Scheda riassuntiva di Geometria 1

Alcuni accenni alla geometria di $\mathbb{R}^3$

Si definisce prodotto scalare la forma bilineare simmetrica unicamente determinata da $\langle \underline{e_i}, \underline{e_j} \rangle = \delta_{ij}$. Vale la seguente identità: $\langle (x,y,z), (x',y',z') \rangle = xx' + yy' + zz'$.

Inoltre $\langle \underline{a}, \underline{b} \rangle = |\underline{a}|\,|\underline{b}| \cos(\theta)$, dove θ è l'angolo compreso tra i due vettori. Due vettori $\underline{a}$, $\underline{b}$ si dicono ortogonali se e solo se $\langle \underline{a}, \underline{b} \rangle = 0$.

Si definisce prodotto vettoriale la forma bilineare alternante da $\mathbb{R}^3 \times \mathbb{R}^3$ in $\mathbb{R}^3$ tale che $\underline{e_1} \times \underline{e_2} = \underline{e_3}$, $\underline{e_2} \times \underline{e_3} = \underline{e_1}$, $\underline{e_3} \times \underline{e_1} = \underline{e_2}$ e $\underline{e_i} \times \underline{e_i} = \underline{0}$. Dati due vettori (x,y,z) e (x',y',z'), si può determinarne il prodotto vettoriale informalmente come:

$$\begin{vmatrix} \underline{e_1} & \underline{e_2} & \underline{e_3} \\ x & y & z \\ x' & y' & z' \end{vmatrix}.$$

Vale l'identità $|\underline{a} \times \underline{b}| = |\underline{a}|\,|\underline{b}| \sin(\theta)$, dove θ è l'angolo con cui, ruotando di θ in senso antiorario $\underline{a}$, si ricade su $\underline{b}$. Due vettori $\underline{a}$, $\underline{b}$ si dicono paralleli se $\exists\, k \mid \underline{a} = k\underline{b}$, o equivalentemente se $\underline{a} \times \underline{b} = \underline{0}$. Altrettanto si può dire se $\langle \underline{a}, \underline{b} \rangle = |\underline{a}|\,|\underline{b}|$ (i.e. $\cos(\theta) = 1 \implies \theta = 0$).

Una retta in $\mathbb{R}^3$ è un sottospazio affine della forma $\underline{v} + \operatorname{Span}(\underline{r})$. Analogamente un piano è della forma $\underline{v} + \operatorname{Span}(\underline{x}, \underline{y})$.

Nella forma cartesiana, un piano è della forma $ax + by + cz = d$, dove (a,b,c) è detta normale del piano. Una retta è l'intersezione di due piani, e dunque è un sistema lineare di due equazioni di un piano. Due piani sono perpendicolari fra loro se e solo se le loro normali sono ortogonali. Due piani sono paralleli se e solo se le loro normali sono parallele. Il vettore $\underline{r}$ che genera lo Span di una retta che è intersezione di due piani può essere computato come prodotto vettoriale delle normali dei due piani.

Valgono le seguenti identità:

- $\underline{a} \times (\underline{b} \times \underline{c}) = \langle \underline{a}, \underline{c} \rangle\, \underline{b} - \langle \underline{a}, \underline{b} \rangle\, \underline{c}$ (*identità di Lagrange*),
- $\underline{a} \times (\underline{b} \times \underline{c}) + \underline{b} \times (\underline{c} \times \underline{a}) + \underline{c} \times (\underline{a} \times \underline{b}) = \underline{0}$ (*identità di Jacobi*).

Dati tre punti $\underline{a}$, $\underline{b}$, $\underline{c}$, il volume del parallelepipedo individuato da questi punti è:

$$\left| \det \begin{pmatrix} \underline{a} \\ \underline{b} \\ \underline{c} \end{pmatrix} \right| = |\langle \underline{a}, \underline{b} \times \underline{c} \rangle|.$$

Tre punti sono complanari se e solo se il volume di tale parallelpipedo è nullo (infatti questo è equivalente a dire che almeno uno dei tre punti si scrive come combinazione lineare degli altri due).

Proprietà generali di uno spazio vettoriale

Uno spazio vettoriale V su un campo $\mathbb{K}$ soddisfa i seguenti assiomi:

- $(V, +)$ è un gruppo abeliano,
- il prodotto esterno da $\mathbb{K} \times V$ in V è associativo rispetto agli scalari (i.e. $a(b\underline{v}) = (ab)\underline{v}$),
- $1_{\mathbb{K}} \cdot \underline{v} = \underline{v}$,
- il prodotto esterno è distributivo da ambo i lati (i.e. $(a+b)\underline{v} = a\underline{v} + b\underline{v}$ e $a(\underline{v} + \underline{w}) = a\underline{v} + a\underline{w}$.

Un insieme di vettori I si dice linearmente indipendente se una qualsiasi combinazione lineare di un suo sottinsieme finito è nulla se e solo se i coefficienti dei vettori sono tutti nulli. Si dice linearmente dipendente in caso contrario.

Un insieme di vettori G si dice generatore di V se ogni vettore di V si può scrivere come combinazione lineare di un numero finito di elementi di G, ossia se $V = \operatorname{Span}(G)$.

Una base è un insieme contemporaneamente linearmente indipendente e generatore di V. Equivalentemente una base è un insieme generatore minimale rispetto all'inclusione e un insieme linearmente indipendente massimale, sempre rispetto all'inclusione. Ogni spazio vettoriale, anche quelli non finitamente generati, ammettono una base. La dimensione della base è unica ed è il numero di elementi dell'insieme che è base.

Dato un insieme linearmente indipendente I in uno spazio di dimensione finita, tale insieme, data una base $\mathcal{B}$, può essere esteso a una base T che contiene I e che è completato da elementi di $\mathcal{B}$.

Analogamente, dato un insieme generatore finito G, da esso si può estrarre sempre una base dello spazio.

Uno spazio vettoriale fondato su un campo infinito con un insieme di vettori infinito non è mai unione finita di sottospazi propri. Un insieme linearmente indipendente di V con esattamente $\dim V$ elementi è una base di V. Analogamente, un insieme generatore di V con esattamente $\dim V$ elementi è una base di V. In generale, quando il campo su cui fonda lo spazio vettoriale è ambiguo, si scrive $\dim_{\mathbb{K}} V$ o $[V : \mathbb{K}]$ per indicarne la dimensione relativa al campo $\mathbb{K}$ (per esempio un $\mathbb{C}$-spazio è compatibilmente anche un $\mathbb{R}$-spazio).

Sia $\mathcal{B} = \{\underline{v_1}, \ldots, \underline{v_n}\}$ una base ordinata dello spazio vettoriale V.

