cdots - a_{\rho-1n}x_n)$.
Sostituendo x_ρ risulta che $x_{\rho-1}$ dipende dalle variabile (libere) $x_{\rho+1}, \cdots, x_n$,
cioè $x_{\rho-1} = f_{\rho-1}(x_{\rho+1}, \cdots, x_n)$. Si prosegue fino a ricavare $x_1 = f_1(x_{\rho+1}, \cdots, x_n)$.
Le variabili $x_1, \cdots, x_\rho$ dipendono dalle variabili libere $x_{\rho+1}, \cdots, x_n$. Dando valori alle variabili indipendenti si ottengono le seguenti soluzioni del sistema.
Ponendo $x_{\rho+1} = 1, x_{\rho+2} = 0, \cdots, x_n = 0$ si ottiene la soluzione
$s_1 = (f_1(1,0,\cdots,0), f_2(1,0,\cdots,0), \ldots, f_\rho(1,0,\cdots,0), 1, 0, \cdots, 0)$.
Ponendo $x_{\rho+1} = 0, x_{\rho+2} = 1, \cdots, x_n = 0$ si ottiene la soluzione
$s_2 = (f_1(0,1,\cdots,0), f_2(0,1,\cdots,0), \ldots, f_\rho(0,1,\cdots,0), 0, 1, \cdots, 0)$.
Ponendo $x_{\rho+1} = 0, x_{\rho+2} = 0, \cdots, x_n = 1$ si ottiene la soluzione
$s_{n-\rho} = (f_1(0,0,\cdots,1), f_2(0,0,\cdots,1), \ldots, f_\rho(0,0,\cdots,1), 0, 0, \cdots, 1)$.
Per costruzione le soluzioni risultano essere linearmente indipendenti.

SISTEMI OMOGENEI

Sia $A \in M(m,n;\mathbf{K})$. Un sistema lineare $AX = B$ si dice *omogeneo* se B è il vettore nullo. Si osservi che un sistema omogeneo è compatibile ($\mathbf{0}$ è soluzione).

Teorema: L'insieme delle soluzioni del sistema $AX = \mathbf{0}$ è un sottospazio vettoriale di $\mathbf{K}^n$.

Teorema: Lo spazio delle soluzioni del sistema $AX = \mathbf{0}$ ha dimensione $n-\rho$, dove n è il numero delle incognite e ρ è il rango di A.

Dimostrazione: Siccome il sistema omogeneo è risolubile le soluzioni si costruiscono come prima. Le soluzioni $\{s_1, \cdots, s_{n-\rho}\}$ generano il sottospazio delle soluzioni del sistema omogeneo $\{Z \in \mathbf{K}^n : \; AZ = 0\}$. La matrice che ha per righe le soluzioni $s_1, \cdots, s_{n-\rho}$ è (per costruzione) ridotta per righe, quindi $\{s_1, \cdots, s_{n-\rho}\}$ è linearmente indipendente, dunque costituisce una base del sottospazio delle soluzioni del sistema $AX = \mathbf{0}$ per cui la dimensione è $n-\rho$.

SOLUZIONE GENERALE DI UN SISTEMA LINEARE

Teorema: Ogni soluzione del sistema $AX = B$ è della forma $Y = Y_0 + Z$ dove Y_0 è una soluzione particolare del sistema $AX = B$, e Z è una soluzione del sistema omogeneo $AX = \mathbf{0}$.

L'insieme delle soluzioni del sistema $AX = B$ è un sottospazio affine, cioè è della forma

$$Y_0 + \{Z \in \mathbf{K}^n : \ AZ = 0\}.$$

RANGO E DETERMINANTE

Sia A una matrice quadrata di ordine n.

Teorema (Regola di Cramer): Se A è invertibile allora il sistema lineare $AX = B$ ammette una unica soluzione:

$$X = \frac{1}{\det A} \begin{pmatrix} \Delta_1 \\ \vdots \\ \Delta_n \end{pmatrix}$$

dove Δ_i è il determinante della matrice ottenuta da A sostituendo la colonna i-esima A^i con la colonna B dei termini noti:

$$\Delta_1 = \det \begin{pmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \cdots, \Delta_n = \det \begin{pmatrix} a_{11} & \cdots & a_{1(n-1)} & b_1 \\ \vdots & \ldots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n(n-1)} & b_n \end{pmatrix}$$

Dimostrazione:

moltiplicando $AX = B$ per A^{-1} si ha $A^{-1}(AX) = A^{-1}B \Rightarrow (A^{-1}A)X = IX = A^{-1}B$

Per la definizione di matrice inversa e per le proprietà della trasposta si ha

$X = A^{-1}B \quad = \frac{1}{\det A}(A_{ij})^t B \quad = \frac{1}{\det A}(B^t(A_{ij}))^t$

Molteplicando le matrici

$$\begin{aligned} B^t(A_{ij}) &= \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ldots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \\ &= (b_1 A_{11} + \ldots + b_n A_{n1}, \cdots, b_1 A_{1n} + \ldots + b_n A_{nn}) \\ &= (\Delta_1, \cdots, \Delta_n) \end{aligned}$$

L'ultima uguaglianza si ottiene applicando il Teorema di Laplace.

Definizione: Il *rango* di A è l'ordine massimo dei minori che si possono estrarre da A con determinante diverso da zero.

Corollario: Se A è una matrice quadrata di ordine n, le seguenti affermazioni sono equivalenti:

a) rango $A < n$,
b) le colonne di A sono LD,
c) le righe di A sono LD,
d) $\det A = 0$.

Corollario: Il sistema omogeneo $AX = \vec{0}$ ha soluzioni non nulle se e solo se $\det A = 0$

SOLUZIONE DI UN SISTEMA A INCOGNITE VETTORIALI

Siano $X_i = (x_{i1}, \dots, x_{ip}) \in \mathbf{K}^p,\ i = 1, \dots, n; \quad B_j = (b_{j1}, \dots, b_{jp}) \in \mathbf{K}^p,\ j = 1, \dots, m.$ Si consideri il sistema a incognite vettoriali e termini noti vettoriali

$$\begin{cases} a_{11}X_1 + \cdots + a_{1n}X_n & = & B_1 \\ & \vdots & \\ a_{m1}X_1 + \cdots + a_{mn}X_n & = & B_m \end{cases}$$

Ponendo $X_1, \dots, X_n$ come righe di una matrice $X \in M(n, p, \mathbf{K})$, e $B_1, \dots, B_m$ come righe di una matrice $B \in M(m, p, \mathbf{K})$ il sistema si puó riscrivere in notazione matriciale come $AX = B$.

Esplicitamente

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix} = \begin{pmatrix} b_{11} & \dots & b_{1p} \\ \vdots & \dots & \vdots \\ b_{m1} & \dots & b_{mp} \end{pmatrix}$$

considerando p-sistemi: $AX^i = B^i,\ i = 1, \dots p$ si ha che il sistema $AX = B$ è risolubile se e solo se ogni sistema $AX^i = B^i$ è risolubile, cioè se e solo se rango $A =$ rango $(A|B^i)$.

Teorema: Il sistema a incognite vettoriali $AX = B$ è risolubile se e solo se rango $A =$ rango $(A|B)$.

APPLICAZIONE: Calcolo della matrice inversa col metodo della riduzione per righe.

Se pone $B = I$ dove $I \in M(n, \mathbf{K})$ è la matrice identità e si applica l'ultimo Teorema al sistema a incognite vettoriali

$$AX = I$$

La soluzione del sistema è la matrice A^{-1}, inversa della matrice A.

PRODOTTO SCALARE in $\mathbf{R}^n$.
Definizione: Siano $v = (x_1, \ldots, x_n)$ e $w = (y_1, \ldots, y_n)$ due vettori in $\mathbf{R}^n$, il *prodotto scalare* tra v e w è il numero reale $v.w =< v, w >= x_1 y_1 + \ldots + x_n y_n$.
Proprietà:

1) $< v, w >=< w, v > \ \forall v, w \in \mathbf{R}^n$ (commutativa)
2) $< (u + v), w >=< u, w > + < v, w > \ \forall u, v, w \in \mathbf{R}^n$ (distributiva rispetto alla somma)
3) $\lambda < u, v >=< \lambda u, v >=< u, \lambda v > \ \forall u, v \in \mathbf{R}^n \ \forall \lambda \in \mathbf{R}$ (omogenea)
4) $< v, v >\geq 0, \ \forall\, v \in \mathbf{R}^n; < v, v >= 0 \Leftrightarrow \ v = 0$ (positivitá)

Definizione: La *norma* di un vettore è $v = (x_1, \ldots, x_n)$ è $\|v\| := \sqrt{< v, v >} = \sqrt{x_1^2 + \ldots + x_n^2}$.
La *distanza* tra due punti $P(x_1, \ldots, x_n)$ e $Q(y_1, \ldots, y_n)$ è la lunghezza del segmento $\overline{PQ}$, pari alla norma del vettore $v = \vec{PQ} = (y_1 - x_1, \ldots, y_n - x_n)$ che ha punto iniziale P e punto finale Q:

$$\mathrm{dist}(Q, P) := \|v\| = \sqrt{(y_1 - x_1)^2 + \ldots + (y_n - x_n)^2}$$

Teorema del coseno:
$\forall v, w \in \mathbf{R}^n$ si ha che $< v, w >= \|v\|\|w\| \cos\alpha$, dove α è l'angolo fra i vettori v e w .

Dimostrazione: Si consideri il triangolo di vertici A, B, C, sia α l'angolo tra i vettori $v = \vec{AB}$ e $w = \vec{AC}$. Quindi i lati opposti ai vertici hanno lunghezza $a = \|v - w\|$, $b = \|w\|$, $c = \|v\|$.
Sia H il punto d'intersezione tra la retta che contiene il vettore v e la retta perpendicolare a v che passa per il punto C. Il triangolo CHB è rettangolo, i cateti hanno norma $\|\vec{HC}\| = b\sin\alpha$ e $\|\vec{HB}\| = c - b\cos\alpha$. Quindi per il teorema di Pitagora si ha che

$$\begin{aligned} a^2 &= (b\sin\alpha)^2 + (c - b\cos\alpha)^2 \\ &= b^2\sin^2\alpha + c^2 - 2bc\cos\alpha + b^2\cos^2\alpha \\ &= b^2 - 2bc\cos\alpha + c^2. \end{aligned}$$

Sostituendo a, b, c si ha $\|v - w\|^2 = \|v\|^2 - 2\|v\|\|w\|\cos\alpha + \|w\|^2$
Confrontando con $\|v - w\|^2 =< v - w, v - w >=< v, v > -2 < v, w > + < w, w >$,
si ottiene $< v, w >= \|v\|\|w\|\cos\alpha$.

Corollario: v e w sono perpendicolari se e solo se $< v, w >= 0$.

Corollario (disuguaglianza di Cauchy-Schwarz): $\forall v, w \in \mathbf{R}^n \ |v.w| \leq \|v\| \ \|w\|$.
Dimostrazione: $|\cos\alpha| \leq 1$.

Proposizione:
a) $\forall v, w \in \mathbf{R}^n \ \|v + w\| \leq \|w\| + \|v\|$ (disuguaglianza triangolare).
b) $\forall v \in \mathbf{R}^n \ \forall\lambda \in \mathbf{R} \ \|\lambda v\| = |\lambda| \ \|v\|$.
Dimostrazione: a) $v + w$ è la diagonale del parallelogramma di lati v e w.
b) $\|\lambda v\|^2 = (\lambda x_1)^2 + \ldots + (\lambda x_n)^2 = \lambda^2(x_1^2 + \ldots + x_n^2)$ allora $\|\lambda v\| = \sqrt{\lambda^2}\ \|v\| = |\lambda| \ \|v\|$.

Proposizione: L'area del triangolo di lati $v = (x_1, y_1)$ e $w = (x_2, y_2)$ è $A = \frac{1}{2}\left|\det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\right|$

Dimostrazione: $A = \frac{1}{2}bh = \frac{1}{2}\|v\| \ \|w\| \ \sin\theta$ dove θ è l'angolo tra v e w.

$$\begin{aligned} 4A^2 &= \|v\|^2 \ \|w\|^2 \ \sin^2\theta = \|v\|^2 \ \|w\|^2 \ (1 - \cos^2\theta) \\ &= \|v\|^2 \ \|w\|^2 \ - \|v\|^2 \ \|w\|^2 \ \cos^2\theta \\ &= \|v\|^2 \ \|w\|^2 \ - (< v, w >)^2 \\ &= (x_1^2 + y_1^2)(x_2^2 + y_2^2) - (x_1x_2 + y_1y_2)^2 \\ &= (x_1^2x_2^2 + x_1^2y_2^2 + y_1^2x_2^2 + y_1^2y_2^2) - (x_1^2x_2^2 + 2x_1x_2y_1y_2 + y_1^2y_2^2) \\ &= (x_1y_2 - x_2y_1)^2 \end{aligned}$$

Quindi $2A = |x_1y_2 - x_2y_1|$.

PRODOTTO VETTORIALE IN $\mathbf{R}^3$

Definizione geometrica del prodotto vettoriale nello spazio $\mathbf{R}^3$.