- $\{\underline{0}\}$ e V sono detti sottospazi banali,
- lo Span di n vettori è il più piccolo sottospazio di V contenenti tali vettori,
- $\operatorname{Span}(\mathcal{B}) = V$,
- $\operatorname{Span}(\varnothing) = \{\underline{0}\}$,
- $U \cap W$ è sempre un sottospazio se U e W sono due sottospazi di V,
- $U \cup W$ è un sottospazio se e solo se $U \subseteq W$ o $U \supseteq W$ (e quindi $U \cup W = U$ o $U \cup W = W$),
- dato X generatore di V, $X \setminus \{\underline{x_0}\}$ genera $V \iff \underline{x_0} \in \operatorname{Span}(X \setminus \{\underline{x_0}\})$,
- $X \subseteq Y$ è un sottospazio di $Y \iff \operatorname{Span}(X) = X$,
- $\operatorname{Span}(X) \subseteq Y \iff X \subseteq Y$, se Y è uno spazio,
- $\operatorname{Span}(\operatorname{Span}(A)) = \operatorname{Span}(A)$,
- se I è un insieme linearmente indipendente di V, allora $|I| \leq \dim V$,
- se G è un insieme generatore di V, allora $|G| \geq \dim V$,
- $[\underline{v}]_{\mathcal{B}}$ è la rappresentazione di $\underline{v}$ nella base ordinata $\mathcal{B}$, ed è un vettore di $\mathbb{K}^n$ che alla coordinata i-esima associa il coefficiente di $\underline{v_i}$ nella combinazione lineare di $\underline{v}$ nella base $\mathcal{B}$,
- la rappresentazione nella base $\mathcal{B}$ è sempre unica ed esiste sempre (è quindi un isomorfismo tra V e $\mathbb{K}^n$),
- si definisce base canonica di $\mathbb{K}^n$ la base $e = \{\underline{e_1}, \ldots, \underline{e_n}\}$, dove $\underline{e_i}$ è un vettore con tutte le coordinate nulle, eccetto per la i-esima, che è pari ad 1 (pertanto $\dim \mathbb{K}^n = n$),
- una base naturale di $M(m,n,\mathbb{K})$ è data da $\mathcal{B} = \{E_{11}, E_{12}, \ldots, E_{1n}, \ldots, E_{mn}\}$, dove E_{ij} è una matrice con tutti gli elementi nulli, eccetto quello nel posto (i,j), che è pari ad 1 (dunque $\dim M(m,n,\mathbb{K}) = mn$),
- le matrici A di taglia n tali che $A^\top = A$ formano il sottospazio $\operatorname{Sym}(n,\mathbb{K})$ di $M(n,\mathbb{K})$, detto sottospazio delle matrici simmetriche, la cui base naturale è data da $\mathcal{B}' = \{E_{ij} + E_{ji} \in \mathcal{B} \mid i < j\} \cup \{E_{ij} \in \mathcal{B} \mid i = j\}$, dove $\mathcal{B}$ è la base naturale di $M(m,n,\mathbb{K})$ (dunque $\dim \operatorname{Sym}(n,\mathbb{K}) = \frac{n(n+1)}{2}$),
- le matrici A di taglia n tali che $A^\top = -A$ formano il sottospazio $\Lambda(n,\mathbb{K})$ di $M(n,\mathbb{K})$, detto sottospazio delle matrici antisimmetriche, la cui base naturale, se $\operatorname{char}\mathbb{K} \neq 2$, è data da $\mathcal{B}' = \{E_{ij} - E_{ji} \in \mathcal{B} \mid i < j\}$, dove $\mathcal{B}$ è la base naturale di $M(m,n,\mathbb{K})$ (dunque $\dim \Lambda(n,\mathbb{K}) = \frac{n(n-1)}{2}$),
- se invece $\operatorname{char}\mathbb{K} = 2$, le matrici antisimmetriche sono esattamente le matrici simmetriche,
- poiché $\operatorname{Sym}(n,\mathbb{K}) \cap \Lambda(n,\mathbb{K}) = \{\underline{0}\}$ e $\dim \operatorname{Sym}(n,\mathbb{K}) + \dim \Lambda(n,\mathbb{K}) = \dim M(n,\mathbb{K})$, vale che $M(n,\mathbb{K}) = \operatorname{Sym}(n,\mathbb{K}) \oplus \Lambda(n,\mathbb{K})$,
- una base naturale di $\mathbb{K}[x]$ è data da $\mathcal{B} = \{x^n \mid n \in \mathbb{N}\}$, mentre una di $\mathbb{K}_t[x]$ è data da $\mathcal{B} \cap \mathbb{K}_t[x] = \{x^n \mid n \in \mathbb{N} \wedge n \leq t\}$ (quindi $\dim \mathbb{K}[x] = \infty$ e $\dim \mathbb{K}_t[x] = t+1$),
- una base naturale di $\mathbb{K}$ è $1_{\mathbb{K}} = \{1_{\mathbb{K}}\}$ (quindi $\dim \mathbb{K} = 1$),
- un sottospazio di dimensione 1 si definisce *retta*, uno di dimensione 2 *piano*, uno di dimensione 3 *spazio*, e, infine, uno di dimensione $n-1$ un iperpiano,

- un iperpiano Π è sempre rappresentabile da un'equazione cartesiana nelle coordinate della rappresentazione della base (infatti ogni iperpiano è il kernel di un funzionale $f \in V^*$, e $M_{1_\mathbb{K}}^{\mathcal{B}}(f)\,[\underline{v}]_{\mathcal{B}} = 0$ è l'equazione cartesiana; è sufficiente prendere una base di Π e completarla a base di V con un vettore $\underline{t}$, considerando infine $\operatorname{Ker} \underline{t}^*$),
- un iperpiano Π, rappresentato da un'equazione cartesiana $\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$, è in $\mathbb{K}^n$ esattamente il sottospazio ortogonale a $(\alpha_1, \ldots, \alpha_n)^\top$ tramite il prodotto scalare standard,
- in generale, un sistema di equazioni omogenee è l'intersezione di più sottospazi ortogonali,
- se $\mathbb{F}$ è un'estensione di campo di $\mathbb{K}$, allora vale $[V : \mathbb{K}] = [V : \mathbb{F}][\mathbb{F} : \mathbb{K}]$ (*teorema delle torri algebriche*).

Applicazioni lineari, somme dirette, quozienti e prodotti diretti

Un'applicazione da V in W si dice applicazione lineare se:

- $f(\underline{v} + \underline{w}) = f(\underline{v}) + f(\underline{w})$,
- $f(\alpha \underline{v}) = \alpha f(\underline{v})$.

Si definisce $\mathcal{L}(V, W) \subseteq W^V$ come lo spazio delle applicazioni lineari da V a W. Si definisce $\operatorname{End}(V)$ come lo spazio degli endomorfismi di V, ossia delle applicazioni lineari da V in V, dette anche operatori. Un'applicazione lineare si dice isomorfismo se è bigettiva. La composizione di funzioni è associativa.

Dato un sottospazio A di V, si definisce lo spazio quoziente V/A come l'insieme quoziente $V/\!\sim$ della relazione di equivalenza $\underline{a} \sim \underline{b} \iff a - b \in A$ dotato dell'usuale somma e prodotto esterno. Si scrive $[\underline{v}]_A$ come $\underline{v} + A$ e vale che $A = \underline{0} + A$. In particolare $\underline{v} + A = A \iff \underline{v} \in A$.

Siano $f : V \to W$, $h : V \to W$, $g : W \to Z$ tre applicazioni lineari. $\mathcal{B}_V$ e $\mathcal{B}_W$ sono due basi rispettivamente di V e W. In particolare sia $\mathcal{B}_V = \{\underline{v_1}, \ldots, \underline{v_n}\}$. Si ricorda che $\operatorname{rg}(f) = \dim \operatorname{Im} f$. Siano e ed e' le basi canoniche rispettivamente di $\mathbb{K}^n$ e $\mathbb{K}^m$.

- $f(\underline{0}_V) = \underline{0}_W$,
- $\operatorname{Ker} f = f^{-1}(\underline{0}_W)$ è un sottospazio di V,
- $\operatorname{Im} f = f(V)$ è un sottospazio di W,
- $\operatorname{Im} f = \operatorname{Span}(f(\underline{v_1}), \ldots, f(\underline{v_n}))$,
- f è iniettiva $\iff \operatorname{Ker} f = \{\underline{0}\}$,
- $V/\operatorname{Ker} f \cong \operatorname{Im} f$ (*primo teorema d'isomorfismo*),
- $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$ (*teorema del rango*, o formula delle dimensioni, valido se la dimensione di V è finita),
- $g \circ f$ è un'applicazione lineare da V in Z,
- la composizione di funzioni è associativa e distributiva da ambo i lati,
- $g \circ (\alpha f) = \alpha(g \circ f) = (\alpha g) \circ f$, se $\alpha \in \mathbb{K}$,
- $\operatorname{Ker} f \subseteq \operatorname{Ker}(g \circ f)$,
- $\operatorname{Im}(g \circ f) \subseteq \operatorname{Im} g$,
- $\dim \operatorname{Im}(g \circ f) = \dim \operatorname{Im} g|_{\operatorname{Im} f} = \dim \operatorname{Im} f - \dim \operatorname{Ker} g|_{\operatorname{Im} f} = \dim \operatorname{Im} f - \dim(\operatorname{Ker} g \cap \operatorname{Im} f)$ (è sufficiente applicare la formula delle dimensioni sulla composizione),
- $\dim \operatorname{Im}(g \circ f) \leq \min\{\dim \operatorname{Im} g, \dim \operatorname{Im} f\}$,
- $\dim \operatorname{Ker}(g \circ f) \leq \dim \operatorname{Ker} g + \dim \operatorname{Ker} f$ (è sufficiente applicare la formula delle dimensioni su $(g \circ f)|_{\operatorname{Ker}(g \circ f)}$),
- f iniettiva $\implies \dim V \leq \dim W$,
- f surgettiva $\implies \dim V \geq \dim W$,
- f isomorfismo $\implies \dim V = \dim W$,
- $g \circ f$ iniettiva $\implies f$ iniettiva,
- $g \circ f$ surgettiva $\implies g$ surgettiva,
- f surgettiva $\implies \operatorname{rg}(g \circ f) = \operatorname{rg}(g)$,
- g iniettiva $\implies \operatorname{rg}(g \circ f) = \operatorname{rg}(f)$,
- $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) = \left([f(\underline{v_1})]_{\mathcal{B}_W} \mid \cdots \mid [f(\underline{v_n})]_{\mathcal{B}_W} \right)$ è la matrice associata a f sulle basi $\mathcal{B}_V$, $\mathcal{B}_W$,
- $M_W^V(f + h) = M_W^V(f) + M_W^V(h)$,
- $M_Z^V(g \circ f) = M_Z^W(g) M_W^V(f)$,
- data $A \in M(m, n, \mathbb{K})$, sia $f_A : \mathbb{K}^n \to \mathbb{K}^m$ tale che $f_A(\underline{x}) = A\underline{x}$, allora $M_{e'}^{e}(f_A) = A$,
- f è completamente determinata dai suoi valori in una qualsiasi base di V ($M_{\mathcal{B}_W}^{\mathcal{B}_V}$ è un isomorfismo tra $\mathcal{L}(V, W)$ e $M(\dim W, \dim V, \mathbb{K})$),
- $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$ (dall'isomorfismo di sopra),
- $[\,]_{\mathcal{B}_W}^{-1} \circ M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \circ [\,]_{\mathcal{B}_V} = f$,
- $[f(\underline{v})]_{\mathcal{B}_W} = M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \cdot [\underline{v}]_{\mathcal{B}_V}$,
- $\operatorname{Im}(f) = [\,]_{\mathcal{B}_W}^{-1} \left(\operatorname{Im} M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \right)$
- $\operatorname{rg}(f) = \operatorname{rg}\left(M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \right)$,
- $\operatorname{Ker}(f) = [\,]_{\mathcal{B}_V}^{-1} \left(\operatorname{Ker} M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \right)$,
- $\dim \operatorname{Ker}(f) = \dim \operatorname{Ker} M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)$.