Siano v e $w \in \mathbf{R}^3$ due vettori, il *prodotto vettoriale* $v \wedge w$ è il vettore di $\mathbf{R}^3$ definito come segue:
a) se v e w sono paralleli allora $v \wedge w$ è il vettore nullo (include i casi $v = \vec{0}$ o $w = \vec{0}$)
b) se v e w non sono paralleli allora $v \wedge w$ è il vettore
  i) la cui *direzione* è la retta perpendicolare al piano generato dai vettori v e w,
  ii) il *verso* è ottenuto applicando la regola della mano destra.
  iii) la *norma* $\|v \wedge w\|$ è pari all'area del parallelogramma che ha per lati i vettori v e w

Definizione del prodotto vettoriale $\wedge\colon \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$, $(v, w) \mapsto v \wedge w$ in coordinate
se $v = (x_1, y_1, z_1)$ e $w = (x_2, y_2, z_2)$ allora $v \wedge w =: (y_1z_2 - z_1y_2, -x_1z_2 + z_1x_2, x_1y_2 - y_1x_2)$

Regola mnemotecnica per il calcolo del prodotto vettoriale:

$$v \wedge w = \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$$

Proprietà del prodotto vettoriale

1) è anti-commutativo: $w \wedge v = -v \wedge w$
Esempi: $\vec{i} \wedge \vec{j} = \vec{k}$, $\vec{j} \wedge \vec{k} = \vec{i}$, $\vec{k} \wedge \vec{i} = \vec{j}$
  $\vec{j} \wedge \vec{i} = -\vec{k}$, $\vec{k} \wedge \vec{j} = -\vec{i}$, $\vec{i} \wedge \vec{k} = -\vec{j}$
2) è distributivo rispetto alla somma $(u + v) \wedge w = u \wedge w + v \wedge w$
3) è omogeneo: $(\lambda v) \wedge w = v \wedge (\lambda w) = \lambda(v \wedge w)$
4) NON è associativo.
Esempio: $\vec{0} = (\vec{i} \wedge \vec{i}) \wedge \vec{j} \neq \vec{i} \wedge (\vec{i} \wedge \vec{j}) = \vec{i} \wedge \vec{k} = -\vec{j}$.
5) Inoltre soddisfa

$$u \wedge (v \wedge w) =< u, w > v- < u, v > w$$

6) ldentità di Jacobi:

$$u \wedge (v \wedge w) + v \wedge (w \wedge u) + w \wedge (u \wedge v) = \vec{0}$$

PRODOTTO MISTO IN $\mathbf{R}^3$

Definizione: $\mathbf{R}^3 \times \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$, $(u, v, w) \mapsto < u, v \wedge w > \in \mathbf{R}$.

Se $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ e $w = (x_3, y_3, z_3)$ allora $< u, v \wedge w >:= \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$

Proprietà del prodotto misto:

i) $< u, v \wedge w >= - < u, w \wedge v >$.
ii)] $< u, v \wedge w >=< v, w \wedge u >=< w, u \wedge v >$.

Le seguente affermazioni sono equivalenti:

a) $< u, v \wedge w >= 0$,
b) $u \perp v \wedge w$,
c) u, v e w giacciono sullo stesso piano,
d) $\{u, v, w\}$ è linearmente dipendente.

Definizione: Una base $\{v, w, u\}$ si dice *positivamente orientata* se e solo se il determinante della matrice che ha per righe (o colonne) i vettori v, w, u è positivo.

Dalla definizione del prodotto vettoriale e del prodotto misto (in coordinate) ne segue che:
a) $v \parallel w \Leftrightarrow w \wedge v = \vec{0}$
b) Se $\{v, w\}$ è LI,
i) $< v \wedge w, w >= 0$, $< v \wedge w, v >= 0$.
ii) $\{v, w, v \wedge w\}$ è una base positivamente orientata.
iii) $\|v \wedge w\| = \|v\|\|w\| \sin\alpha$ dove α è l'angolo tra v e w

Verifica di iii):

$$\begin{aligned}
\|v\|^2\|w\|^2 \sin^2\alpha &= \|v\|^2\|w\|^2(1 - \cos^2\alpha) \\
&= \|v\|^2\|w\|^2 - \|v\|^2\|w\|^2 \cos^2\alpha \\
&= \|v\|^2\|w\|^2 - < v, w >^2 \\
&= (x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2) - (x_1x_2 + y_1y_2 + z_1z_2)^2 \\
&= y_1^2z_2^2 - 2y_1z_2z_1y_2 + z_1^2y_2^2 + x_1^2z_2^2 - 2x_1z_2z_1x_2 + z_1^2x_2^2 + x_1^2y_2^2 - 2x_1y_2y_1x_2 + y_1^2x_2^2 \\
&= (y_1z_2 - z_1y_2)^2 + (-x_1z_2 + z_1x_2)^2 + (x_1y_2 - y_1x_2)^2 \\
&= \|v \wedge w\|^2.
\end{aligned}$$

Notare che : $\|v \wedge w\|^2 = \|v\|^2\|w\|^2 - < v, w >^2$

PROIEZIONE ORTOGONALE

Il vettore proiezione $\pi_v(w)$ di un vettore w su una retta l passante per l'origine di $\mathbf{R}^n$ e parallela a un vettore v non nullo è il vettore (parallelo al vettore v)

$$\pi_v(w) = \frac{< w, v >}{\|v\|^2}\, v$$

Il vettore proiezione $\pi_\alpha(w)$ di un vettore w su un piano α passante per l'origine di $\mathbf{R}^n$ e perpendicolare a un vettore n non nullo è il vettore

$$\pi_\alpha(w) = w - \pi_n(w) = w - \frac{< w, n >}{\|n\|^2}\, n.$$

VOLUME

Il volume del parallelepipedo di spigoli u, v e w è uguale all'area della base per l'altezza.
L'altezza è $\|\pi_{v\wedge w}(u)\|$,

$$\pi_{v\wedge w}(u) = \frac{< u, v \wedge w >}{\|v \wedge w\|^2} v \wedge w = \frac{< u, v \wedge w >}{\|v \wedge w\|} \frac{v \wedge w}{\|v \wedge w\|} \Longrightarrow \|\pi_{v\wedge w}(u)\| = \left| \frac{< u, v \wedge w >}{\|v \wedge w\|} \right|$$

L'area della base è $\|v \wedge w\|$, quindi

$$V = |< u, v \wedge w >|\,.$$

Il parallelepipedo di spigoli $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$, $w = (x_3, y_3, z_3)$ ha volume

$$V = \left| \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \right|.$$

EQUAZIONE PARAMETRICA DI UNA RETTA

Sia r la retta passante per P_0 e parallela al vettore v.
Per ogni punto Q (generico) sulla retta r il vettore di componenti $Q - P_0$ (coordinate di Q meno coordinate di P_0) è parallelo al vettore v. Quindi $Q - P_0 = tv$ per un qualche $t \in \mathbf{R}$. Al variare di t si ottiene l'equazione parametrica (vettoriale) della retta r: $Q = P_0 + tv$.

EQUAZIONE CARTESIANA DI UN PIANO IN $\mathbf{R}^3$

Sia π il piano passante per P_0 e perpendicolare al vettore di componenti $n = (a, b, c)$.
Per ogni punto Q (generico) sul piano π il vettore v di componenti $Q - P_0$ (coordinate di Q meno coordinate di P_0) è perpendicolare a n. Ne segue che $v.n = (Q - P_0).n = 0$, per cui l'equazione del piano passante per P_0 è $Q.n = P_0.n$. Chiamando $d = P_0.n$ si ottiene $Q.n = d$. Se le coordinate del punto generico Q sono (x, y, z) si ha l'equazione cartesiana del piano π : $ax + by + cz = d$.

PARALLELISMO E ORTOGONALITA IN $\mathbf{R}^3$:

TRA RETTE

Sia r la retta parallela al vettore v (di equazione vettoriale parametrica $P = P_0 + tv,\ t \in \mathbf{R}$) e sia r la retta parallela al vettore v' (di equazione vettoriale parametrica $P = P_0' + sv',\ s \in \mathbf{R}$).

Definizione: r è parallela a r' ($r \parallel r'$) se e solo se $v \parallel v'$ ($\Leftrightarrow v \wedge v' = 0$).
r è ortogonale a r' ($r \perp r'$) se e solo se $v \perp v'$ ($\Leftrightarrow v.v' = 0$).

Osservazioni: Siano r una retta e Q un punto esterno a r
a) $\exists!$ retta r' tale che $r' \parallel r$ passante per Q.
b) $\exists\infty$ rette r' tale che $r' \perp r$ passante per Q.
c) $\exists!$ rette r' tale che $r' \perp r$ passante per Q con $r' \cap r \neq \emptyset$.

TRA PIANI

Sia π il piano perpendicolare al vettore n (di equazione vettoriale cartesiana $(P - P_0).n = 0$) e sia π' il piano perpendicolare al vettore n' (di equazione vettoriale cartesiana $(P - P_0').n' = 0$).

Definizione: π è parallelo a π' ($\pi \parallel \pi'$) se e solo se $n \parallel n'$ ($\Leftrightarrow n \wedge n' = 0$).
π è ortogonale a π' ($\pi \perp \pi'$) se e solo se $n \perp n'$ ($\Leftrightarrow n.n' = 0$).

Osservazioni: Siano π un piano e Q un punto esterno a π
a) $\exists!$ piano π' tale che $\pi \parallel \pi'$ passante per Q.
b) $\exists\infty$ piani π' tale che $\pi \perp \pi'$ passante per Q.

TRA RETTA E PIANO

Sia r la retta parallela al vettore v (di equazione vettoriale parametrica $P = P_0 + tv,\ t \in \mathbf{R}$) e sia π il piano perpendicolare al vettore n (di equazione vettoriale cartesiana $(P - P_0).n = 0$).

Definizione: r è parallela a π ($r \parallel \pi$) se e solo se $v \perp n$ ($\Leftrightarrow v.n = 0$).
r è ortogonale a π ($r \perp \pi$) se e solo se $v \parallel n$ ($\Leftrightarrow v \wedge n = 0$).

Osservazioni: Siano π un piano e Q un punto esterno a π
a) $\exists!$ retta r tale che $r \perp \pi$ passante per Q.
b) $\exists\infty$ rette r tali che $r \parallel \pi$ passante per Q.

Siano r una retta e Q un punto esterno a r
c) $\exists\infty$ piani π tali che $\pi \parallel r$ passante per Q.
d) $\exists!$ piano π tale che $\pi \perp r$ passante per Q.

INTERSEZIONE IN $\mathbf{R}^3$:

TRA PIANI

Sia $\quad ax+by+cz=d$ l'equazione del piano π e sia $\quad a'x+b'y+c'z=d'$ l'equazione del piano π'.

La intersezione dei piani si trova risolvendo il sistema lineare $\begin{cases} ax+by+cz=d \\ a'x+b'y+c'z=d' \end{cases}$.

La matrice associata è $A=\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$ e la matrice completa è $(A|B)=\left(\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \end{array}\right)$.

Se rango $A=1=$ rango $(A|B)$ il sistema è compatibile, i piani sono coincidenti e paralleli (perche $n=(a,b,c) \parallel n'=(a',b',c')$).

Se rango $A=1<$ rango $(A|B)=2$ il sistema è incompatibile, quindi non c'e intersezione, cioé i piani sono paralleli con $\pi \cap \pi' = \emptyset$.

Se rango $A=2=$ rango $(A|B)$ il sistema è risolubile e la intersezione fra i piani è una retta.

Definizione: L'equazione cartesiana di una retta in $\mathbf{R}^3$ e data dalla intersezione di due piani non paralleli che la contengono.

Se $r:\begin{cases} ax+by+cz=d \\ a'x+b'y+c'z=d' \end{cases}$ allora il vettore parallelo alla retta r è $v_r=(a,b,c)\wedge(a',b',c')$.

TRA PIANO E RETTA

Sia π il piano di equazione $ax+by+cz=d$ e sia r la retta di equazione $\begin{cases} a_1x+b_1y+c_1z=d_1 \\ a_2x+b_2y+c_2z=d_2 \end{cases}$.

La matrice associata è $A=\begin{pmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$ la matrice completa è $(A|B)=\left(\begin{array}{ccc|c} a & b & c & d \\ a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{array}\right)$.

Se rango $A=2=$ rango $(A|B)$ il sistema è compatibile, $r\cap\pi=r\subset\pi$.

Se rango $A=2<$ rango $(A|B)=3$ il sistema è incompatibile, (il piano π è parallelo a r con $\pi\cap r'=\emptyset$).

Se rango $A=3=$ rango $(A|B)$ il sistema è risolubile e la soluzione è unica (quindi la intersezione $\pi\cap r'$ è un punto).

TRA RETTE (e posizione relativa)

Si considerino due rette $r:\begin{cases} a_1x+b_1y+c_1z=d_1 \\ a_2x+b_2y+c_2z=d_2 \end{cases}$ ed $s:\begin{cases} a_3x+b_3y+c_3z=d_3 \\ a_4x+b_4y+c_4z=d_4 \end{cases}$.

Sia $A=\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{pmatrix}$ la matrice associata e $(A|B)=\left(\begin{array}{ccc|c} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{array}\right)$ la matrice completa.

Proposizione:

a) Se rango $A=2=$ rango $(A|B)$ le rette sono coincidenti ($r\cap s=r=s$).

b) Se rango $A=2<$ rango $(A|B)=3$ le rette sono parallele non coincidenti ($r\parallel s\,; r\cap s=\emptyset$).

c) Se rango $A=3=$ rango $(A|B)$ le rette sono incidenti.

d) Se rango $A=3<$ rango $(A|B)=4$ le rette sono sghembe. ($r\not\parallel s$; $r\cap s=\emptyset$)

Definizioni: Due rette si dicono *complanari* se c'e un piano che le contiene. È il caso delle rette parallele e delle rette incidenti. Due rette si dicono *sghembe* se non sono complanari.

Notare che se r ed s sono due rette sghembe rispettivamente parallele ai vettori v e w, si ha che i piani perpendicolari al vettore $v \wedge w$ sono paralleli a tutte e due le rette. Quindi le rette r ed s sono contenute in piani paralleli.

Dimostrazione della proposizione:

Caso a): Se rango $A = 2 =$ rango $(A|B)$ il sistema è compatibile. Per il teorema di Rouchè-Capelli la dimensione dell'insieme delle soluzioni $r \cap s$ è 1. Siccome $(r \cap s) \subseteq r$ e $\dim(r \cap s) = \dim r$ allora $r \cap s = r$, $(r \cap s) \subseteq s$ e $\dim(r \cap s) = \dim s$ allora $r \cap s = s$. Quindi le rette sono coincidenti ($r = s$).

Caso b) Se rango $A = 2 <$ rango $(A|B) = 3$ il sistema è incompatibile, Quindi $r \cap s = \emptyset$.
Si denoti con $n_i = (a_i, b_i, c_i)$ il vettore perpendicolare al piano α_i di equazione $a_i x + b_i y + c_i z = d_i$.

Se rango $A = 2$ allora $\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = 0$ e $\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_4 & b_4 & c_4 \end{pmatrix} = 0$.

Dunque $< n_1 \wedge n_2, n_3 >= 0$ e $< n_1 \wedge n_2, n_4 >= 0$, cioè i vettori n_1, n_2, n_3 sono complanari e i vettori n_1, n_2, n_4 sono complanari. Di conseguenza i vettori n_3 e n_4 appartengono a $\mathrm{Span}\{n_1, n_2\}$. Siccome $\{n_3, n_4\}$ è LI si ha che $\vec{0} \neq n_3 \wedge n_4 \perp \mathrm{Span}\{n_1, n_2\}$. Quindi i vettori paralleli alle rette sono paralleli: $v_s = n_3 \wedge n_4 \parallel n_1 \wedge n_2 = v_r$. Cioé le rette sono parallele (non coincidenti).
Nota: in questo caso $(n_3 \wedge n_4) \wedge (n_1 \wedge n_2) = \vec{0}$.

Caso c) Se rango $A = 3$ si ha che $\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \neq 0$ oppure $\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_4 & b_4 & c_4 \end{pmatrix} \neq 0$.

Dunque $n_3 \notin \mathrm{Span}\{n_1, n_2\}$ oppure $n_4 \notin \mathrm{Span}\{n_1, n_2\}$. Quindi $n_1 \wedge n_2$ non è parallelo a $n_3 \wedge n_4$.

Verifica: si usa la seguente formula

$$u \wedge (v \wedge w) =< u, w > v - < u, v > w$$

Ponendo $u = n_3 \wedge n_4,\ v = n_1,\ w = n_2$, siccome $\{n_1, n_2\}$ è LI si ha che
$(n_3 \wedge n_4) \wedge (n_1 \wedge n_2) =< n_3 \wedge n_4, n_2 > n_1 - < n_3 \wedge n_4, n_1 > n_2 \neq \vec{0}$.

Quindi se rango $A = 3$ le rette non sono parallele. Se anche rango $(A|B) = 3$ per il teorema di R-C sulla risolubilità il sistema è risolubile, inoltre per il teorema R-C sul numero di soluzioni, la soluzione è unica. Quindi la intersezione $r \cap s$ è un punto. Le rette sono incidenti.

Caso d) Se rango $A = 3 <$ rango $(A|B) = 4$ per il teorema di R-C il sistema è incompatibile, quindi $r \cap s = \emptyset$, cioé non sono incidenti. Siccome rango $A = 3$ le rette non sono parallele per ciò non c'e un piano che le contiene. Dunque le rette sono sghembe.

Equazione della retta perpendicolare a due rette sghembe:
Date due rette sghembe r ed s, rispettivamente parallele a v e a w, esiste una e solo una retta l perpendicolare a tutte e due e che le interseca.

La retta l è parallela al vettore $u = v \wedge w$ e giace nella intersezione del piano α parallelo a u che contiene r con il piano β paralelo a u che contiene s.
Cioè $l = \alpha \cap \beta$, dove α è il piano parallelo a u e v ($\perp u \wedge v$) che passa per un punto $R \in r$, β è il piano parallelo a u e w ($\perp u \wedge w$) che passa per un punto $S \in s$.

DISTANZA IN $\mathbf{R}^3$:

TRA DUE PUNTI

La distanza tra due punti di coordinate $P_0 = (x_0, y_0, z_0)$ e $P_1 = (x_1, y_1, z_1)$ è uguale alla norma del vettore $\vec{P_0 P_1}$ di componenti $P_1 - P_0$.

$$d(P_0, P_1) = \|P_1 - P_0\| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}.$$

DA UN PUNTO A UN PIANO

Sia π il piano di equazione $(P - P_1).n = 0$ dove $n = (a, b, c)$; $d = P_1.n$ e sia $P_0 \notin \pi$

$$d(P_0, \pi) = |\ (P_0 - P_1) \cdot \frac{n}{\|n\|}\ | = \frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_- z_1)|}{\sqrt{a^2 + b^2 + c^2}} = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

DA UN PUNTO A UNA RETTA

Sia r una retta di equazione $P = tv + P_1$ e sia $P_0 \notin r$.
Si trova il piano α perpendicolare a r passante per P_0. Si trova il punto Q_0 dove il piano α interseca la retta $r:\ Q_0 = \alpha \cap r$. La distanza da P_0 a r è la distanza tra i punti P_0 e Q_0:

$$d(P_0, r) = d(P_0, Q_0).$$

TRA DUE PIANI

Siano α e α' due piani paralleli.
La distanza tra i piani è uguale alla distanza da uno di loro a un qualunque punto dell'altro: $d(\alpha, \alpha') = d(\alpha, P')$ dove P' e un punto arbitrario di α'.

TRA DUE RETTE PARALLELE

Siano r ed r' due rette parallele (parallele a un vettore v).
Si calcola l'equazione di un piano α perpendicolare alle rette (perpendicolare al vettore v), si trovano le intersezioni delle rette col piano α: $P = r \cap \alpha$ e $P' = r' \cap \alpha$, la distanza tra le rette è uguale alla distanza tra i punti P e P':

$$d(r, r') = d(P, P').$$

Alternativamente, si trovano un punto Q su r e un punto Q' su r'. Si considera il parallelogramma che abbia per lati il vettore $u = Q' - Q$ e il vettore v (applicato nel punto Q). L'area del parallelogramma è uguale alla norma del loro prodotto vettoriale. Se prendiamo come base $\|v\|$, l'altezza corrisponde alla distanza tra le rette r ed r', quindi

$$d(r, r') = \frac{\|u \wedge v\|}{\|v\|}.$$

TRA DUE RETTE SGHEMBE

Siano r ed s due rette sghembe $(r \parallel v,\ s \parallel w)$.
Si calcola l'equazione di un piano β parallelo alle due rette $(\beta \perp v \wedge w)$ e che contenga una di loro, ad esempio $r \subset \beta$. La distanza tra le rette è uguale alla distanza da β a un qualunque punto $Q \in s$:

$$d(r, s) = d(Q, \beta).$$

Alternativamente, si trovano un punto R su r e un punto S su s. Il volume del parallelepipedo di lati $u = R - S$, v e w è uguale all'area della base per l'altezza. La base è il parallelogramma di lati v e w e la sua area è $\|v \wedge w\|$, mentre l'altezza è la norma del vettore proiezione di u sul vettore $v \wedge w$ perpendicolare alla base. La distanza tra le rette sghembe r ed s è pari all'altezza del parallelepipedo, quindi:

$$d(r, s) = \frac{| < v \wedge w, u > |}{\|v \wedge w\|}.$$

Siano V e W spazi vettoriali sul campo $\mathbf{K}$. Una applicazioni $f : V \to W$ si dice lineare se

1) $f(v + v') = f(v) + f(v')\ \forall v, v' \in V$

2) $f(\alpha v) = \alpha f(v)\ \forall v \in V\ \forall \alpha \in \mathbf{K}$.

Esempi di applicazioni lineari.

1) L'applicazione nulla. 2) L'applicazione identità $id : V \to V\ \ id(v) = v\ \forall v \in V$.
3) $f : \mathbf{R}^2 \to \mathbf{R}^2\ f(x, y) = (x + y, 2y)$.
4) $g : \mathbf{R}^3 \to \mathbf{R}^2\ g(x, y, z) = (x - y + z, 2y + 3z)$.
5) Sia $V = \{f : [a, b] \to \mathbf{R}$ derivabile$\}$, l'applicazione $D(f) = f'$.
6) Sia $V = \{f : [a, b] \to \mathbf{R}$ continue$\}$, sia $\phi : V \to \mathbf{R}$ definita da $\phi(f) = \int_a^b f(x)dx$.
7) traccia: $M(n, \mathbf{R}) \to \mathbf{R}$, $A \mapsto$ traccia$A$.
8) $T : \mathbf{K}^n \to \mathbf{K}^m$, $T(X) = AX$ dove $A \in M(m, n; \mathbf{K})$

Esempi di applicazioni non-lineari.

1) det: $M(2, \mathbf{R}) \to \mathbf{R}$, $A \mapsto \det A$.
2) $x \mapsto x + 1$. 3) $x \mapsto x^2$. 4) $(x, y) \mapsto xy$.

Lemma: Se $f : V \to W$ è lineare allora $f(0_V) = 0_W$ e $f(-v) = -f(v)$.
Dimostrazione: Siccome $0_V = 0v$, allora $f(0_V) = f(0v) = 0f(v) = 0_W\ \forall v \in V$.
Siccome $-v = (-1)v\ \forall v \in V$, allora $f(-v) = f((-1)v) = (-1)f(v)\ \forall v \in V$.

NUCLEO e IMMAGINE

Il nucleo (kernel) di un'applicazione lineare è l'insieme $\ker f := \{v \in V$ tali che $f(v) = 0_W\} \subset V$, l'immagine è l'insieme $\mathrm{Im}\ f := \{w \in W$ tali che $\exists v \in V$ che soddisfa $w = f(v)\} \subset W$.

Proposizione: Se $f : V \to W$ è lineare allora i) $\ker f$ è un sottospazio vettoriale di V,
ii) $\mathrm{Im} f$ è un sottospazio vettoriale di W.

APPLICAZIONI INIETTIVE, SURIETTIVE E BIUNIVOCHE

$f : V \to W$ è *iniettiva* se $v_1 \neq v_2 \Rightarrow f(v_1) \neq f(v_2)$ (equivalentemente $f(v_1) = f(v_2) \Rightarrow v_1 = v_2$).
$f : V \to W$ è *suriettiva* se $\mathrm{Im} f = W$ (cioé se $\forall w \in W\ \exists v \in V\ / f(v) = w$).

Proposizione: Se $f : V \to W$ è lineare allora i) f è iniettiva se e solo se $\ker f = \{0_V\}$,
ii) Se $\mathcal{B}\{v_1, \ldots, v_n\}$ è una base di V allora $\mathrm{Im} f = \mathrm{Span}\{f(v_1), \ldots, f(v_n)\}$.

Dimostrazione: i) $\Rightarrow$) Se $v \in \ker f$, per la linearità di f si ha che $f(v) = 0_W = f(0_V)$. La iniettività di f implica che $v = 0_V$. $\Leftarrow$) Se $f(v) = f(v')$, per linearità $f(v - v') = 0_W$, per cui $v - v' \in \ker f$. L'ipotesi implica che $v - v' = 0_V$, se ha dunque che f è iniettiva.
ii) Per ogni $v \in V$ esistono $a_1, \ldots, a_n \in \mathbf{K}$ tali che $v = a_1v_1 + \ldots + a_nv_n$.
Applicando f si ha che $f(v) = f(a_1v_1 + \ldots + a_nv_n) = a_1f(v_1) + \ldots + a_nf(v_n)$.
Quindi $\{f(v_1), \ldots, f(v_n)\}$ è un insieme di generatori di $\mathrm{Im} f$.

Definizione: Un isomorfismo è una applicazione lineare biunivoca.

Corollario: $f : V \to W$ è un isomorfismo se e solo se $\ker f = \{0_V\}$ e $\mathrm{Im} f = W$.

Lemma: La scelta di una base $\mathcal{B}$ di V individua un isomorfismo fra V e $\mathbf{K}^n$.
Dimostrazione: Al vettore $v \in V$ si fa corrispondere la n-upla $(a_1, \ldots, a_n) \in \mathbf{K}^n$ delle componenti di v nella base $\mathcal{B}$.

TEOREMA DELLA DIMENSIONE

Proposizione: Se $f : V \to W$ è lineare allora

1) Se $v_1, \dots, v_k \in V$ sono linearmente dipendenti (L D) allora $f(v_1), \dots, f(v_k) \in W$ sono L D.
2) Se $f(v_1), \dots, f(v_k) \in W$ sono linearmente indipendenti (L I) allora $v_1, \dots, v_k \in V$ sono L I.
3) Se $\ker f = \{0_V\}$ e $v_1, \dots, v_k \in V$ sono L I allora $f(v_1), \dots, f(v_k) \in W$ sono L I.

Dimostrazione:

1) Se $\{v_1, \dots, v_k\}$ è LD esistono coefficienti $a_1, \dots, a_k$ non tutti nulli tali che $a_1v_1 + \dots + a_kv_k = 0_V$. La linearità di f implica che $f(0_V) = 0_W$ e $f(a_1v_1 + \dots + a_kv_k) = a_1f(v_1) + \dots + a_kf(v_k)$ si ha dunque $a_1f(v_1) + \dots + a_kf(v_k) = 0_W$.
2) Se $\vec{0}_V = a_1v_1 + \dots + a_kv_k$, applicando f si ha $0_W = f(\vec{0}_V) = f(a_1v_1 + \dots + a_kv_k)$ $= f(a_1v_1) + \dots + f(a_kv_k) = a_1f(v_1) + \dots + a_kf(v_k)$.
L'indipendenza lineare di $\{f(v_1), \dots, f(v_k)\}$, implica che $a_1 = \dots = a_k = 0$.
3) Se $a_1f(v_1) + \dots + a_kf(v_k) = 0_W$, per la linearità di f si ha $f(a_1v_1 + \dots + a_kv_k) = 0_W$. Quindi $a_1v_1 + \dots + a_kv_k \in \ker f$. Per ipotesi $\ker f = \{0_V\}$ allora $a_1v_1 + \dots + a_kv_k = 0_V$. $\{v_1, \dots, v_k\}$ LI implica $a_1 = \dots = a_k = 0$. Quindi $\{f(v_1), \dots, f(v_k)\}$ è L I.

Teorema (della dimensione): Se $\dim V < \infty$ e se $f : V \to W$ è lineare allora i) $\dim \ker f \le \dim V$, ii) $\dim \mathrm{Im}\ f \le \dim V$ e iii) $\dim \ker f + \dim \mathrm{Im}\ f = \dim V$.

Dimostrazione: i) e ii) seguono dalla proposizione precedente.
Per dimostrare iii) consideriamo tre casi:
caso 1) Se dim $\ker f = \dim V$ allora $\ker f = V$ quindi $f(v) = 0_W$ $\forall v \in V$, si ha che dim Im $f = 0$.
caso 2) Se dim Ker$f = 0$ e $\mathcal{B}\{v_1, \dots, v_n\}$ è una base di V allora $\{f(v_1), \dots, f(v_n)\}$ sono linearmente indipendenti quindi dim Im f = dim V.
caso 3) $0 < \dim \ker f = q < \dim V$. Siano $\{u_1, \dots, u_q\}$ una base di $\ker f$ e $\{w_1, \dots, w_p\}$ una base di Imf, allora $\exists v_1, \dots, v_p \in V$ tali che $w_1 = f(v_1), \dots, w_p = f(v_p)$. Basta dimostrare che $\{u_1, \dots, u_q, v_1, \dots, v_p\}$ è una base di V.