Siano $\mathcal{B}'_V$, $\mathcal{B}'_W$ altre due basi rispettivamente di V e W. Allora vale il *teorema del cambiamento di base*:

$$M_{\mathcal{B}'_W}^{\mathcal{B}'_V}(f) = M_{\mathcal{B}'_W}^{\mathcal{B}_W}(id_W)\, M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)\, M_{\mathcal{B}_V}^{\mathcal{B}'_V}(id_V).$$

Siano A e B due sottospazi di V. $\mathcal{B}_A$ e $\mathcal{B}_B$ sono due basi rispettivamente di A e B.

- $A + B = \{\underline{a} + \underline{b} \in V \mid \underline{a} \in A, \underline{b} \in B\}$ è un sottospazio,
- $\dim(A + B) = \dim A + \dim B - \dim(A \cap B)$ (*formula di Grassmann*),
- A e B sono in somma diretta $\iff A \cap B = \{\underline{0}\} \iff$ ogni elemento di $A + B$ si scrive in modo unico come somma di $\underline{a} \in A$ e $\underline{b} \in B$ $\iff \dim(A + B) = \dim A + \dim B$ (in tal caso si scrive $A + B = A \oplus B$),
- $\dim V/A = \dim V - \dim A$ (è sufficiente applicare il teorema del rango alla proiezione al quoziente),
- $\dim V \times W = \dim V + \dim W$ ($\mathcal{B}_V \times \{\underline{0}_W\} \cup \{\underline{0}_V\} \times \mathcal{B}_W$ è una base di $V \times W$).

Si definisce *immersione* da V in $V \times W$ l'applicazione lineare i_V tale che $i_V(\underline{v}) = (\underline{v}, \underline{0})$. Si definisce *proiezione* da $V \times W$ in V l'applicazione lineare p_V tale che $p_V(\underline{v}, \underline{w}) = \underline{v}$. Analogamente si può fare con gli altri spazi del prodotto cartesiano.

Si dice che B è un supplementare di A se $V = A \oplus B$ (ossia $\iff \dim A + \dim B = \dim V \wedge A \cap B = \{\underline{0}\}$). Il supplementare non è per forza unico. Per trovare un supplementare di A è sufficiente completare $\mathcal{B}_A$ ad una base $\mathcal{B}$ di V e considerare $B := \operatorname{Span}(\mathcal{B} \setminus \mathcal{B}_A)$.

Somma diretta di più sottospazi

Si dice che i sottospazi W_1, ..., W_k di V sono in somma diretta, e si scrive $W_1 + \ldots + W_k = W_1 \oplus \ldots \oplus W_k$, se la rappresentazione di un vettore della somma di questi sottospazi è unica, ossia se esistono unici $\underline{w_1} \in W_1$, ..., $\underline{w_k} \in W_k$ tali per cui $\underline{w} \in W_1 + \ldots + W_k$ si scrive come $\underline{w} = \underline{w_1} + \ldots + \underline{w_k}$. In generale, sono equivalenti i seguenti fatti:

(i) W_1, ..., W_k sono in somma diretta,

(ii) Se esistono $\underline{w_1} \in W_1$, ..., $\underline{w_k} \in W_k$ tali per cui $\underline{w_1} + \ldots + \underline{w_k} = \underline{0}$, allora $\underline{w_1} = \cdots = \underline{w_k} = \underline{0}$ (è sufficiente considerare due scritture alternative e poi farne la differenza per dimostrare un'implicazione),

(iii) Se $\mathcal{B}_{W_1}$, ..., $\mathcal{B}_{W_k}$ sono basi di W_1, ..., W_k, allora $\bigcup_{i=1}^{k} \mathcal{B}_{W_i}$ è base di $W_1 + \ldots + W_k$ (è sufficiente considerare l'indipendenza lineare per dimostrare un'implicazione),

(iv) $\dim(W_1 + \ldots + W_k) = \dim W_1 + \ldots + \dim W_k$ (si dimostra facilmente che è equivalente a (iii), e quindi che lo è alle altre proposizioni),

(v) $W_i \cap (W_1 + \ldots + W_{i-1}) = \{\underline{0}\}$ $\forall\, 2 \leq i \leq k$ (è sufficiente spezzare la somma in $(W_1 + \ldots + W_{i-1}) + W_i$ e ricondursi al caso di due sottospazi, mostrando in particolare, per induzione, l'equivalenza con (iv), da cui seguono le altre equivalenze),

(vi) $W_i \cap (W_1 + \ldots + W_{i-1} + \widehat{W_i} + W_{i+1} + W_k) = \{\underline{0}\}$ $\forall\, 1 \leq i \leq k$, ossia W_i, intersecato con la somma dei restanti sottospazi, è di dimensione nulla (è facile ricondursi alla proposizione (v) per induzione).

Proprietà generali delle matrici

Si dice che una matrice $A \in M(n, \mathbb{K})$ è singolare se $\det(A) = 0$, o equivalentemente se non è invertibile. Compatibilmente, si dice che una matrice $A \in M(n, \mathbb{K})$ è non singolare se $\det(A) \neq 0$, ossia se A è invertibile.

Si definisce la matrice trasposta di $A \in M(m, n, \mathbb{K})$, detta $A^\top$, in modo tale che $A_{ij} = A^\top_{ji}$.

- $(AB)^\top = B^\top A^\top$,
- $(A + B)^\top = A^\top + B^\top$,
- $(\lambda A)^\top = \lambda A^\top$,
- $(A^\top)^\top = A$,
- se A è invertibile, $(A^\top)^{-1} = (A^{-1})^\top$,
- $\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\left(\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right) = \left(\begin{array}{c|c} AE+BG & AF+BH \\ \hline CE+DG & CF+DH \end{array}\right)$.

Siano $A \in M(m, n, \mathbb{K})$ e $B \in M(n, m, \mathbb{K})$.

Si definisce $\mathrm{GL}(n, \mathbb{K})$ come il gruppo delle matrici di taglia n invertibili sulla moltiplicazione matriciale. Si definisce triangolare superiore una matrice i cui elementi al di sotto della diagonale sono nulli, mentre si definisce triangolare inferiore una matrice i cui elementi nulli sono quelli al di sopra della diagonale.

Si definiscono

$$Z(M(n, \mathbb{K})) = \{A \in M(n, \mathbb{K}) \mid AB = BA \ \forall B \in M(n, \mathbb{K})\},$$

ossia l'insieme delle matrici che commutano con tutte le altre matrici, e

$$Z_{\mathrm{GL}}(M(n, \mathbb{K})) = \{A \in M(n, \mathbb{K}) \mid AB = BA \ \forall B \in \mathrm{GL}(n, \mathbb{K})\},$$

ovvero l'insieme delle matrici che commutano con tutte le matrici di $\mathrm{GL}(n, \mathbb{K})$.

Si definisce $\mathrm{tr} \in M(m, \mathbb{K})^*$ come il funzionale che associa ad ogni matrice la somma degli elementi sulla sua diagonale.

- $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$,
- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$,
- $Z(M(n, \mathbb{K})) = \mathrm{Span}(I_n)$,
- $Z_{\mathrm{GL}}(M(n, \mathbb{K})) = \mathrm{Span}(I_n)$.