Dimostriamo prima che $\{u_1, \dots, u_q, v_1, \dots, v_p\}$ genera V:
Per ogni $v \in V$, esistono $a_1, \dots, a_p \in \mathbf{K}$ tali che
Im $f \ni f(v) = a_1w_1 + \dots + a_pw_p = a_1f(v_1) + \dots + a_pf(v_p)$
Per la linearità di f si ha $f(v) = f(a_1v_1 + \dots + a_pv_p)$. Quindi $f(v - a_1v_1 - \dots - a_pv_p) = 0_W$.
Questo implica che $v - a_1v_1 - \dots - a_pv_p \in \ker f$.
Per tanto esistono $b_1, \dots, b_q \in \mathbf{K}$ tali che $v - a_1v_1 - \dots - a_pv_p = b_1u_1 + \dots + b_qu_q$.
Dunque $v = a_1v_1 + \dots + a_pv_p + b_1u_1 + \dots + b_qu_q$. Quindi Span$\{u_1, \dots, u_q, v_1, \dots, v_p\} = V$.

Per verificare l'indipendenza lineare di $\{u_1, \dots, u_q, v_1, \dots, v_p\}$:
sia $0_V = \alpha_1u_1 + \dots + \alpha_qu_q + \beta_1v_1 + \dots + \beta_pv_p$ si vuole dimostrare che
$\alpha_1 = \dots = \alpha_q = \beta_1 = \dots = \beta_p = 0$.
Applicando f e usando la linearità si ha che
$0_W = f(0_V) = f(\alpha_1u_1 + \dots + \alpha_qu_q + \beta_1v_1 + \dots + \beta_pv_p)$
$\quad = f(\alpha_1u_1 + \dots + \alpha_qu_q) + f(\beta_1v_1 + \dots + \beta_pv_p)$.
$\alpha_1u_1 + \dots + \alpha_qu_q \in \ker f \Rightarrow f(\alpha_1u_1 + \dots + \alpha_qu_q) = 0_W$
Dunque $0_W = f(\beta_1v_1 + \dots + \beta_pv_p) = \beta_1f(v_1) + \dots + \beta_pf(v_p) = \beta_1w_1 + \dots + \beta_pw_p$.
Siccome $\{w_1, \dots, w_p\}$ è LI allora $\beta_1 = \dots = \beta_p = 0$. Per tanto $\alpha_1u_1 + \dots + \alpha_qu_q = 0_V$.
La indipendenza lineare de $\{u_1, \dots, u_q\}$ implica che $\alpha_1 = \dots = \alpha_q = 0$.

Definizione: il rango di un'applicazione lineare è la dimensione del sottospazio Imf.

Corollario: Se $f : V \to W$ è lineare allora

1) f è iniettiva se e solo se rango $f = \dim V$
2) f è suriettiva se e solo se rango $f = \dim W$.

MATRICE RAPPRESENTATIVA di una applicazione lineare fra spazi vettoriali di dimensione finita.

Siano V e W due spazi vettoriali tali che $\dim V = n$ e $\dim W = m$. Il numero di colonne di ogni matrice che rappresenta una applicazione lineare $f \colon V \to W$ è pari alla dimensione di V mentre il numero di righe è pari alla dimensione di W, cioè: le matrici che rappresentano f sono di ordine $m \times n$.
Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e sia $\mathcal{B}' = \{w_1, \ldots, w_m\}$ una base di W. La matrice $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ rappresentativa di $f \colon V \to W$ nelle basi $\mathcal{B}$, $\mathcal{B}'$ è definita nel seguente modo:

la j-esima colonna è costituita dalle componenti rispetto alla base $\mathcal{B}'$ dell'imagine $f(v_j)$ del j-esimo vettore v_j della base $\mathcal{B}$.

Dunque, se

$$\begin{aligned} f(v_1) &= a_{11}w_1 + \ldots + a_{m1}w_m \\ &\vdots \\ f(v_n) &= a_{1n}w_1 + \ldots + a_{mn}w_m \end{aligned} \qquad \text{allora} \qquad M^{\mathcal{B}}_{\mathcal{B}'}(f) = \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

La matrice $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ dipende dalle basi $\mathcal{B}$ di V e $\mathcal{B}'$ di W, vale a dire che se si scelgono altre basi per V o per W si ottiene una matrice diversa da $M^{\mathcal{B}}_{\mathcal{B}'}(f)$.

Siano $x_1, \ldots, x_n$ le componenti di un vettore $v \in V$ rispetto alla base $\mathcal{B}$, cioè $v = x_1v_1 + \cdots + x_nv_n$. La linearità di f implica che

$$\begin{aligned} f(v) &= x_1 f(v_1) + \cdots + x_n f(v_n) \\ &= x_1(a_{11}w_1 + \cdots + a_{m1}w_m) + \cdots + x_n(a_{1n}w_1 + \cdots + a_{mn}w_m)\,. \\ &= (x_1a_{11} + \cdots + x_na_{1n})w_1 + \cdots + (x_1a_{m1} + \cdots + x_na_{mn})w_m \end{aligned}$$

Quindi, ponendo

$$\begin{aligned} y_1 &= x_1a_{11} + \cdots + x_na_{1n} \\ &\vdots \\ y_m &= x_1a_{m1} + \cdots + x_na_{mn} \end{aligned} \qquad \text{otteniamo} \qquad AX = Y,$$

dove $A = M^{\mathcal{B}}_{\mathcal{B}'}(f)$. $X = v_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $Y = [f(v)]_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ sono rispettivamente il vettore delle componenti di v nella base $\mathcal{B}$ e il vettore delle componenti di $f(v)$ nella base $\mathcal{B}'$.

BASI di $\ker f$ e di $\mathrm{Im} f$
Sia A la matrice che rappresenta la applicazione lineare $f : V \to W$ nelle basi $\mathcal{B}$ di V e $\mathcal{B}'$ di W.

Una base di $\ker f$ si ottiene calcolando una base dello spazio delle soluzioni del sistema lineare omogeneo $AX = 0$ (riducendo A per righe). Le componenti delle soluzioni X sono le componenti degli elementi della base di $\ker f$ rispetto alla base $\mathcal{B}$ di V.

Siccome $\mathrm{Im} f = \mathrm{Span}\{f(v_1), \ldots, f(v_n)\}$. Per trovare una base di $\mathrm{Im} f$ si riduce la matrice A per colonne. Le colonne non nulle della matrice ridotta (per colonne) sono le componenti, rispetto alla base $\mathcal{B}'$ di W, di una base di $\mathrm{Im} f$. Alternativamente si applica alle colonne di A il metodo degli scarti successivi.

Un terzo modo di definire una applicazione lineare è data dalla seguente

Proposizione:
Siano V e W due spazi vettoriale, sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e $w_1, \ldots, w_n$ vettori in W.
Allora esiste una unica applicazione lineare $f : V \to W$ tale che $f(v_i) = w_i,\ i = 1, \ldots, n$.
L'applicazione f soddisfa $f(a_1v_1 + \ldots + a_nv_n) = a_1w_1 + \ldots + a_nw_n$.

MATRICE DI CAMBIO DI BASE
Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, \ldots, w_n\}$ due basi di uno spazi vettoriali V sul campo $\mathbf{K}$
Si scrivono i vettori della base $\mathcal{B}'$ nella base $\mathcal{B}$. Le componenti degli elementi della base $\mathcal{B}'$ rispetto alla base $\mathcal{B}$ formano le colonne di una matrice P che è detta matrice di cambio di base da $\mathcal{B}$ a $\mathcal{B}'$, come segue:

$$\begin{aligned} w_1 &= p_{11}v_1 + \cdots + p_{n1}v_n \\ &\vdots \\ w_n &= p_{1n}v_1 + \cdots + p_{nn}v_n \end{aligned} .$$

La matrice $P = \begin{pmatrix} p_{11} & \ldots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \ldots & p_{nn} \end{pmatrix}$ può essere vista come:

1) $P = M_{\mathcal{B}}^{\mathcal{B}}(\phi)$ la matrice dell'applicazione lineare $\phi : (V, \mathcal{B}) \to (V, \mathcal{B})$ tale che $\phi(v_i) = w_i$ (nella stessa base). Oppure

2) $P = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id})$ la matrice dell'applicazione lineare $\mathrm{id} : (V, \mathcal{B}') \to (V, \mathcal{B})$ tale che $\mathrm{id}(w_i) = w_i$ nelle basi $\mathcal{B}'$ e $\mathcal{B}$.

OPERAZIONI tra applicazioni lineari

La somma tra due applicazioni f e g da V in W è definita da

$$(f + g)(v) := f(v) + g(v),$$

il prodotto tra uno scalare α ed f è definito da

$$(\alpha f)(v) := \alpha f(v).$$

Se $f : U \to V$ e $g : V \to W$ la composizione $g \circ f : U \to W$ é definita da

$$(g \circ f)(u) := g(f(u)).$$

Proposizione: Se f e g sono applicazioni lineari allora anche $f + g,\ \alpha f,\ f \circ g$. lo sono.

L'applicazione composizione è associativa, distributiva rispetto alla somma, ed omogenea rispetto al prodotto per uno scalare.

In generale la composizione è non commutativa: $f \circ g \neq g \circ f$
ESEMPIO: $V = W = U = \mathbf{R}^3$ $f(x, y, z) = (x, y, 0)$, $g(x, y, z) = (0, z, 0)$
$(g \circ f)(x, y, z) = g(f(x, y, z)) = g(x, y, 0) = (0, 0, 0)$.
$(f \circ g)(x, y, z) = f(g(x, y, z)) = f(0, z, 0) = (0, z, 0)$.

Definizione: Una applicazione lineare $f : V \to W$ si dice *invertibile* se esiste una applicazione $g : W \to V$ lineare tale che la composizione $g \circ f : V \to V$ è l'applicazione identità di V e la composizione $f \circ g : W \to W$ è l'applicazione identità di W.

Se una tale applicazione g esiste allora è unica e la si denota con f^{-1}.

Proposizione: Se $f : V \to W$ è lineare, allora f è invertibile se e solo se f è biiettiva.

Dimostrazione: ($\Rightarrow$) Se g è l'inversa di f e, $f(v_1) = f(v_2)$ allora $v_1 = g(f(v_1)) = g(f(v_2)) = v_2$ quindi f è iniettiva.
Sia $w \in W$, siccome $f \circ g = id_W$ si ha $w = (f \circ g)(w)$. Ponendo $v = g(w) \in V$ si ottiene $w = f(v)$, quindi f è suriettiva.
($\Leftarrow$) Se f è suriettiva si ha che $\forall\ w \in W \exists v \in V\colon w = f(v)$. Allora si definisce l'applicazione $g : W \to V$ che a w assegna v.
Verifica che g è ben definita: Se $v_1, v_2 \in V$ tali che $g(w) = v_1$ e $g(w) = v_2$ applicando f si ottiene che $f(v_1) = f(g(w)) = f(v_2)$. L'ipotesi che f sia iniettiva implica che $v_1 = v_2$.

Corollario: Se $f : V \to W$ è lineare e $\dim V = \dim W$ allora sono equivalenti:
i) f è un isomorfismo,
ii) f è iniettiva,
iii) f è suriettiva.

Proposizione: Se $f : V \to W$ è lineare ed invertibile, allora $f^{-1} : W \to V$ è lineare.

Dimostrazione:
a) Siano $w_1, w_2 \in W$ per la suriettività di f $\exists\ v_1, v_2 \in V$ tali che $w_1 = f(v_1), w_2 = f(v_2)$
$f^{-1}(w_1 + w_2) = f^{-1}[f(v_1) + f(v_2)] = f^{-1}(f(v_1 + v_2)) = [f^{-1} \circ f](v_1 + v_2) = v_1 + v_2$
$= f^{-1}(w_1) + f^{-1}(w_2)$.
b) $f^{-1}(aw) = f^{-1}(af(v)) = f^{-1}(f(av)) = [f^{-1} \circ f](av) = av = af^{-1}(w)$.

MATRICE rappresentativa delle operazione tra applicazioni lineari:

Prpoposizione: Se $M(f) = A$ e $M(g) = B$ si ha che

$$M(f + g) = A + B,\ M(\alpha f) = \alpha A,\ \ M(g \circ f) = BA \ \text{ e } \ M(f^{-1}) = A^{-1}.$$

Dimostrazione:
si considerino le applicazioni lineari

$$f,\ g,\ f + g,\ \alpha f\colon (V, \mathcal{B}_V) \to (W, \mathcal{B}_W),$$

dove $\mathcal{B}_V = \{v_1, \dots, v_n\}$ è una base di V e $\mathcal{B}_W = \{w_1, \dots, w_m\}$ è una base di W.
Siano $A = M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ e $B = M^{\mathcal{B}_V}_{\mathcal{B}_W}(g)$ le matrici rappresentative di f e g.

$f(v_j) = \Sigma a_{ij} w_i$ e $g(v_j) = \Sigma b_{ij} w_i \Longrightarrow (f + g)(v_j) = \Sigma a_{ij} w_i + \Sigma b_{ij} w_i = \Sigma (a_{ij} + b_{ij}) w_i$.
Quindi $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f + g) = A + B$.

$(\alpha f)(v_j) = \alpha(\Sigma a_{ij} w_i) = \Sigma \alpha a_{ij} w_i$. Quindi $M^{\mathcal{B}_V}_{\mathcal{B}_W}(\alpha f) = \alpha A$.

Si consideri

$$g \circ f : (U, \mathcal{B}_U) \to (V, \mathcal{B}_V) \to (W, \mathcal{B}_W)$$

dove $\mathcal{B}_U = \{u_1, \dots, u_h\}$, $\mathcal{B}_V = \{v_1, \dots, v_n\}$, e $\mathcal{B}_W = \{w_1, \dots, w_m\}$ sono le rispettive basi.

Siano $A = M^{\mathcal{B}_U}_{\mathcal{B}_V}(f)$ e $B = M^{\mathcal{B}_V}_{\mathcal{B}_W}(g)$
$f(u_j) = \Sigma_k a_{kj} v_k$
$g(f(u_j)) = g(\Sigma_k a_{kj} v_k) = \Sigma_k a_{kj} g(v_k) = \Sigma_k a_{kj}(\Sigma_i b_{ik} w_i) = \Sigma_i(\Sigma_k b_{ik} a_{kj}) w_i = \Sigma_i (BA)_{ij} w_i$
Quindi $M^{\mathcal{B}_U}_{\mathcal{B}_W}(g \circ f) = BA$.

Se $f^{-1}\colon W \to V$ è l'inversa di $f\colon V \to W$. Siano $A = M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ e $B = M^{\mathcal{B}_W}_{\mathcal{B}_V}(f^{-1})$
Siccome $id = f^{-1} \circ f : (V, \mathcal{B}_V) \to (W, \mathcal{B}_W) \to (V, \mathcal{B}_V)$ si ha che
$I = M^{\mathcal{B}_V}_{\mathcal{B}_V}(id) = M^{\mathcal{B}_V}_{\mathcal{B}_V}(f^{-1} \circ f) = BA \Rightarrow B = A^{-1}$.

Sia V uno spazio vettoriale su un campo $\mathbf{K}$.
Definizioni: Un *endomorfismo* di V è una applicazioni lineare da V in se stesso: $f : V \to V$ lineare. Un $\lambda \in \mathbf{K}$ è un *autovalore* dell'endomorfismo f se esiste un vettore non nullo $v \in V$ tale che $f(v) = \lambda v$. Un tale vettore è detto *autovettore* di f relativo all'autovalore λ. Il sottospazio vettoriale $V_\lambda = \{v \in V \text{ tale che } f(v) = \lambda v\}$ è detto *l'autospazio* associato all'autovalore λ.

Esempi:
1) L'unico autovalore dell'applicazione identità è 1 e tutti i vettori non nulli sono autovettori.
2) Gli autovalori di $f(x, y) = (2x, 3y)$ sono 2 e 3. $V_2 = \text{Span}\{(1, 0)\}$ e $V_3 = \text{Span}\{(0, 1)\}$.
3) Gli autovalori di $g(x, y) = (-y, -x)$ sono 1 e -1. $V_1 = \text{Span}\{(1, -1)\}$ e $V_3 = \text{Span}\{(1, 1)\}$.
4) Sia $h(x, y) = (-y, x)$. Se $V = \mathbf{R}^2$, h non ha autovalori. Invece se $V = \mathbf{C}^2$ gli autovalori di h sono i e $-i$.
5) Se $Df = f'$. $f'(x) = \lambda f(x)\ \forall x \in \mathbf{R} \Rightarrow f(x) = ce^{\lambda x}$. Quindi ogni $\lambda \in \mathbf{R}$ è autovalore di D e $V_\lambda = \text{Span}\{e^{\lambda x}\}$, $\dim V_\lambda = 1$.

Osservazioni:
1) 0 è autovalore di f se e solo se $V_0 = \ker f \neq \{\vec{0}\}$, cioè se e solo se f è non-iniettiva.
2) Se $\lambda \neq 0$, sia v un autovettore relativo a λ allora $v \in \text{Im} f$.
3) Se λ è un autovalore di f allora λ^2 è autovalore di f^2.

POLINOMIO CARATTERISTICO

Lemma: Se λ_0 è un autovalore di $f : V \to V$ allora $V_{\lambda_0} = \ker(f - \lambda_0 \text{id}) \neq \{0\}$, quindi l'applicazione lineare $(f - \lambda_0 \text{id}) : V \to V$ è non-iniettiva.

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V^n e siano $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$e $I = M^{\mathcal{B}}_{\mathcal{B}}(\text{id})$ le matrici rappresentative di f e della'identità nella base $\mathcal{B}$ rispettivamente. Sia v un autovettore relativo a λ_0 e $X = v_{\mathcal{B}}$ dunque si ha che $(A - \lambda_0 I)X = \vec{0}$, quindi $\det(A - \lambda_0 I) = 0$.

Definizione: si chiama *polinomio caratteristico* di f al polinomio nella indeterminata λ di grado n

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

Ne segue dal lemma che gli autovalori di $f : V \to V$ sono le radici del polinomio caratteristico che appartengono al campo $\mathbf{K}$ (cioé le soluzioni dell'equazione $\det(A - \lambda I) = 0$ che appartengono a $\mathbf{K}$).

Proposizione: Se B è un'altra matrice rappresentativa di f allora $\det(A - \lambda I) = \det(B - \lambda I)$.

Dimostrazione: Se $\mathcal{B}'$ è un'altra base di V e B è la matrice rappresentativa di f nella base $\mathcal{B}'$, e P è la matrice di cambio di base da $\mathcal{B}$ a $\mathcal{B}'$. Si ha che $PBP^{-1} = A$ quindi $\det(A - \lambda I) = \det(PBP^{-1} - \lambda PIP^{-1}) = \det(P(B - \lambda I)P^{-1}) = \det P \det(B - \lambda I) \det P^{-1} = \det(B - \lambda I)$.

Quindi il polinomio caratteristico non dipende dalla matrice rappresentativa, cioè non dipende dalla base scelta per rappresentare f e lo denotiamo con p_f.

MOLTEPLICITA' ALGEBRICA E GEOMETRICA

La molteplicità algebrica m_{λ_0} dell'autovalore λ_0 è la molteplicità di λ_0 come radice del polinomio caratteristico.
La molteplicità geometrica μ_{λ_0} dell'autovalore λ_0 è la dimensione dell'autospazio associato V_{λ_0}.

Proposizione: Se $\dim V = n$ allora $1 \leq \mu_{\lambda_0} \leq m_{\lambda_0} \leq n$.

Dimostrazione: Completiamo una base $\{u_1, \dots, u_{\mu_0}\}$ di V_{λ_0} a una base $\{u_1, \dots, u_{\mu_0}, v_{\mu_0+1}, \dots v_n\}$ di V. La matrice di f in questa base è $A = M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_0 I_{\mu_0} & B \\ 0 & C \end{pmatrix}$.

Il polinomio caratteristico $p_f(\lambda) = \det(A - \lambda_0 I) = (\lambda - \lambda_0)^{\mu_0} q(\lambda)$. Quindi $(\lambda - \lambda_0)^{\mu_0}$ divide $p_f(\lambda)$. Per definizione di molteplicità algebrica m è il più grande intero tale che $(\lambda - \lambda_0)^{\mu}$ divide $p_f(\lambda)$, quindi $\mu_0 \leq m_0$.

Definizione di somma diretta di più di due sottospazi vettoriale di V:
la somma $V_1 + \cdots + V_k = \{v_1 + \cdots + v_k \in V$ tali che $v_i \in V_i, i = 1, \ldots, k\}$ è diretta se e solo se ogni elemento non nullo di $V_1 + \cdots + V_k$ si scrive in modo unico come somma di elementi di $V_1, \ldots, V_k$.
Caratterizzazione: $V_1 + \cdots + V_k = V_1 \oplus \cdots \oplus V_k \Leftrightarrow V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots V_k) = \{\vec{0}\}$, inoltre $\dim(V_1 \oplus \cdots \oplus V_k) = \dim V_1 + \ldots + \dim V_k$.

Teorema: Se $\lambda_1, \ldots, \lambda_k$ sono autovalori di $f : V \to V$ distinti a due a due e siano $v_1, \ldots, v_k$ autovettori associati a $\lambda_1, \ldots, \lambda_k$ rispettivamente allora $\{v_1, \ldots, v_k\}$ è linearmente indipendente.

Dimostrazione: Supporre che $\{v_1, \ldots, v_k\}$ sia LD. Allora esiste i tale che $\{v_1, \ldots, v_i\}$ è LI e $v_{i+1} = a_1 v_1 + \cdots + a_i v_i$. Applicando f si ottiene che $\lambda_1 a_1 v_1 + \cdots + \lambda_i a_i v_i = \lambda_{i+1} v_{i+1} = \lambda_{i+1}(a_1 v_1 + \cdots + a_i v_i)$, si ha dunque che $(\lambda_1 - \lambda_{i+1}) a_1 v_1 + \cdots + (\lambda_i - \lambda_{i+1}) a_i v_i = \vec{0}$ il che implica che $a_1 = \ldots = a_i = 0$ e quindi $v_{i+1} = \vec{0}$, il che non è possibile perche v_{i+1} è un autovettore.

Corollario: Se $\lambda_1, \ldots, \lambda_k$ sono autovalori di $f : V \to V$ distinti a due a due allora
1) la somma $V_{\lambda_1} + \cdots + V_{\lambda_k}$ è diretta,
2) $\dim(V_{\lambda_1} + \cdots + V_{\lambda_k}) = \dim V_{\lambda_1} + \ldots + \dim V_{\lambda_k}$,
3) $\dim V_{\lambda_1} + \ldots + \dim V_{\lambda_k} \leq \dim V$.

Corollario: Sia $f : V \to V$ un endomorfismo, se $\dim V = n < \infty$ allora f ha al più n autovalori.

ENDOMORFISMI SEMPLICI

Definizione: Un endomorfismo $f : V \to V$ è *semplice o diagonalizzabile* se lo spazio vettoriale V ammette una base costituita di autovettori di f.

Osservazione: Se f è semplice allora $V = \ker f \oplus \ \mathrm{Im} f$

Teorema (criteri di diagonalizzazione):

Siano V uno spazio vettoriale di dimensione finita n su un campo $\mathbf{K}$, f un endomorfismo e $\lambda_1, \ldots, \lambda_k$ autovalori di f distinti a due a due, allora la seguente affermazioni sono equivalenti:
1) f è semplice,
2) $V_{\lambda_1} \oplus ... \oplus V_{\lambda_k} = V$,
3) $\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_k} = \dim V$,
4) tutte le radici del polinomio caratteristico di f appartengono a $\mathbf{K}$ e $\mu_{\lambda_i} = m_{\lambda_i}$ $\forall i = 1, \ldots, k$.

Dimostrazione: 1⇒2) Se esiste una base $\{v_1, \ldots, v_n\}$ di V costituita da autovettori di f allora $\forall v \in V$ $v = \Sigma_{i=1}^{n} v_i$. Riordinando i vettori della base si otiene $v = \Sigma_{i=1}^{k} w_i$ dove $w_i \in V_{\lambda_i}$, quindi $V = V_{\lambda_1} + \cdots + V_{\lambda_k}$. Segue dal corollario che la somma è diretta.

2⇒3) $V_{\lambda_1} + \cdots + V_{\lambda_k} = V$ quindi $\dim V_{\lambda_1} + \ldots + \dim V_{\lambda_k} = \dim V$.

3⇒4) $n = \dim V = \mu_{\lambda_1} + \ldots + \mu_{\lambda_k} \leq m_{\lambda_1} + \ldots + m_{\lambda_k} = n \Rightarrow \mu_{\lambda_i} = m_{\lambda_i}$.

4⇒1) siccome le radici del polinomio caratteristico appartengono a $\mathbf{K}$ allora sono autovalori. Se $\mu_{\lambda_i} = m_{\lambda_i}$ allora una base di V_{λ_i} ha m_{λ_i} vettori, cioè $\mathcal{B}_{V_{\lambda_i}} = \{u_1, \ldots, u_{m_{\lambda_i}}\}$. Quindi l'unione delle basi di tutti gli autospazi è una base di V di autovettori di f: $\mathcal{A} = \mathcal{B}_{V_{\lambda_1}} \cup \ldots \cup \mathcal{B}_{V_{\lambda_k}}$.

Osservazione: la matrice rappresentativa di un endomorfismo semplice in una base $\mathcal{A}$ di autovettori di f è diagonale. Gli autovalori giacciono sulla diagonale.

Definizione: A e B si dicono coniugate se esiste una matrice P invertibile tale che $P^{-1}AP = B$.

Se $A = M_{\mathcal{B}}(f)$ è la matrice che rappresenta l'endomorphismo f in una base $\mathcal{B}$, diagonalizzare A significa trovare (se esiste) una matrice P invertibile tale che A sia coniugata a una matrice diagonale D, cioè $P^{-1}AP = D$, dove P è la matrice di passaggio dalla base $\mathcal{B}$ alla base $\mathcal{A}$.

ESEMPI Si considerino in $\mathbf{R}^2$ gli endomorfismi rappresentati dalle matrici:

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

Il polinomio caratteristico di A è $p_A(\lambda) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$, gli autovalori di A sono 1 e -1, quindi A è semplice e la base di $\mathbf{R}^2$ di autovettori di A è $\{(1, 1); (-1, 1)\}$.

Il polinomio caratteristico di B, $p_B(\lambda) = \lambda^2 + 1$ non ha radici reali, quindi sul campo reale $\mathbf{R}$ non è semplice. Invece sul campo complesso $\mathbf{C}$ $\lambda^2 + 1 = (\lambda - i)(\lambda + i)$, quindi gli autovalori sono i e $-i$ e una base di $\mathbf{C}^2$ di autovettori di B è $\{(1, -i); (1, i)\}$. Quindi B è semplici su $\mathbf{C}$.

Il polinomio caratteristico di C è $p_C(\lambda) = (\lambda - 1)^2$, quindi 1 è autovalore di C con molteplicià algebrica 2. Una base di V_1 è $\{(1, 0)\}$ per ciò la molteplicità geometrica è $\dim V_1 = 1$. Quindi C non è semplice (o diagonalizzabile).

OSSERVAZIONI

1) Autovalori e autovettori della matrice l'inversa:
Se A è una matrice invertibile e λ é un autovalore di A allora λ^{-1} è un autovalore di A^{-1}.

Verifica: se λ é un autovalore di A allora $\lambda \neq 0$. Sia v un suo autovettore, quindi $Av = \lambda v$ molteplicando per A^{-1} si ottiene che $A^{-1}(Av) = (A^{-1}A)v = \lambda(A^{-1}v)$, quindi $A^{-1}v = \lambda^{-1}v$. Notare che A e A^{-1} hanno gli stessi autovettori.

2) Autovalori e autovettori della matrice trasposta: A e A^t hanno gli stessi autovalori.
Verifica: Per ogni matrice M si ha che $\det M = \det M^t$.
Ponendo $M = (A - \lambda I)$ si ha $\det(A - \lambda I) = \det(A - \lambda I)^t = \det(A^t - \lambda I)$
quindi A e A^t hanno lo stesso polinomio caratteristico, quindi gli stessi autovalori.
In generale A e A^t non hanno gli stessi autovettori.
Esempio:

Gli autovettori di $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ sono i vettori $\begin{pmatrix} x \\ 0 \end{pmatrix}$ con $x \neq 0$,

mentre gli autovettoti di $A^t = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ sono i vettori $\begin{pmatrix} 0 \\ y \end{pmatrix}$ con $y \neq 0$.