Sia $A \in M(n, \mathbb{K})$. Sia $C_A \in \mathrm{End}(M(n, \mathbb{K}))$ definito in modo tale che $C_A(B) = AB - BA$. Allora $\mathrm{Ker}\, C_A = M(n, \mathbb{K}) \iff A \in \mathrm{Span}(I_n)$. Siano I un insieme di n^2 indici distinti, allora l'insieme

$$T = \left\{A^i \mid i \in I\right\}$$

è linearmente dipendente (è sufficiente notare che se così non fosse, se $A \notin \mathrm{Span}(I_n)$, tale T sarebbe base di $M(n, \mathbb{K})$, ma così $\mathrm{Ker}\, C_A = M(n, \mathbb{K}) \implies A \in \mathrm{Span}(I_n)$, $\lightning$, e che se $A \in \mathrm{Span}(I_n)$, T è chiaramente linearmente dipendente).

In generale esiste sempre un polinomio $p(X) \in \mathbb{K}[x]$ di grado n tale per cui $p(A) = 0$, dove un tale polinomio è per esempio il polinomio caratteristico di p, ossia $p(\lambda) = \det(\lambda I_n - A)$ (*teorema di Hamilton-Cayley*).

Si elencano adesso i tipi principali di matrici:

- A è simmetrica $\overset{\text{def}}{\iff} A^\top = A$,
- A è antisimmetrica $\overset{\text{def}}{\iff} A^\top = -A$,
- A è hermitiana $\overset{\text{def}}{\iff} A^* := \left(\overline{A}\right)^\top = A$,
- A è ortogonale ($A \in O(n)$ o O_n) $\overset{\text{def}}{\iff} AA^\top = A^\top A = I_n$,
- A è unitaria ($A \in U(n)$ o U_n) $\overset{\text{def}}{\iff} AA^* = A^*A = I_n$.

Se $A \in M(m, n, \mathbb{R})$, allora $\mathrm{Ker}\, A^\top A = \mathrm{Ker}\, A$. Infatti, se $\underline{x} \in \mathrm{Ker}\, A^\top A$, allora $A^\top A\underline{x} = \underline{0} \implies \underline{x}^\top A^\top A\underline{x} = \underline{0} \implies q(A\underline{x}) = \underline{0} \implies A\underline{x} = \underline{0} \implies \underline{x} \in \mathrm{Ker}\, A$, dove q è la forma quadratica derivante dal prodotto scalare standard di $\mathbb{R}^n$. Da questo risultato si deduce anche che $\mathrm{rg}(A^\top A) = \mathrm{rg}(A)$.

Se $A \in M(m, n, \mathbb{C})$, allora $\mathrm{Ker}\, A^*A = \mathrm{Ker}\, A$ e $\mathrm{rg}(A^*A) = \mathrm{rg}(A)$ (si segue la stessa linea di dimostrazione di sopra).

Rango di una matrice

Si definisce rango di una matrice A il numero di colonne linearmente indipendenti di A. Siano A, $B \in M(m, n, \mathbb{K})$.

- $\mathrm{rg}(A) = \mathrm{rg}(A^\top)$ (i.e. il rango è lo stesso se calcolato sulle righe invece che sulle colonne),
- $\mathrm{rg}(A) \leq \min\{m, n\}$ (come conseguenza dell'affermazione precedente),
- $\mathrm{rg}(A + B) \leq \mathrm{rg}(A) + \mathrm{rg}(B) \impliedby \mathrm{Im}(A + B) \subseteq \mathrm{Im}(A) + \mathrm{Im}(B)$,
- $\mathrm{rg}(A + B) = \mathrm{rg}(A) + \mathrm{rg}(B) \implies \mathrm{Im}(A + B) = \mathrm{Im}(A) \oplus \mathrm{Im}(B)$ (è sufficiente applicare la formula di Grassmann),
- $\mathrm{rg}(A)$ è il minimo numero di matrici di rango uno che sommate restituiscono A (è sufficiente usare la proposizione precedente per dimostrare che devono essere almeno $\mathrm{rg}(A)$),
- $\mathrm{rg}(A) = 1 \implies \exists\, B \in M(m, 1, \mathbb{K})$, $C \in M(1, n, \mathbb{K}) \mid A = BC$ (infatti A può scriversi come $\begin{pmatrix} \alpha_1 A^i & \cdots & \alpha_n A^i \end{pmatrix}$ per un certo $i \leq n$ tale che $A^i \neq \underline{0}$).

Siano $A \in M(m, n, \mathbb{K})$, $B \in M(n, k, \mathbb{K})$ e $C \in M(k, t, \mathbb{K})$.

- $\mathrm{rg}(AB) \geq \mathrm{rg}(A) + \mathrm{rg}(B) - n$ (*disuguaglianza di Sylvester* – è sufficiente usare la formula delle dimensioni ristretta alla composizione $f_A \circ f_B$),
- $\mathrm{rg}(ABC) \geq \mathrm{rg}(AB) + \mathrm{rg}(BC) - \mathrm{rg}(B)$ (*disuguaglianza di Frobenius*, di cui la proposizione precedente è un caso particolare con $B = I_n$ e $k = n$),
- $\mathrm{rg}(AB) = \mathrm{rg}(B) \impliedby \mathrm{Ker}\, A = \{\underline{0}\}$ (è sufficiente usare la formula delle dimensioni ristretta alla composizione $f_A \circ f_B$),
- $\mathrm{rg}(AB) = \mathrm{rg}(A) \impliedby f_B$ surgettiva (come sopra).

Sia $A \in M(n, \mathbb{K})$.

- se A è antisimmetrica e il campo su cui si fonda lo spazio vettoriale non ha caratteristica 2, allora $\mathrm{rg}(A)$ è pari,
- $\mathrm{rg}(A) = n \iff \dim \mathrm{Ker}\, A = 0 \iff \det(A) \neq 0 \iff A$ è invertibile,

Sistemi lineari, algoritmo di eliminazione di Gauss ed SD-equivalenza

Un sistema lineare di m equazioni in n variabili può essere rappresentato nella forma $A\underline{x} = B$, dove $A \in M(m, n, \mathbb{K})$, $\underline{x} \in \mathbb{K}^n$ e $B \in \mathbb{K}^m$. Un sistema lineare si dice omogeneo se $B = \underline{0}$. In tal caso l'insieme delle soluzioni del sistema coincide con $\mathrm{Ker}\, A = \mathrm{Ker}\, f_A$, dove $f_A : \mathbb{K}^n \to \mathbb{K}^m$ è l'applicazione lineare indotta dalla matrice A. Le soluzioni di un sistema lineare sono raccolte nel sottospazio affine $\underline{s} + \mathrm{Ker}\, A$, dove $\underline{s}$ è una qualsiasi soluzione del sistema completo.

- $A\underline{x} = B$ ammette soluzione se e solo se $B \in \mathrm{Span}(A^1, \ldots, A^n) \iff \mathrm{Span}(A^1, \ldots, A^n, B) = \mathrm{Span}(A^1, \ldots, A^n) \iff \dim \mathrm{Span}(A^1, \ldots, A^n, B) = \dim \mathrm{Span}(A^1, \ldots, A^n) \iff \dim \mathrm{Im}(A \mid B) = \dim \mathrm{Im}\, A \iff \mathrm{rg}(A \mid B) = \mathrm{rg}(A)$ (*teorema di Rouché-Capelli*),
- Data un'applicazione lineare $f : \mathbb{K}^n \to \mathbb{K}^m$ determinata dalla matrice M, il sistema di equazioni cartesiane che rappresenta $\mathrm{Im}\, f$ si ottiene imponendo la validità del teorema di Rouché-Capelli sul vettore $\underline{x} = (x_1, \ldots, x_m)$, ossia imponendo $\mathrm{rg}(M) = \mathrm{rg}(M \mid \underline{x})$,
- $A\underline{x} = B$, se la ammette, ha un'unica soluzione se e solo se $\mathrm{Ker}\, A = \{\underline{0}\} \iff \mathrm{rg}\, A = n$.

Si definiscono tre operazioni sulle righe di una matrice A:

1. l'operazione di scambio di riga,
2. l'operazione di moltiplicazione di una riga per uno scalare non nullo,
3. la somma di un multiplo non nullo di una riga ad un'altra riga distinta.