3) Autovalori e autovettori del prodotto e somma di matrici:

Se λ è un autovalore di A e μ è un autovalore di B, in generale $\lambda\mu$ nonè autovalore di AB, e $\lambda + \mu$ non è autovalore di $A + B$.
Esempio:

0 è l'unico autovalore di $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ e di $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

mentre gli autovalori di $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ sono 1 e 0,

e gli autovalori di $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix})$ sono 1 e -1.

I COEFFICIENTI del polinomio caratteristico $p_f(\lambda)$

Teorema: Il polinomio caratteristico e quindi i suoi coefficienti sono invarianti del endomorfismo.

Se $p_f(\lambda) = \alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_1\lambda^1 + \alpha_0$, i suoi coefficiente sono

$\alpha_{n-k} = (-1)^{n-k}\Sigma \det(\text{minori principali di ordine k})$.

In particolare

$\alpha_n = (-1)^n$,

$\alpha_{n-1} = (-1)^{n-1}\text{traccia } A$,

$\alpha_0 = \det A$.

Dimostrazione: per induzione sull'ordine della matrice A.

Se $n = 2$, il polinomio caratteristico di $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} \\ &= (a-\lambda)(d-\lambda) - cb \\ &= ad - a\lambda - d\lambda + \lambda^2 - cb \\ &= \lambda^2 - (a+d)\lambda + ad - cb \\ &= \lambda^2 - \text{ traccia } A\lambda + \det A. \end{aligned}$$

Se $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \in M(n, \mathbf{K})$

$$p_A(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} a_{11}-\lambda & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn}-\lambda \end{pmatrix}$$

$$= (a_{11}-\lambda)\det \begin{pmatrix} a_{22}-\lambda & \ldots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \ldots & a_{nn}-\lambda \end{pmatrix} + \text{ termini di grado al più } n-2.$$

Per ipotesi induttiva il polinomio caratteristico

$p_A(\lambda) = (a_{11}-\lambda)[(-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}(a_{22} + \cdots + a_{nn})\lambda^{n-2} + \cdots] + \cdots$

$p_A(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(\text{ traccia } A)\lambda^{n-1} + \cdots$.

Il termine noto $\alpha_0 = p_A(0) = \det(A - 0I) = \det A$.

La traccia e il determinante

Siccome il polinomio caratteristico e quindi i suoi coefficienti sono invarianti del endomorfismo, se $\lambda_1, \ldots, \lambda_n$ sono le radici del polinomio caratteristico della matrice A allora

$$\text{traccia } A = a_{11} + \cdots + a_{nn} = \lambda_1 + \cdots + \lambda_n.$$

e il prodotto delle radici del polinomio caratteristico è il determinante di A

$$\det A = \lambda_1 \ldots \lambda_n.$$

APPLICAZIONI

I. La potenza k-essima della matrice A.

Siano A e $P \in M(n, \mathbf{K})$ tale che P sia invertibile, si osservi che

$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P$,

e per ogni intero positivo k si ha che $(P^{-1}AP)^k = P^{-1}A^kP$.

Se A è diagonalizzabile e $P^{-1}AP = D$ è diagonale allora $P^{-1}A^kP = (P^{-1}AP)^k = D^k$, quindi $A^k = PD^kP^{-1}$.

II. La traccia della matrice A^k.

Si osservi che tr (AB) = tr (BA), quindi se A è diagonalizzabile e $\lambda_1, \dots, \lambda_n$ sono i suoi autovalori allora $\mathrm{tr\ A = a_{11} + \cdots + a_{nn} = \lambda_1 + \cdots + \lambda_n}$.
Per il risultato precedente si ha che $\mathrm{tr\ A^k = \lambda_1^k + \cdots + \lambda_n^k}$.

III. La serie esponenziale della matrice A.

La matrice esponenziale è definita da $e^{At} = I + At + \frac{1}{2}(At)^2 + \ldots + \frac{1}{k!}(At)^k + \ldots$.
La serie e^{At} converge sempre e la sua derivata è Ae^{At}.
Se A è diagonalizzabile e se $A = PDP^{-1}$ allora $e^{At} = Pe^{Dt}P^{-1}$.
Se $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A allora $e^{\lambda_1 t}, \dots, e^{\lambda_n t}$ sono gli autovalori di e^{At}.

Sia $\mathbf{K} = \mathbf{R}$ o $\mathbf{C}$ e V uno spazio vettoriale su $\mathbf{K}$. Un prodotto scalare su V è una funzione che ad ogni coppia ordinata di vettori $v, w \in V$ assegna uno scalare $< v, w > \in \mathbf{K}$ tale che

1) $< u + v, w > = < u, w > + < v, w > \ \forall u, v, w \in V$
2) $< u, v > = \overline{< v, u >} \, \forall u, v \in V$
3) $< \lambda u, v > = \lambda < u, v > \ \forall u, v \in V \forall \lambda \in \mathbf{K}$
4) $< v, v > \geq 0, \ \forall \ v \in V; < v, v > = 0 \Leftrightarrow \ v = 0$

Dalle condizioni 2) e 3) segue che $< u, \lambda v > = \bar{\lambda} < u, v > \ \forall u, v \in V \forall \lambda \in \mathbf{K}$

Esempi:
1) Siano $v = (x_1, \ldots, x_n)$ e $w = (y_1, \ldots, y_n)$ due vettori di $\mathbf{R}^n$, il prodotto scalare standard o canonico in $\mathbf{R}^n$ è definito da $< v, w > = x_1 y_1 + \ldots + x_n y_n$.

2) Siano $v = (x_1, \ldots, x_n)$ e $w = (y_1, \ldots, y_n)$ due vettori in $\mathbf{C}^n$, il prodotto scalare standard o Hermitiano in $\mathbf{C}^n$ è definito da $< v, w > = x_1 \bar{y}_1 + \ldots + x_n \bar{y}_n$.

3) Siano $v = (x_1, x_2)$ e $w = (y_1, y_2)$ due vettori in $\mathbf{R}^2$, un'altro prodotto scalare in $\mathbf{R}^2$ è definito da $< v, w > = x_1 y_1 - x_2 y_1 - x_1 y_2 + 4 x_2 y_2$.

4) In $V = M(n, \mathbf{K})$ si denota per B^* la matrice coniugata trasposta della matrice B (i.e. $b^*_{kj} = \bar{b}_{jk}$), $< A, B > = \mathrm{tr}(AB^*) = \mathrm{tr}(B^* A) = \sum_j (AB^*)_{jj} = \sum_j \sum_k a_{jk} b^*_{kj} = \sum_j \sum_k a_{jk} \bar{b}_{jk}$ definisce un prodotto scalare su V.
5) Sia $V = C^0([a, b]) = \{f : [a, b] \to \mathbf{R} \text{ continua } \}$, $< f, g > = \int_a^b f(x) g(x) dx$ è un prodotto scalare su V.

Definizione: La norma (o lunghezza) di v è $\|v\| := \sqrt{< v, v >}$. La distanza tra due vettori $v, w \in V$ è $\|v - w\|$. L'angolo θ fra $v, w \in V$ è dato da $\cos \theta = \frac{< v, w >}{\|v\| \|w\|}$.

Definizione: Uno spazio vettoriale V provvisto da un prodotto scalare $<, >$ è detto spazio metrico. Se reale si chiama Euclideo. Se complesso si dice spazio Hermitiano (oppure metrico unitario).

Teorema: Se V è uno spazio vettoriale con prodotto scalare su un campo $\mathbf{K}$ allora

1) $\|v\| > 0 \, \forall v \neq 0$
2) $\|\lambda v\| = |\lambda| \, \|v\| \ \forall v \in V, \forall \lambda \in \mathbf{K}$
3) $| < v, w > | \leq \|v\| \, \|w\| \ \forall v, w \in V$ (Cauchy-Schwarz inequality).
4) $\|v + w\| \leq \|v\| + \|w\| \ \forall v, w \in V$ (triangular inequality).

Dimostrazione di 3): Se $v = 0$ l'uguaglianza è soddisfatta. Se $v \neq 0$, scrivere $u = w - \frac{< w, v >}{\|v\|^2} v$. La disequazione segue da $< u, v > = 0$ e da $0 \leq \|u\|^2 = < w - \frac{< w, v >}{\|v\|^2} v, w - \frac{< w, v >}{\|v\|^2} v > = < w, w > - \frac{< w, v > < v, w >}{\|v\|^2} = \|w\|^2 - \frac{| < w, v > |^2}{\|v\|^2}$.
Dimostrazione di 4): $\|v + w\|^2 = \|v\|^2 + < v, w > + < w, v > + \|w\|^2 = \|v\|^2 + 2\mathrm{Re} < v, w > + \|w\|^2 \leq \|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$

Definizione: Sia V uno spazio metrico, due vettori $v, w \in V$ sono detti *ortogonali* se $< v, w > = 0$. Un insieme $S \subset V$ è detto *ortogonale* se i suoi elementi sono a due a due ortogonali. Un insieme $S \subset V$ è detto *ortonormale* se i suoi elementi oltre a essere ortogonale hanno norma 1.

Lemma: Siano $v_1, \ldots v_k$ vettori non-nulli di V. Se $\{v_1, \ldots v_k\}$ è ortogonale allora è linearmente indipendente.

Teorema (Gram-Schmidt): Ogni spazio vettoriale di dimensione finita provvisto di prodotto scalare ha una base ortogonale.

Dimostrazione: Si procede per induzione su n. Sia $\{v_1, \ldots, v_n\}$ una base qualsiasi di V. Scrivere $w_1 := v_1$. Se $\dim V = 1$ non c'e nulla da dimostrare. Supporre per induzione che il Teorema sia vero per ogni dimensione minore di $\dim V$. Supporre che $\{w_1, \ldots, w_{n-1}\}$ sono scelti in modo che formano una base ortogonale di Span $\{v_1, \ldots, v_{n-1}\}$. Scrivere $w_n = v_n - \sum_j \frac{<v_n, w_j>}{\|w_j\|^2} w_j$. Allora $w_n \neq 0$, altrimenti v_n sarebbe combinazione lineare di $w_1, \ldots, w_{n-1}$ e quindi di $v_1, \ldots, v_{n-1}$. Inoltre $< w_n, w_j >= 0 \ \forall j = 1, \ldots, n-1$. Quindi siccome $\{w_1, \ldots, w_{n-1}, w_n\}$ è un insieme ortogonale di vettori non-nulli allora devono essere linealmente indipendenti.

Teorema: Ogni spazio vettoriale metrico di dimensione finita ha una base ortonormale.
Dimostrazione: Sostituire il vettore w_k della base ortogonale per $u_k = \frac{w_k}{\|w_k\|}$.

Formulazione matriciale:
Sia V uno spazio metrico n-dimensionale e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Sia $g_{ij} =< v_i, v_j >$. Se $v = \sum_i x_i v_i$ e $w = \sum_j y_j v_j$ allora

$$< v, w >= \sum_{i,j} x_i g_{ij} y_j = X^t G Y,$$

dove $X^t = v_{\mathcal{B}} = (x_1, \ldots, x_n)$ e $Y^t = w_{\mathcal{B}} = (y_1, \ldots, y_n)$ e $G = (g_{ij})$. Quindi $< v_i, v_j >$ è noto se i valori g_{ij} sono noti. Viceversa, dati valori arbitrari $g_{ij} \in \mathbf{K}$, e definendo un prodotto $< v, w >$ che soddisfi le regole 1) - 4) si ottiene un prodotto scalare.
Ora se $\mathcal{B}' = \{w_1, \ldots, w_n\}$ è un'altra base e si ha $w_i = \sum_k p_{ki} v_k$ e $w_j = \sum_h p_{hj} v_h$ allora

$$g'_{ij} =< w_i, w_j >=< \sum_k p_{ki} v_k, \sum_h p_{hj} v_h >= \sum_{k,h} p_{ki} g_{kh} p_{hj}.$$

Quindi $G' = P^t G P$, dove $G' = (g'_{ij})$ and P è la matrice che ha per colonne le componenti dei vettori w_i relativi alla base $\mathcal{B}$.

Lemma: Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base ortogonale di V, allora $\forall v \in V \ v = \sum_j \frac{<v, v_j>}{\|v_j\|^2} v_j$.
Se $\mathcal{B}$ è ortonormale allora $\forall v \in V \ v = \sum_j < v, v_j > v_j$.

Corollario: Se $\mathcal{B}$ è una base ortogonale di V, allora $\forall v, w \in V \ \ < v, w >= \sum_j \frac{<v, v_j><v_j, w>}{\|v_j\|^2}$
e $\forall v \in V \ \|v\|^2 = \sum_j \frac{|<v, v_j>|^2}{\|v_j\|^2}$.
Se $\mathcal{B}$ è ortonormale allora $\forall v, w \in V \ \ < v, w >= \sum_j < v, v_j >< v_j, w >$ (Parseval)
e $\forall v \in V \ \|v\|^2 = \sum_j | < v, v_j > |^2$ (Pitagora).

Definizione: I numeri $\frac{<v, v_j>}{\|v_j\|^2}$ sono detti coefficienti di Fourier di v relativi alla base $\mathcal{B}$.

IL COMPLEMENTO ORTOGONALE DI UN SOTTOSPAZIO
Sia V uno spazio metrico, il complemento ortogonale $W^\perp$ di un sottospazio $W \subset V$, è il sottospazio vettoriale

$$W^\perp := \{v \in V \text{ tale che } < v, w >= 0 \ \forall w \in W\}.$$

Notare che se $U \perp W$ allora $U \subset W^\perp$, cioè: $W^\perp$ è il più grande sottospazio di V ortogonale a W.

Teorema: Sia V uno spazio euclideo di dimensione n e sia $W \subset V$ un sottospazio vettoriale di dimensione k. Allora 1) $V = W \oplus W^\perp$, 2) $(W^\perp)^\perp = W$.