A queste operazioni è associato il prodotto a sinistra per delle particolari matrici. In particolare, l'operazione di scambio della riga i-esima con quella j-esima corrisponde alla moltiplicazione a sinistra per la matrice $S_{i,j}$, detta matrice elementare di permutazione, dove:

$$S_{i,j} = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}.$$

In particolare, la stessa matrice $S_{i,j}$ si ottiene scambiando la riga i-esima e la j-esima. Per esempio, scambiare due righe in una matrice 2×2 corrisponde a moltiplicare a sinistra per $S_{1,2}$, dove:

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

All'operazione di moltiplicazione della riga i-esima per uno scalare $\lambda \neq 0$ corrisponde invece la matrice $M_{i,\lambda}$, detta elementare di dilatazione, dove:

$$M_{i,\lambda} = \left(\begin{array}{c|c|c} I_{i-1} & 0 & 0 \\ \hline 0 & \lambda & 0 \\ \hline 0 & 0 & I_{n-i} \end{array}\right).$$

All'operazione di somma della riga j-esima moltiplicata per $\lambda \neq 0$ alla riga i-esima corrisponde invece la matrice $M_{i,j,\lambda}$, detta elementare di trasvezione (o di tosatura, dall'inglese *shear matrix*), dove:

$$M_{i,j,\lambda} = I_n + \lambda E_{i,j}.$$

Se $\lambda \neq 0$, tutte queste matrici sono invertibili ed in particolare valgono le seguente relazioni:

- $S_{i,j}^{-1} = S_{i,j}$, da cui si osserva che l'inversa di una matrice elementare di permutazione è ancora una matrice dello stesso tipo),
- $M_{i,\lambda}^{-1} = M_{i,\frac{1}{\lambda}}$, come sopra,
- $M_{i,j,\lambda}^{-1} = M_{i,j,-\lambda}$, come sopra,
- $M_{i,\lambda} = M_{i,i,\lambda-1}$,
- le matrici elementari generano il gruppo delle matrici invertibili $\mathrm{GL}(n,\mathbb{K})$, ossia ogni matrice invertibile si scrive come prodotto di matrici elementari (è sufficiente applicare l'algoritmo di eliminazione di Gauss per righe su $A \in \mathrm{GL}(n,\mathbb{K})$ e osservare che ne deve risultare obbligatoriamente una matrice diagonale che, normalizzata sugli elementi, restituisce esattamente I_n; allora poiché applicare l'algoritmo equivale a moltiplicare a sinistra per delle matrici elementari $\mathcal{E}_1$, ..., $\mathcal{E}_k$, si verifica che $\mathcal{E}_1 \cdots \mathcal{E}_k A = I_n \implies A = \mathcal{E}_k^{-1} \cdots \mathcal{E}_1^{-1}$, dove si conclude la dimostrazione osservando che l'inversa di una matrice elementare è ancora una matrice elementare).

Queste operazioni non variano né $\ker A$ né $\mathrm{rg}(A)$.
Permettendo di variare $\ker A$ si possono effettuare le stesse medesime operazioni sulle colonne (lasciando però invariato $\mathrm{Im}\, A$, e quindi $\mathrm{rg}(A)$): tali operazioni corrispondono a moltiplicare a destra per una matrice invertibile, analogamente a come accade per le righe.
Le matrici per cui si moltiplica a destra per operare sulle colonne sono esattamente le stesse matrici impiegate per le operazioni di riga, sebbene trasposte (e quindi sono ancora matrici elementari). In particolare le matrici elementari di permutazione (per scambiare le righe) e di dilatazione (per moltiplicare una riga per uno scalare non nullo) coincidono. Pertanto, se A è una matrice simmetrica (i.e. se $A \in \mathrm{Sym}(n,\mathbb{K})$), operare mediante le stesse operazioni sulle righe e sulle colonne permette di individuare matrici congruenti ad A.

L'algoritmo di eliminazione di Gauss procede nel seguente modo:

1. se A ha una riga, l'algoritmo termina;
2. altrimenti si prenda la prima riga di A con il primo elemento non nullo e la si scambi con la prima riga di A (in caso non esista, si proceda all'ultimo passo),
3. per ogni riga di A con primo elemento non nullo, esclusa la prima, si sottragga un multiplo della prima riga in modo tale che la riga risultante abbia il primo elemento nullo,
4. si ripeta l'algoritmo considerando come matrice A la matrice risultante dall'algoritmo senza la prima riga e la prima colonna (in caso tale matrice non possa esistere, l'algoritmo termina).

Si definiscono *pivot* di una matrice l'insieme dei primi elementi non nulli di ogni riga della matrice. Il rango della matrice iniziale A è pari al numero di *pivot* della matrice risultante dall'algoritmo di eliminazione di Gauss. Una matrice che processata dall'algoritmo di eliminazione di Gauss restituisce sé stessa è detta matrice a scala.

Agendo solo attraverso operazioni per riga, l'algoritmo di eliminazione di Gauss non modifica $\ker A$ (si può tuttavia integrare l'algoritmo con le operazioni per colonna, perdendo quest'ultimo beneficio).

Agendo su una matrice a scala con operazioni per riga considerando la matrice riflessa (ossia dove l'elemento $(1,1)$ e (m,n) sono scambiati), si può ottenere una matrice a scala ridotta, ossia un matrice dove tutti i pivot sono 1 e dove tutti gli elementi sulle colonne dei pivot, eccetto i pivot stessi, sono nulli.

Si definisce:

$$I_r^{m\times n} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M(m,n,\mathbb{K}).$$

Per ogni applicazione lineare $f : V \to W$, con $\dim V = n$ e $\dim W = m$ esistono due basi $\mathcal{B}_V$, $\mathcal{B}_W$ rispettivamente di V e W tale che $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) = I_r^{m\times n}$, dove $r = \mathrm{rg}(f)$ (è sufficiente completare con I a base di V una base di $\ker f$ e poi prendere come base di W il completamento di $f(I)$ su una base di W).

Si definisce SD-equivalenza la relazione d'equivalenza su $M(m,n,\mathbb{K})$ indotta dalla relazione
$A \sim_{SD} B \iff \exists\, P \in \mathrm{GL}(m,\mathbb{K}),\, Q \in \mathrm{GL}(n,\mathbb{K}) \mid A = PBQ$.
L'invariante completo della SD-equivalenza è il rango:
$\mathrm{rg}(A) = \mathrm{rg}(B) \iff A \sim_{SD} B$ (infatti
$\mathrm{rg}(A) = r \iff A \sim_{SD} I_r^{m\times n}$ – è sufficiente applicare il cambio di base e sfruttare il fatto che esistono sicuramente due basi per cui f_A ha $I_r^{m\times n}$ come matrice associata).

Poiché $I_r^{m\times n}$ ha sempre rango r, l'insieme quoziente della SD-equivalenza su $M(m,n,\mathbb{K})$ è il seguente:

$$M(m,n,\mathbb{K})/\!\sim_{SD} = \left\{ \underline{[0]}, \left[I_1^{m\times n}\right], \ldots, \left[I_{\min\{m,n\}}^{m\times n}\right] \right\},$$

contenente esattamente $\min\{m,n\}$ elementi. L'unico elemento di $\underline{[0]}$ è $\underline{0}$ stesso.

La regola di Cramer

Qualora $m = n$ e A fosse invertibile (i.e. $\det(A) \neq 0$), per calcolare il valore di $\underline{x}$ si può applicare la regola di Cramer.
Si definisce:

$$A_i^* = \begin{pmatrix} A^1 & \cdots & A^i \to B & \cdots & A^n \end{pmatrix},$$

dove si sostituisce alla i-esima colonna di A il vettore B.
Allora vale la seguente relazione:

$$\underline{x} = \frac{1}{\det(A)} \begin{pmatrix} \det(A_1^*) \\ \vdots \\ \det(A_n^*) \end{pmatrix}.$$

L'inverso (generalizzato e non) di una matrice

Si definisce matrice dei cofattori di una matrice $A \in M(n,\mathbb{K})$ la seguente matrice:

$$\mathrm{Cof}\, A = \begin{pmatrix} \mathrm{Cof}_{1,1}(A) & \ldots & \mathrm{Cof}_{1,n}(A) \\ \vdots & \ddots & \vdots \\ \mathrm{Cof}_{n,1}(A) & \ldots & \mathrm{Cof}_{n,n}(A), \end{pmatrix},$$

dove, detta $A_{i,j}$ il minore di A ottenuto eliminando la i-esima riga e la j-esima colonna, si definisce il cofattore (o complemento algebrico) nel seguente modo:

$$\mathrm{Cof}_{i,j}(A) = (-1)^{i+j} \det(A_{i,j}).$$

Si definisce inoltre l'aggiunta classica:

$$\mathrm{adj}(A) = (\mathrm{Cof}\, A)^\top.$$

Allora, se A ammette un inverso (i.e. se $\det(A) \neq 0$), vale la seguente relazione:

$$A^{-1} = \frac{1}{\det(A)} \mathrm{adj}(A).$$

Quindi, per esempio, A^{-1} è a coefficienti interi $\iff \det(A) = \pm 1$.

Siano A, $B \in M(n,\mathbb{K})$.

- $\mathrm{adj}(AB) = \mathrm{adj}(B)\,\mathrm{adj}(A)$,
- $\mathrm{adj}(A^\top) = \mathrm{adj}(A)^\top$.