Dimotrazione: a) Sia $\{v_1, \dots, v_k\}$ una b.o.n. di W e $\{v_1, \dots, v_n\}$ una b.o.n. di V. Basta dimostrare che $\text{Span}\{v_{k+1}, \dots, v_n\} = W^\perp$. Si noti che $\text{Span}\{v_{k+1}, \dots, v_n\} \subset W^\perp$. Sia $v \in W^\perp$, siccome $< v, v_1 >= \dots =< v, v_k >= 0$ allora $v =< v, v_{k+1} > v_{k+1} + \dots + < v, v_n > v_n$. Quindi $W^\perp \subset \text{Span}\{v_{k+1}, \dots, v_n\}$.
b) $W \perp W^\perp$, quindi $W \subset (W^\perp)^\perp$, ne segue da a) che $V = W^\perp \oplus (W^\perp)^\perp$, e siccome $\dim V = \dim W^\perp + \dim (W^\perp)^\perp = \dim W + \dim W^\perp$ risulta che hanno la stessa dimensione, per cui coincidono.

LA PROIEZIONE ORTOGONALE
Se V è uno spazio metrico di dimensione finita, $W \subset V$ un sottospazio vettoriale e $\{w_1, \dots, w_k\}$ una b.o.n. di W, il vettore proiezione ortogonale di v su W è definito da

$$P_W(v) := \sum_j^k < v, w_j > w_j.$$

Proposizione: $P_W(v)$ è il vettore in W più vicino a v. Cioè $\|v - P_W(v)\| < \|v - w\| \; \forall w \in W$ tale che $w \neq P_W(v)$.
Dimostrazione: $\|v - w\|^2 = \|v - P_W(v)\|^2 + \|P_W(v) - w\|^2$ e $\|P_W(v) - w\|^2 > 0$.

Definizione: L'applicazione $P_W : V \to V$ che a ogni vettore $v \in V$ assegna la sua proiezione ortogonale $P_W(v)$ su W è detta la proiezione ortogonale di V su W.

Si osservi che $P_W(w) = w \; \forall w \in W$ quindi $W \subset \text{Im}\, P_W$ e che $\ker P_W = W^\perp$. Inoltre $\forall\, v \in V \; v = P_W(v) + (v - P_W(v))$, dove $v - P_W(v) \in \ker P_W$, per cui $V = \ker P_W + \text{Im}\, P_W$ e quindi si ha che $V = \ker P_W \oplus \text{Im}\, P_W$. In particolare si ha che $\text{Im}\, P_W = W$.

Nota: La dimostrazione del teorema di Gram-Schmidt fornisce un algoritmo per costruire una base ortogonale da una base data. Ogni vettore w_j è ottenuto sottraendo dal vettore dato v_j la sua proiezione ortogonale su il sottospazio generato dai vettori precedenti. L'idea è ripetuta ad ogni paso. Per ottenere una base ortonormale si divide ogni vettore per la propria lunghezza (norma).

APPLICAZIONE AGGIUNTA

Definizione: Siano $(V, <,>_V)$ e $(W, <,>_W)$ due spazi metrici. L'aggiunta di una applicazione lineare $f : V \to W$ è l'applicazione $f^* : W \to V$ che soddisfa

$$< f(v), w >_W = < v, f^*(w) >_V \quad \forall v \in V, \forall w \in W.$$

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortonormale di $(V, <,>_V)$. Per ogni $v \in V$ e $w \in W$ si ha che $< v, f^*(w) >_V = < \sum_j < v, v_j >_V v_j, f^*(w) >_V = \sum_j < v, v_j >_V < v_j, f^*(w) >_V = \sum_j < v, v_j >_V < f(v_j), w >_W = < v, \sum_j < f(v_j), w >_W v_j >_V$. Quindi si definisce

$$f^*(w) := \sum_j \overline{< f(v_j), w >}_W v_j.$$

L'applicazione f^* è lineare giacché $< v, f^*(\alpha w + u) > = < f(v), \alpha w + u > = \bar{\alpha} < f(v), w > + < f(v), u > = \bar{\alpha} < v, f^*(w) > + < v, f^*(u) > = < v, \alpha f^*(w) + f^*(u) > \ \forall v \in V, \forall w \in W, \forall \alpha \in \mathbf{K}$.

Teorema: Sia $f : V \to W$ un'applicazione lineare, siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{w_1, \dots, w_m\}$ base ortonormali di V e W rispettivamente, inoltre sia A la matrice di f relativa alle basi $\mathcal{B}$ e $\mathcal{B}'$. Allora la matrice $f^*: W \to V$ rispetto alle basi $\mathcal{B}'$ e $\mathcal{B}$ è la trasposta coniugata di A (i.e. $M_{\mathcal{B}}(f^*) = \bar{A}^t := A^*$).
Dimostrazione: Se $f(v_i) = \sum_k a_{ki} w_k$ e $f^*(w_j) = \sum_k b_{kj} v_k$, si ha che

$$a_{ji} = < \sum_k a_{ki} w_k, w_j > = < f(v_i), w_j > = < v_i, f^*(w_j) > = < v_i, \sum_k b_{kj} w_k > = \bar{b}_{ij}.$$

Teorema: Sia $f: (V, <,>_V) \to (W, <,>_W)$ lineare. Allora 1) $V = \ker f \oplus \ \mathrm{Im}\ f^*$ e
2) $W = \ker f^* \oplus \ \mathrm{Im}\ f$.

Dimostrazione: 1) $V = \ker f \oplus (\ker f)^\perp$ quindi basta dimostrare che $(\ker f)^\perp = \ \mathrm{Im}\ f^*$.
Analogamente per 2) basta verificare che $(\ \mathrm{Im}\ f)^\perp = \ker f^*$:
Se $w \in \ker f^* \Rightarrow 0 = < v, f^*(w) > = < f(v), w > \Rightarrow w \in (\ \mathrm{Im}\ f)^\perp$. Viceversa, se $w \in (\ \mathrm{Im}\ f)^\perp \subset W \Rightarrow 0 = < f(v), w > = < v, f^*(w) > \forall v \in V \Rightarrow f^*(w) = 0 \Rightarrow w \in \ker f^*$.

Nel caso $V = \mathbf{R}^n$, $W = \mathbf{R}^m$ ed $f(X) = AX$ dove $A \in M(m, n; \mathbf{R})$ si ha che $\mathbf{R}^n = \ker A \oplus \ \mathrm{Im}\ A^t$ e $\mathbf{R}^m = \ker A^t \oplus \ \mathrm{Im}\ A$.

Esempio: Siano $V = \mathbf{R}^3$, $W = \mathbf{R}$ e $f: V \to W$ definita da $f(v) = < v, u >_V \quad \forall v \in V$, dove $u = (a, b, c) \in V$. Si vuole calcolare $f^*: W \to V$, dove il prodotto scalare in W è il prodotto di numeri reali. $< v, f^*(x) >_V = < f(v), x >_W = f(v)x = < v, u >_V x = < v, xu >_V \quad \forall v \in V, \ \forall x \in W \Rightarrow f^*(x) = xu$. Si verifica quindi che la matrice dell'applicazione aggiunta f^* di f è $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

ENDOMORFISMI AUTOAGGIUNTI

Definizione: Un endomorfismo $f : (V, <,>) \to (V, <,>)$ è detto *autoaggiunto* se $f = f^*$. Se V è reale f è detto *simmetrico*, se V è complesso f è chiamato *Hermitiano*. Cioè, se

$$< f(v), w > = < v, f(w) > \ \forall v, w \in V.$$

Proposizione: 1) Gli autovalori di un endomorfismo autoaggiunto sono reali. 2) Gli autovettori corrispondenti a autovalori distinti sono ortogonali.

Dimostrazione: 1) Presso un autovettore v relativo all'autovalore λ, siccome $v \neq 0$ si ha che $\lambda < v, v >=< \lambda v, v >=< f(v), v >=< v, f^*(v) >=< v, \lambda v >= \bar{\lambda} < v, v >$ implica che $\lambda = \bar{\lambda}$.
2) Siano v e w autovettori relativi agli autovalori λ e μ rispettivamente, siccome $\lambda \neq \mu$ allora $\lambda < v, w >=< \lambda v, w >=< f(v), w >=< v, f^*(w) >=< v, \mu v >= \mu < v, w >$ implica $< v, w >= 0$.

Definizione: Sia $f : V \to V$ lineare, un sottospazio vettoriale $W \subset V$ è detto f-invariante se $f(W) \subset W$.

Proposizione: Se f è un endomorfismo autoaggiunto e W un sottospazio f-invariante allora il sottospazio $W^\perp$ è anche esso f-invariante.

Dimostrazione: se $f = f^*$ allora se $v \in W^\perp$ si ha che $< f(v), w >=< v, f(w) >= 0 \ \forall w \in W$. Inoltre, se $f(w) \in W \ \forall w \in W$ si ha dunque che $f(v) \in W^\perp$.

Se lo spazio vettoriale V è euclideo si ha il seguente risultato.

Proposizione: Ogni applicazione autoaggiunta su V ha un autovettore reale.

Dimostrazione: Per il Teorema Fondamentale dell'Algebra ogni applicazione lineare su uno spazio vettoriale di dimensione finita sul campo $\mathbf{C}$ ha un autovalore λ e quindi un autovettore. Siccome f è autoaggiunta λ è un numero reale. Se V è reale A e $A - \lambda I$ hanno coefficienti reali. Inoltre la matrice $A - \lambda I$ è non-singolare quindi il sistema lineare $(A - \lambda I)X = 0$ ha una soluzione reale non banale ($X \neq 0$).

Teorema spettrale: Sia $(V, <,>)$ uno spazio metrico di dimensione finita e sia $f : V \to V$ un endomorfismo autoaggiunto. Allora c'e una base ortonormale di V costituita di autovettori di f.

Dimostrazione: Si procede per induzione sulla dimensione di V. Se $\dim V > 0$ allora f ha un autovettore v. Si consideri $v_1 = \frac{v}{\|v\|}$. Se $\dim V = 1$ il Teorema è vero. Supporre che il Teorema sia valido per ogni spazio metrico di dimensione minore di $\dim V$. Sia W il sottospazio di dimensione 1 generato da v_1. Notare che W, e quindi il suo complemento ortogonale $W^\perp$ sono f-invarianti. Inoltre il prodotto scalare di V ristretto a $W^\perp$ è un prodotto scalare. Denotare per g la restrizione di f a $W^\perp$. Si verifica facilmente che g è auttoaggiunta rispetto al prodotto scalare di $W^\perp$. Per ipotesi induttiva $W^\perp$ ha una base ortonormale $\{v_2, \ldots, v_n\}$ che consiste di autovettori di g. Ognuno di questi vettori è anche autovettore di f, e siccome $V = W \oplus W^\perp$, $\{v_1, v_2, \ldots, v_n\}$ è la base cercata.

Il viceversa del Teorema Spettrale è vero

Proposizione: Sia V uno spazio euclideo di dimensione finita e $f : V \to V$ una applicazione lineare. Allora f è autoaggiunta se e solo se V ha una base ortonormale costituita di autovettori di f.

Dimostrazione: rimane dimostrare che se V ha una base ortonormale $\mathcal{B} = \{u_1, \ldots, u_n\}$ di autovettori di f allora $f = f^*$. In questa base $v = \sum_i < v, u_i > u_i$ e $w = \sum_j < w, u_j > u_j$ quindi $f(v) = \sum_i < v, u_i > f(u_i) = \sum_i < v, u_i > \lambda_i u_i$ e $f(w) = \sum_j < w, u_j > f(u_j) = \sum_j < w, u_j > \lambda_j u_j$. Allora $< f(v), w >=< \sum_i < v, u_i > \lambda_i u_i, \sum_j < w, u_j > u_j >= \sum_i \sum_j \lambda_i < v, u_i >< w, u_j >< u_i, u_j >= \sum_j \lambda_j < v, u_j >< w, u_j >= \sum_i \sum_j \lambda_j < v, u_i >< w, u_j >< u_i, u_j >=< \sum_i < v, u_i > u_i, \sum_j < w, u_j > \lambda_j u_j >=< v, f(w) >$.

Esempio: La proiezione ortogonale $P_W : V \to V$ di V su W è autoaggiunta:
sia $\{w_1, \ldots, w_k\}$ una base ortonormale di W allora $\forall\, v, v' \in V < P_W(v), v' >=$
$=< \sum_j^k < v, w_j > w_j, v' >=< \sum_j^k < v, w_j >< w_j, v' >=< v, \sum_j^k w_j < w_j, v' >=< v, P_W(v') >$.

ENDOMORFISMI ANTISIMMETRICI

Definizione: Un endomorfismo f è detto antisimmetrico se $f = -f^*$. Cioè, se

$$< f(v), w >= - < v, f(w) >, \forall v, w \in V.$$

In particolare $< f(v), v >= 0\ \forall v \in V$. Inoltre se λ è un autovalore di f allora $\lambda = 0$. Giacchè se v è un autovettore relativo a λ allora $< f(v), v >= \lambda < v, v >= 0 \Rightarrow \lambda = 0$.

Esempi: 1) Se $\dim V = 2, A = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$. Quindi, se $v = (x, y) \Rightarrow Av = a(-y, x)$. Osservare che Av si ottiene ruotando di $\frac{\pi}{2}$ in senso antiorario e moltiplicando per a.
2) Se $V = \mathbf{R}^3$, sia $f(v) = w \wedge v\ \forall v \in V$. Notare che $w \in \ker f$ quindi w è autovettore di f relativo all'autovalore 0. Se $w = (a, b, c)$ la matrice di f nella base canonica è $\begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$.

ISOMETRIE LINEARI

Definizione: Siano $(V, <,>_V)$ e $(W, <,>_W)$ due spazi metrici su un campo $\mathbf{K}$. Una applicazione lineare $f : V \to W$ conserva i prodotti scalari se

$$< f(v), f(v') >_W =< v, v' >_V \quad \forall v, v' \in V.$$

Se $\dim V = \dim W$ una tale applicazione è detta isometria (lineare).

Esempio: Si consideri $\mathbf{R}^3$ con il prodotto scalare canonico e sia V lo spazio delle matrici antisimmetriche di ordine 3 con il prodotto scalare $< A, B >= \frac{1}{2}\mathrm{tr}(AB^t)$. L'applicazione lineare $f\colon \mathbf{R}^3 \to V$ definita da $f(x_1, x_2, x_3) = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}$ è una isometria, giacchè se $A = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{pmatrix}$ si ha che $\mathrm{tr}(AB^t) = 2(x_1y_1 + x_2y_2 + x_3y_3)$.

Proposizione: Se V e W sono spazi metrici su un campo $\mathbf{K}$. Allora $f\colon V \to W$ è una isometria se e solo se trasforma basi ortonormali di V in basi ortonormali di W.