Si definisce inverso generalizzato di una matrice $A \in M(m,n,\mathbb{K})$ una matrice $X \in M(n,m,\mathbb{K}) \mid AXA = A$.
Ogni matrice ammette un inverso generalizzato (è sufficiente considerare gli inversi generalizzati di $I_r^{m\times n}$ e la SD-equivalenza di A con $I_r^{m\times n}$, dove $\mathrm{rg}(A) = r$). Se $m = n$ ed A è invertibile, allora A^{-1} è l'unico inverso generalizzato di A.
Gli inversi generalizzati di $I_r^{m\times n}$ sono della forma:

$$X = \left(\begin{array}{c|c} I_r & B \\ \hline C & D \end{array}\right) \in M(m,n,\mathbb{K}).$$

Endomorfismi e similitudine

Si definisce la similitudine tra matrici su $M(n,\mathbb{K})$ come la relazione di equivalenza determinata da
$A \sim B \iff \exists\, P \in \mathrm{GL}(n,\mathbb{K}) \mid A = PBP^{-1}$.
$A \sim B \implies \mathrm{rg}(A) = \mathrm{rg}(B)$, $\mathrm{tr}(A) = \mathrm{tr}(B)$, $\det(A) = \det(B)$, $P_\lambda(A) = P_\lambda(B)$ (invarianti *non completi* della similitudine). Vale inoltre che $A \sim B \iff A$ e B hanno la stessa forma canonica di Jordan, a meno di permutazioni dei blocchi di Jordan (invariante *completo* della similitudine). La matrice identità è l'unica matrice identica a sé stessa.

Sia $p \in \mathrm{End}(V)$. Si dice che un endomorfismo è un automorfismo se è un isomorfismo. Gli automorfismi formano un sottospazio vettoriale di $\mathrm{End}(V)$ denotato con $\mathrm{Aut}(V)$ o $\mathrm{GL}(V)$. Siano $\mathcal{B}$, $\mathcal{B}'$ due qualsiasi basi di V.

- p automorfismo $\iff$ p iniettivo $\iff$ p surgettivo (è sufficiente applicare la formula delle dimensioni),
- $M^{\mathcal{B}}_{\mathcal{B}'}(id_V) M^{\mathcal{B}'}_{\mathcal{B}}(id_V) = I_n$ (dunque entrambe le matrici sono invertibili e sono l'una l'inverso dell'altra),
- se p è un automorfismo, $M^{\mathcal{B}}_{\mathcal{B}'}(p^{-1}) = M^{\mathcal{B}'}_{\mathcal{B}}(p)^{-1}$,
- $M^{\mathcal{B}}_{\mathcal{B}}(p) = \underbrace{M^{\mathcal{B}'}_{\mathcal{B}}(id_V)}_{P}\, M^{\mathcal{B}'}_{\mathcal{B}'}(p)\, \underbrace{M^{\mathcal{B}}_{\mathcal{B}'}(id_V)}_{P^{-1}}$ (ossia $M^{\mathcal{B}}_{\mathcal{B}}(p) \sim M^{\mathcal{B}'}_{\mathcal{B}'}(p)$).

$M^{\mathcal{B}}_{\mathcal{B}'}(id_V) M^{\mathcal{B}'}_{\mathcal{B}}(id_V) = I_n$. Dunque entrambe le matrici sono invertibili. Inoltre $M^{\mathcal{B}}_{\mathcal{B}}(id_V) = I_n$.

Si definisce un analogo della similitudine anche per gli endomorfismi: due endomorfismi f, $g \in \mathrm{End}(V)$ si dicono coniugati se e solo se $\exists\, h \in \mathrm{GL}(V) \mid f = h\, g\, h^{-1}$. Il coniugio induce in particolare un'altra relazione di equivalenza. Due endomorfismi f e g sono coniugati se e solo se le loro matrici associate nella stessa base $\mathcal{B}$ sono simili.

Duale, biduale e annullatore

Si definisce duale di uno spazio vettoriale V lo spazio $V^* = \mathcal{L}(V,\mathbb{K})$, i cui elementi sono detti funzionali. Analogamente il biduale è il duale del duale di V: $V^{**} = (V^*)^* = \mathcal{L}(V^*,\mathbb{K})$.

Sia data una base $\mathcal{B} = \{\underline{v_1},\dots,\underline{v_n}\}$ di uno spazio vettoriale V di dimensione n. Allora
$\dim V^* = \dim \mathcal{L}(V,\mathbb{K}) = \dim V \cdot \dim \mathbb{K} = \dim V$. Si definisce il funzionale $\underline{v_i}^*$ come l'applicazione lineare univocamente determinata dalla relazione:

$$\underline{v_i}^*(\underline{v_j}) = \delta_{ij}.$$

Sia $\mathcal{B}^* = \{\underline{v_1}^*,\dots,\underline{v_n}^*\}$. Allora $\mathcal{B}^*$ è una base di V^*. Poiché V e V^* hanno la stesso dimensione, tali spazi sono isomorfi, sebbene non canonicamente. Ciononostante, V e V^{**} sono canonicamente isomorfi tramite l'isomorfismo:

$$\varphi^{**} : V \to V^{**},\ \underline{v} \mapsto \mathrm{val}\,|_{V^*},$$

che associa ad ogni vettore $\underline{v}$ la funzione di valutazione in una funzionale in $\underline{v}$, ossia:

$$\mathrm{val}\,|_{V^*} : V^* \to \mathbb{K},\ f \mapsto f(\underline{v}).$$

Sia $U \subseteq V$ un sottospazio di V. Si definisce il sottospazio di $\mathcal{L}(V,W)$:

$$\mathrm{Ann}_{\mathcal{L}(V,W)}(U) = \{f \in \mathcal{L}(V,W) \mid f(U) = \{\underline{0}\}\}.$$

Se V è a dimensione finita, la dimensione di $\mathrm{Ann}_{\mathcal{L}(V,W)}(U)$ è pari a $(\dim V - \dim U) \cdot \dim W$ (è sufficiente prendere una base di U, completarla a base di V e notare che $f(U) = \{\underline{0}\} \iff$ ogni valutazione in f degli elementi della base di U è nullo $\iff$ la matrice associata di f ha tutte colonne nulle in corrispondenza degli elementi della base di U).

Si scrive semplicemente $\mathrm{Ann}(U)$ quando $W = \mathbb{K}$ (ossia quando le funzioni sono funzionali di V). In tal caso $\dim \mathrm{Ann}(U) = \dim V - \dim U$.

- $\varphi^{**}(U) \subseteq \mathrm{Ann}(\mathrm{Ann}(U))$,
- se V è a dimensione finita, $\varphi^{**}(U) = U^{**} = \mathrm{Ann}(\mathrm{Ann}(U))$ (è sufficiente applicare la formula delle dimensioni $\varphi^{**}|_U$ e notare l'uguaglianza tra le due dimensioni),
- se V è a dimensione finita e W è un altro sottospazio di V, $U = W \iff \mathrm{Ann}(U) = \mathrm{Ann}(W)$ (è sufficiente considerare $\mathrm{Ann}(\mathrm{Ann}(U)) = \mathrm{Ann}(\mathrm{Ann}(W))$ e applicare la proposizione precedente, ricordandosi che φ^{**} è un isomorfismo, ed è dunque iniettivo).

Si definisce l'applicazione trasposta $^\top$ da $\mathcal{L}(V,W)$ a $\mathcal{L}(W^*,V^*)$ in modo tale che $f^\top(g) = g \circ f \in V^*$. Siano f, $g \in \mathcal{L}(V,W)$ e sia $h \in \mathcal{L}(W,Z)$.

- $(f+g)^\top = f^\top + g^\top$,
- $(\lambda f)^\top = \lambda f^\top$,
- se f è invertibile, $(f^{-1})^\top = (f^\top)^{-1}$,
- $(h \circ f)^\top = f^\top \circ h^\top$.

Siano $\mathcal{B}_V$, $\mathcal{B}_W$ due basi rispettivamente di V e di W. Allora vale la seguente relazione:

$$M^{\mathcal{B}_W^*}_{\mathcal{B}_V^*}(f^\top) = M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)^\top.$$

Applicazioni multilineari

Sia $f : V_1 \times \ldots \times V_n \to W$ un'applicazione, dove V_i è uno spazio vettoriale $\forall\, i \le n$, così come W. Tale applicazione si dice n-lineare ed appartiene allo spazio $\mathrm{Mult}(V_1 \times \ldots \times V_n, W)$, se è lineare in ogni sua coordinata, ossia se:

- $f(x_1,\dots,x_i+y_i,\dots,x_n) = f(x_1,\dots,x_i,\dots,x_n) + f(x_1,\dots,y_i,\dots,x_n)$,
- $f(x_1,\dots,\alpha x_i,\dots,x_n) = \alpha f(x_1,\dots,x_i,\dots,x_n)$.

Sia $W = \mathbb{K}$, e siano tutti gli spazi V_i fondati su tale campo: allora $\mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K})$ si scrive anche come $V_1^* \otimes \ldots \otimes V_n^*$, e tale spazio è detto prodotto tensoriale tra V_1, ..., V_n. Sia V_i di dimensione finita $\forall\, i \le n$. Siano $\mathcal{B}_{V_i} = \left\{\underline{v_1^{(i)}},\dots,\underline{v_{k_i}^{(i)}}\right\}$ base di V_i, dove $k_i = \dim V_i$. Si definisce l'applicazione n-lineare $\underline{v_{j_1}^{(1)}}^* \otimes \cdots \otimes \underline{v_{j_n}^{(n)}}^* \in \mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K})$ univocamente determinata dalla seguente relazione:

$$\underline{v_{j_1}^{(1)}}^* \otimes \cdots \otimes \underline{v_{j_n}^{(n)}}^*(\underline{w_1},\dots,\underline{w_n}) = \underline{v_{j_1}^{(1)}}^*(\underline{w_1}) \cdots \underline{v_{j_n}^{(n)}}^*(\underline{w_n}).$$

Si definisce l'insieme $\mathcal{B}_\otimes$ nel seguente modo:

$$\mathcal{B}_\otimes = \left\{\underline{v_{j_1}^{(1)}}^* \otimes \cdots \otimes \underline{v_{j_n}^{(n)}}^* \mid 1 \le j_1 \le k_1,\ \dots,\ 1 \le j_n \le k_n\right\}.$$

Poiché ogni applicazione n-lineare è univocamente determinata dai valori che assume ogni combinazione degli elementi delle basi degli spazi V_i, vi è un isomorfismo tra $\mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K})$ e $\mathbb{K}^{\mathcal{B}_{V_1} \times \cdots \times \mathcal{B}_{V_n}}$, che ha dimensione $\prod_{i=1}^n k_i = k$. Pertanto anche $\dim \mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K}) = k$.

Poiché $\mathcal{B}_\otimes$ genera $\mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K})$ e i suoi elementi sono tanti quanto è la dimensione dello spazio, tale insieme è una base di $\mathrm{Mult}(V_1 \times \ldots \times V_n, \mathbb{K})$.

Se $V_i = V_1 = V\ \forall\, i \le n$, si dice che $\mathrm{Mult}(V^n, \mathbb{K})$ è lo spazio delle forme n-lineari di V.

Applicazioni multilineari simmetriche

Sia V uno spazio di dimensione n. Una forma k-lineare f si dice simmetrica ed appartiene allo spazio $\mathrm{Sym}^k(V)$ se:

$$f(\underline{x_1},\dots,\underline{x_k}) = f(\underline{x_{\sigma(1)}},\dots,\underline{x_{\sigma(k)}}), \quad \forall\, \sigma \in S_k.$$

Poiché ogni applicazione n-lineare simmetrica è univocamente determinata dai valori che assume negli elementi della base disposti in modo non decrescente, $\dim \mathrm{Sym}^k(V) = \binom{n+k-1}{k}$.

Sia $\mathcal{B} = \{\underline{v_1},\dots,\underline{v_n}\}$ una base di V. Dato un insieme di indici non decrescente I, si definisce il prodotto simmetrico (o *prodotto vee*) $\underline{v_{i_1}}^* \vee \cdots \vee \underline{v_{i_k}}^*$ tra elementi della base come la forma k-lineare simmetrica determinata dalla seguente relazione:

$$\underline{v_{i_1}}^* \vee \cdots \vee \underline{v_{i_k}}^* = \sum_{\sigma \in S_k} \underline{v_{i_{\sigma(1)}}}^* \otimes \cdots \otimes \underline{v_{i_{\sigma(k)}}}^*.$$

Si definisce l'insieme:

$$\mathcal{B}_{\mathrm{Sym}} = \left\{\underline{v_{i_1}}^* \vee \cdots \vee \underline{v_{i_k}}^* \mid 1 \le i_1 \le \cdots \le i_k \le n\right\}.$$

L'insieme $\mathcal{B}_{\mathrm{Sym}}$ è sia generatore che linearmente indipendente su $\mathrm{Sym}^k(V)$, ed è dunque base. Allora $\dim \mathrm{Sym}^k(V) = \binom{n+k-1}{k}$.

Applicazioni multilineari alternanti

Sia V uno spazio di dimensione n. Una forma k-lineare f si dice alternante (o antisimmetrica) ed appartiene allo spazio $\Lambda^k(V)$ (talvolta scritto come $\mathrm{Alt}^k(V)$) se:

$$f(x_1,\dots,x_k)=0 \iff \exists\, i,j\le k \mid x_i=x_j.$$

Questo implica che:

$$f(x_1,\dots,x_k)=\mathrm{sgn}(\sigma)f(x_{\sigma(1)},\dots,x_{\sigma(n)}),\quad \forall\,\sigma\in S_k$$

Se $k>n$, un argomento della base di V si ripete sempre nel computo f negli elementi della base, e quindi ogni alternante è pari a $\underline{0}$, ossia $\dim\Lambda^k(V)=0$.

Sia $\mathcal{B}=\{\underline{v_1},\dots,\underline{v_n}\}$ una base di V. Dato un insieme di indici crescente I, si definisce il prodotto esterno (o *prodotto wedge*) $\underline{v_{i_1}}^*\wedge\cdots\wedge\underline{v_{i_k}}^*$ tra elementi della base come la forma k-lineare alternante determinata dalla relazione:

$$\underline{v_{i_1}}^*\wedge\cdots\wedge\underline{v_{i_k}}^*=\sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)\,\underline{v_{i_{\sigma(1)}}}^*\otimes\cdots\otimes\underline{v_{i_{\sigma(k)}}}^*.$$

Si definisce l'insieme:

$$\mathcal{B}_\Lambda=\left\{\underline{v_{i_1}}^*\wedge\cdots\wedge\underline{v_{i_k}}^*\mid 1\le i_1<\cdots<i_k\le n\right\}.$$

L'insieme $\mathcal{B}_\Lambda$ è sia generatore che linearmente indipendente su $\Lambda^k(V)$, ed è dunque base. Allora $\dim\Lambda^k(V)=\binom{n}{k}$. Riassumendo si può scrivere:

$$\dim\Lambda^k(V)=\begin{cases}0 & \text{se } k>n\,,\\ \binom{n}{k} & \text{altrimenti.}\end{cases}$$

Quindi è quasi sempre vero che:

$$\underbrace{\dim\mathrm{Sym}^k(V)}_{=\binom{n+k-1}{k}}+\underbrace{\dim\Lambda^k(V)}_{\le\binom{n}{k}}<\underbrace{\dim\mathrm{Mult}(V^k,\mathbb{K})}_{=n^k},$$

e dunque che $\mathrm{Sym}^k(V)+\Lambda^k(V)\neq\mathrm{Mult}(V^k,\mathbb{K})$.

Determinante di una matrice

Si definisce il determinante det di una matrice di taglia $n\times n$ come l'unica forma n-lineare alternante di $(\mathbb{K}^n)^n$ tale che $\det(\underline{e_1},\dots,\underline{e_n})=1$ (infatti $\dim\Lambda^n(V)=\binom{n}{n}=1$, e quindi ogni forma alternante è multipla delle altre, eccetto per lo zero).

Equivalentemente $\det=\underline{e_1}^*\wedge\cdots\wedge\underline{e_n}^*$.

Siano $A,\,B\in M(n,\mathbb{K})$. Si scrive $\det(A)$ per indicare $\det(A_1,\dots,A_n)$. Vale pertanto la seguente relazione:

$$\det(A)=\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)\,a_{1\sigma(1)}\cdots a_{n\sigma(n)}.$$