Dimostrazione: Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormale di V, allora $\{f(v_1), \ldots, f(v_n)\}$ è una base ortonormale di W, poichè $< f(v_i), f(v_j) >_W =< v_i, v_j >_V = \delta_{ij}$.

Proposizione: Siano V e W due spazi metrici di dimensione finita n sul campo reale $\mathbf{K} = \mathbf{R}$. Allora $f\colon V \to W$ conserva prodotti scalari se e solo se conserva le norme, cioè se $\|f(v)\| = \|v\|\ \forall v \in V$.
Dimostrazione: segue da $\|v\|_V^2 + 2 < v, v' >_V + \|v'\|_V^2 =< v + v', v + v' >_V =< f(v + v'), f(v + v') >_W =< f(v) + f(v'), f(v) + f(v') >_W = \|f(v)\|_W^2 + 2 < f(v), f(v') >_W + \|f(v')\|_W^2 = \|v\|_W^2 + 2 < f(v), f(v') >_W + \|v'\|_W^2$.

TRASFORMAZIONI UNITARIE

Definizione: Sia $(V, <,>)$ uno spazio Hermitiano, un endomorfismo f è detto unitario se è invertibile e $f^{-1} = f^*$.
cioè: f è unitario se e solo se esiste f^* e se $ff^* = f^*f = \mathrm{id}$ dove id denota l'applicazione identità.

Proposizione: f è unitario se e solo conserva il prodotto scalare.

Dimostrazione: $\Rightarrow$). $< f(v), f(v') >=< v, f^* f(v') >=< v, \mathrm{id}(v') >=< v, v' >$ implica f conserva il prodotto scalare.
$\Leftarrow$) $< f(v), v' >=< f(v), f f^{-1}(v') >=< v, f^{-1}(v') >$ quindi $f^* = f^{-1}$

Proposizione: Se $\lambda \in \mathbf{C}$ è un autovalore di una trasformazione unitaria allora $|\lambda| = 1$.
Dimostrazione: $\lambda < v, v >=< \lambda v, v >=< f(v), v >=< v, f^*(v) >=< v, f^{-1}(v) >=< v, \lambda^{-1} v >= \bar{\lambda}^{-1} < v, v >$ siccome $\|v\| \neq 0$ si ha $\lambda = \bar{\lambda}^{-1}$.

Proposizione: Se f un endomorfismo unitario e W un sottospazio f-invariante allora $W^\perp$ è f-invariante.

Teorema: Sia V uno spazio metrico di dimensione finita. Allora l'endomorfismo f è unitario se e solo se esiste una base ortonormale di V rispetto alla quale f è una matrice diagonale con numeri complessi di norma 1 sulla diagonale.

Dimostrazione: segue dalle proposizioni precedenti e del Teorema sugli endomorfismi normali che si studiano nel seguito.

Definizione: Una matrice invertibile A è unitaria se $A^{-1} = A^*$.
Osservare che A è unitaria se e solo se l'insieme delle righe (colonne) è ortogonale.

Definizioni: una matrice reale o complessa è detta ortogonale se $AA^t = I$ $(A^{-1} = A^t)$.
Osservare che una matrice reale ortogonale è unitaria e che una matrice unitaria è ortogonale se e solo se ogni suo coefficiente è reale.

Inoltre osservare che se $I = AA^t$, siccome $\det A = \det A^t$, si ha che $1 = \det I = \det(AA^t) = \det A \det A^t = (\det A)^2$ quindi $\det A = 1$ oppure $\det A = -1$.

Lemma: Sia f un endomorfismo unitario, $\mathcal{B}$ una base ortonormale di V allora la matrice rappresentativa di f rispetto alla base $\mathcal{B}$ è una matrice unitaria.

FORMULAZIONE MATRICIALE DEL TEOREMA SPETRALE

Definizione: Una matrice complessa che soddisfa $A = A^* (:= \bar{A}^t)$ è detta Hermitiana. Una matrice reale che soddisfa $A = A^*$ è detta simmetrica.

Teorema spettrale: Se A è una matrice Hermitiana allora esiste una matrice unitaria P $(P^{-1} = \overline{P}^t)$ tale che $\overline{P}^t AP$ è diagonale.
If A è una matrice reale simmetrica allora esiste una matrice reale ortogonale P $(P^{-1} = P^t)$ tale che $P^t AP$ è diagonale.

TRASFORMAZIONI ORTOGONALI SU SPAZI EUCLIDEI

Definizione: Sia $(V, <, >)$ uno spazio Euclideo, un endomorfismo f è detto ortogonale se è invertibile e se $f^{-1} = f^*$.

Lemma: Siano $(V, <, >)$ uno spazio Euclideo, f un endomorfismo ortogonale e $\mathcal{B}$ una base ortonormale di V allora la matrice A rappresentativa di f relativa alla base $\mathcal{B}$ è una matrice ortogonale. $(A^{-1} = A^t)$.

Definizione: Sia A una matrice ortogonale, se $\det A = 1$ allora è detta *rotazione*. Invece, se $\det A = -1$ è detta *riflessione*. Una rotazione è detta *semplice* se esiste un sottospazio $W \subset V$ A-invariante di $\dim W = 2$, $A|_W$ è una rotazione e $A|_{W^\perp} = I$. Una riflessione è detta *semplice* se esiste $0 \neq v \in V$ tale che $Av = -v$ e $A|_{v^\perp} = I$

MATRICI ORTOGONALI REALI

Matrici ortogonali di ordine 2.

Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con $\det A = ad - cb = 1$. Da $A^t A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ne segue che

$$\begin{aligned}
a^2 + c^2 = 1 &\quad\Rightarrow\quad a = \cos\theta_1,\ c = \sin\theta_1, \\
b^2 + d^2 = 1 &\quad\Rightarrow\quad b = \cos\theta_2,\ d = \sin\theta_2, \\
ab + cd = 0 &\quad\Rightarrow\quad 0 = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 = \cos(\theta_2 - \theta_1), \\
ad - cb = 1 &\quad\Rightarrow\quad 1 = \cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2 = \sin(\theta_2 - \theta_1).
\end{aligned}$$

allora $\theta_2 - \theta_1 = \frac{\pi}{2}$ quindi $A = \begin{pmatrix} \cos\theta_1 & \cos\theta_2 \\ \sin\theta_1 & \sin\theta_2 \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & \cos(\theta_1 + \frac{\pi}{2}) \\ \sin\theta_1 & \sin(\theta_1 + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}$.
Geometricamente produce una rotazione di angolo θ in senso antiorario.

Se $\det A = -1$, da calcoli analoghi si ottiene $\theta_2 - \theta_1 = -\frac{\pi}{2}$ quindi $A = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \sin\theta_1 & -\cos\theta_1 \end{pmatrix}$.

Con $\theta = 0$ si ottiene $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, geometricamente produce una riflessione rispetto all'asse x.

Classificazione delle trasformazioni ortogonali di $\mathbf{R}^2$ secondo la dimensione del sottospazio dei punti fissi $U = \{u \text{ tali che } f(u) = u\}$.

Se $\dim U = 2$ allora tutti i punti sono fissi, quindi $f = id$.
Se $\dim U = 1$, se $U = \text{ Span } \{u\}$ allora $U^\perp = \text{ Span } \{v\}$ è f-invariante quindi $f(v) \perp u$, ne segue che $f(v) = -v$. Di conseguenza la matrice di f nella base $\{\frac{u}{\|u\|}, \frac{v}{\|v\|}\}$ è $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Se $\dim U = 0$ allora f è una rotazione e la sua matrice è $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Classificazione delle trasformazioni ortogonali $\mathbf{R}^3$ secondo la dimensione del sottospazio dei punti fissi $U = \{u \text{ tali che } f(u) = u\}$.

Se $\dim U = 3$ allora tutti i punti sono fissi, quindi $f = id$.
Se $\dim U = 2$, allora $\dim U^\perp = 1$. Quindi se $v \in U^\perp$ siccome $U^\perp$ è f-invariante $f(v) = -v$. Di conseguenza la matrice di f è $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.
Se $\dim U = 1$, allora $\dim U^\perp = 2$. Siccome $U^\perp$ è f-invariante la matrice di $f|_{U^\perp}$ è una rotazione. Quindi in una b.o.n. è $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$.
Il polinomio caratteristico di f ha grado 3 (ed è reale), quindi se z è radice lo è anche $\bar{z}$ dunque una delle tre radici deve essere reale. Se $\dim U = 0$ ne segue che $\exists v \in \mathbf{R}^3$ tale che $f(v) = -v$. Allora $W = v^\perp$ ha dim 2. La restrizione $<,> |_W$ del prodotto scalare a W è un prodotto scalare. Siccome f non ha punti fissi la sua restrizione $f|_W : W \to W$ non ha punti fissi e quindi è una rotazione. Si ha dunque che la matrice di f in una b.o.n. è $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$.

Teorema: Sia V uno spazio metrico di dimensione n sul campo $\mathbf{R}$ e sia f un endomorfismo ortogonale. Allora esiste un sottospazio $W \subset V$ f-invariante con $\dim W = 1$ oppure $\dim W = 2$.

Dimostrazione: Sia λ_0 una radice del polinomio caratteristico, se $\lambda_0 \in \mathbf{R}$ è un autovalore allora esiste un autovettore v_0 a lui relativo. Quindi $W = \text{ Span } \{v_0\}$ è un sottospazio f-invariante di $\dim W = 1$.

Se λ_0 non appartiene a $\mathbf{R}$ allora $\lambda_0 = \alpha + i\beta,\ \alpha \in \mathbf{R},\ \beta \in \mathbf{R}$. Il fatto che $\det(A - \lambda_0 I) = 0$ implica che esiste $Z = X + iY \neq 0 \in \ker(A - \lambda_0 I)$ dove $X, Y \in \mathbf{R}^n$ i.e. $AZ = \lambda_0 Z$. Siccome A è una matrice reale $AX + iAY = (\alpha + i\beta)(X + iY) = (\alpha X - \beta Y) + i(\beta X + \alpha Y)$ si ha $\begin{cases} AX = \alpha X - \beta Y \\ AY = \beta X + \alpha Y \end{cases}$
Allora $W =$ Span $\{X, Y\}$ è A-invariante. Siccome $Z \neq 0$ allora $X \neq 0$ o $Y \neq 0$ o ambedue, si ha dunque $\dim W = 1$ oppure $\dim W = 2$ secondo sia il caso.

Teorema: Sia V uno spazio metrico di dimensione finita, sia $f : V \to V$ una applicazione ortogonale allora esiste una base ortonormale di V relativa alla quale f è una matrice del tipo

$$\begin{pmatrix} \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} & 0 & \cdots & 0 & 0 \\ 0 & \begin{pmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{pmatrix} & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & -I & 0 \\ 0 & 0 & \cdots & 0 & I \end{pmatrix}$$

Dimostrazione: Se $\dim V = 1$ o 2 il Teorema è stato già verificato, supporre quindi che $\dim V \geq 3$. Sia $W \subset V$ un sottospazio f-invariante di $\dim 1$ o 2, allora $f|_W$ è ortogonale quindi esiste una base di W nella quale la matrice rappresentativa è: se $\dim W = 2$ $[f|_W] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ o $[f|_W] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, se $\dim W = 1$, $[f|_W] = (-1)$ or $[f|_W] = (1)$. Inoltre siccome $W^\perp$ è f-invariante e $\dim W < \dim V$ si può applicare il procedimento descritto prima a $W^\perp$ fino a ottenere la base cercata, riordinando i sottospazi invarianti se necessario.

Corollario: Ogni applicazione ortogonale è il prodotto di rotazioni e riflessioni semplici che commutano.

Osservazione: In generale un endomorfismo ortogonale sul campo reale non diagonalizza.

ENDOMORFISMI NORMALI

Definizione: Sia V uno spazio vettoriale su $\mathbf{C}$. Un endomorfismo $f : V \to V$ è detto normale se è invertibile e se $ff^* = f^*f$.

Definizione: Una matrice complessa invertibile A è detta normale se $AA^* = A^*A$. In particolare endomorfismi autoaggiunti, unitari oppure ortogonali sono normali.

Proposizione: Sia $f : V \to V$ un endomorfismo normale e W un sottospazio f-invariante allora $W^\perp$ é f-invariante.

Teorema: Sia $(V, \mathbf{C})$ uno spazio metrico di dimensione finita. Allora un endomorfismo $f : V \to V$ è normale se e solo se esiste una base ortonormale di V relativa alla quale f è diagonale.

Dimostrazione: Basta dimostrare che f è normale se e solo se $f = g + ih$ dove g e h sono endomorfismi autoaggiunti che commutano. Si procede applicando il Teorema spettrale per endomorfismi autoaggiunti e si verifica che g e h diagonalizzano nella stessa base. Prendere $g = \frac{f+f^*}{2}$ e $h = \frac{f-f^*}{2i}$.

Corollario: Se A è una matrice normale allora esiste una matrice unitaria P (i.e. $P^{-1} = \bar{P}^t$) tale che $\bar{P}^t AP$ è diagonale.

MOVIMENTO RIGIDO o ISOMETRIA AFFINE

Definizione: Una varietà affine M di uno spazio vettoriale V è l'imagine di un sottospazio W mediante una traslazione $t_a\colon V \to V$ definita da $t_a(v) = a + v$, dove $a \in V$. (Notare che una traslazione non è lineare).

$M = t_a(W) = a + W = \{a + w \text{ tali che } w \in W\}$. Il sottospazio W è detto la giacitura di M.

Definizione: Un movimento (rigido) o isometria affine $m\colon V \to V$ è la composizione di una isometria (lineare) con una traslazione, cioè $m = t_a \circ f$.

Si osservi che m conserva le distanze: $\|m(v) - m(v')\| = \|v - v'\| \ \forall v, v' \in V$.

Viceversa, se m conserva le distanze allora m è un movimento.

Dimostrazione: 1) Se f conserva le distanze e $f(0) = 0$ allora f conserva il prodotto scalare: $\|f(v)\| = \|f(v) - f(0)\| = \|v - 0\| = \|v\|$, quindi f è una isometria lineare (trasformazione ortogonale). 2) l'applicazione $f\colon V \to V$ definita da $f(v) = m(v) - m(0)$ soddisfa 1), allora $m(v) = f(v) + m(0) = (t_a \circ f)(v)$, dove $a = m(0)$.