- $\det(I_n)=1$,
- $\det\begin{pmatrix}a & b\\ c & d\end{pmatrix}=ad-bc$,
- $\det\begin{pmatrix}a & b & c\\ d & e & f\\ g & h & i\end{pmatrix}=a(ei-fh)-b(di-fg)+c(dh-eg)$,
- $\det(A)\neq0\iff A$ invertibile (ossia non singolare),
- $\det(\lambda A)=\lambda^n A$,
- $\det(A)=\det(A^\top)$ (è sufficiente applicare la definizione di det e manipolare algebricamente il risultato per evidenziare l'uguaglianza),
- se A è antisimmetrica, n è dispari e $\mathrm{char}\,\mathbb{K}\neq2$, $\det(A)=\det(-A^\top)=(-1)^n\det(A^\top)=(-1)^n\det(A)=-\det(A)\implies\det(A)=0$ (quindi ogni matrice antisimmetrica di taglia dispari non è invertibile),
- $\det(AB)=\det(A)\det(B)$ (*teorema di Binet* – è sufficiente considerare la forma $\frac{\det(AB)}{\det(B)}$ in funzione delle righe di A e determinare che tale forma è alternante e che vale 1 nell'identità, e che, per l'unicità del determinante, deve obbligatoriamente essere pari a $\det(A)$),
- se A è invertibile, $\det(A^{-1})=\det(A)^{-1}$,
- $\det\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}=\det(\lambda_1\underline{e_1},\dots,\lambda_n\underline{e_n})=\prod_{i=1}^n\lambda_i$,
- se A è triangolare superiore (o inferiore), allora $\det(A)$ è il prodotto degli elementi sulla sua diagonale principale,
- $\det(A_1,\dots,A_n)=\mathrm{sgn}(\sigma)\det(A_{\sigma(1)},\dots,A_{\sigma(n)})$, $\forall\,\sigma\in S_n$ (infatti det è alternante),
- $\det\left(\begin{array}{c|c}A & B\\ \hline C & D\end{array}\right)=\det(AD-BC)$, se C e D commutano e D è invertibile,
- $\det\left(\begin{array}{c|c}A & B\\ \hline 0 & C\end{array}\right)=\det(A)\det(C)$,
- se A è nilpotente (ossia se $\exists\,k\mid A^k=0$), $\det(A)=0$,
- se A è idempotente (ossia se $A^2=A$), allora $\det(A)=1$ o $\det(A)=0$,
- se A è ortogonale (ossia se $AA^\top=I_n$), allora $\det(A)=\pm1$,
- se $A\in M(n,\mathbb{C})$, $\det(\overline{A})=\overline{\det(A)}$ (segue direttamente dallo sviluppo di Laplace del determinante),
- se A è unitaria (ossia se $AA^*=I_n$), allora $|\det(A)|=1$,
- se A è un'involuzione (ossia se $A^2=I_n$), allora $\det(A)=\pm1$,
- se ogni minore di taglia k di A ha determinante nullo, allora tutti i minori di A taglia maggiore o uguale a k hanno determinante nullo (è una diretta applicazione dello sviluppo di Laplace).

Le operazioni del terzo tipo dell'algoritmo di eliminazione di Gauss (ossia l'aggiunta a una riga di un multiplo di un'altra riga – a patto che le due righe siano distinte) non alterano il determinante della matrice iniziale, mentre lo scambio di righe ne inverte il segno (corrisponde a una trasposizione di S_n). L'operazione del secondo tipo (la moltiplicazione di una riga per uno scalare) altera il determinante moltiplicandolo per tale scalare.

Inoltre, se D è invertibile, vale la decomposizione di Schur:

$$\left(\begin{array}{c|c}A & B\\ \hline C & D\end{array}\right)=\left(\begin{array}{c|c}I_k & BD^{-1}\\ \hline 0 & I_k\end{array}\right)\left(\begin{array}{c|c}A-BD^{-1}C & 0\\ \hline 0 & D\end{array}\right)\left(\begin{array}{c|c}I_k & 0\\ \hline D^{-1}C & I_k\end{array}\right),$$

dove $k\times k$ è la taglia di A. Pertanto vale la seguente relazione, sempre se D è invertibile:

$$\det\left(\begin{array}{c|c}A & B\\ \hline C & D\end{array}\right)=\det(A-BD^{-1}C)\det(D).$$

È possibile computare il determinante di A, scelta la riga i, mediante lo sviluppo di Laplace:

$$\det(A)=\sum_{j=1}^n a_{ij}\,\mathrm{Cof}_{i,j}(A)=\sum_{j=1}^n(-1)^{i+j}a_{ij}\det(A_{i,j}).$$

Si definisce matrice di Vandermonde una matrice $A\in M(n,\mathbb{K})$ della forma:

$$A=\begin{pmatrix}1 & x_1 & x_1^2 & \dots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & x_n^2 & \dots & x_n^{n-1}.\end{pmatrix}$$

Vale allora che:

$$\det(A)=\prod_{1\le i<j\le n}(x_j-x_i),$$

verificabile notando che $\det(A)$ è di grado $\frac{n(n-1)}{2}$ e che ponendo $x_i=x_j$ per una coppia (i,j), tale matrice ha due righe uguali, e quindi determinante nullo

$\implies(x_j-x_i)\mid\det(A)\overset{\text{UFD}}{\implies}\det(A)=\prod_{1\le i<j\le n}(x_j-x_i)$.

Pertanto una matrice di Vandermonde è invertibile se e solo se la sua seconda colonna contiene tutti scalari distinti nelle coordinate. Tale matrice risulta utile nello studio dell'interpolazione di Lagrange (ossia nella dimostrazione dell'unicità del polinomio di $n-1$ grado tale che $p(\alpha_i)=\beta_i$ per i coppie $(\alpha_i,\,\beta_i)$ con α_i tutti distinti).

Rango tramite il determinante degli orlati

Si dicono *sottomatrici* della matrice $A \in M(m,n,\mathbb{K})$ tutte le matrici contenute in A, ossia le matrici B che sono ottenibili da A mantenendo solo alcune sue righe e colonne. In generale, si scrive $A^{j_1,\dots,j_s}_{i_1,\dots,i_t}$ per indicare la sottomatrice ottenuta da A mantenendo le colonne di indice j_1, ..., j_s e le righe di indice i_1, ..., i_t. Quando è omesso l'indice delle colonne o l'indice delle righe, si sottintende di aver mantenuto o tutte le colonne o tutte le righe (e.g. $A_{1,2}$ è la sottomatrice di A ottenuta mantenendo tutte le colonne e le prime due righe). Si dice che M è *minore* di A una sua sottomatrice quadrata. Si chiamano *orlati* di un minore M di taglia k i minori di taglia $k+1$ di A aventi M come minore.

- se B è una sottomatrice di A, allora $\operatorname{rg}(B) \leq \operatorname{rg}(A)$ (è sufficiente prendere un numero massimo di colonne linearmente indipendenti di B e mostrare che le relative colonne in A sono ancora linearmente indipendenti),
- $\operatorname{rg}(A) = \max\{\operatorname{rg}(B) \mid B \text{ sottomatrice di } A\}$ (è sufficiente utilizzare il precedente risultato; infatti A è una sottomatrice di A),
- $\operatorname{rg}(A) = \max\{\operatorname{rg}(B) \mid B \text{ minore invertibile di } A\} = \max\{n \mid \text{esiste un minore di } A \text{ di taglia } n \text{ invertibile}\}$ (è sufficiente utilizzare la prima disuguaglianza e considerare un minore di A composto dalle righe e le colonne linearmente indipendenti di A, che sono dello stesso numero, dal momento che il rango per righe è uguale al rango per colonne),
- $\operatorname{rg}(A)$ è il più piccolo naturale n tale per cui, per ogni minore M di A di taglia maggiore di n, $\det(M) = 0$ (ossia M è singolare; segue direttamente dal precedente risultato),
- $\operatorname{rg}(A)$ è il più piccolo naturale n tale per cui, per ogni minore M di A di taglia $n+1$, $\det(M) = 0$ (ossia M è singolare; segue dal precedente risultato a cui si combina lo sviluppo di Laplace del determinante – se ogni minore di taglia k ha determinante nullo, anche tutti i minori di taglia maggiore di k hanno determinante nullo).
- esiste un minore M di taglia k di A con $\det(M) \neq 0 \implies \operatorname{rg}(A) \geq k$ (deriva direttamente dall'ultimo risultato sul rango),
- per ogni minore M di taglia k di A vale che $\det(M) = 0 \implies \operatorname{rg}(A) < k$ (come sopra).

Si può facilitare lo studio del rango tramite il teorema di Kronecker (o degli orlati): $\operatorname{rg}(A)$ è il più piccolo naturale n tale per cui esista un minore M di taglia k con $\det(M) \neq 0$ e per cui ogni suo orlato O è tale per cui $\det(O) = 0$.

Sia infatti, senza perdità di generalità, $M = A^{1,\dots,k}_{1,\dots,k}$ tale minore (altrimenti è sufficiente considerare una permutazione delle righe e delle colonne per ricadere in questo caso; tale permutazione è ammessa dall'algoritmo di Gauss). Si mostra che $A^j \in \operatorname{Span}(A^1,\dots,A^k)\ \forall\, j > k$. Si consideri ogni orlato M_j di M ottenuto scegliendo la j-esima colonna di A: per ipotesi $\det(M_j) = 0$, ed il rango è almeno k. Quindi $\operatorname{rg}(M_j) = k$; poiché le prime k righe sono linearmente indipendenti, l'ultima riga aggiunta deve certamente appartenere al loro sottospazio generato. Quindi ogni riga di $A^{1,\dots,k,j}$ appartiene al sottospazio $\operatorname{Span}(A_1,\dots,A_k)$, da cui si deduce che $\operatorname{rg}(A^{1,\dots,k,j}) = k$, e quindi che $\operatorname{rg}(A^{1,\dots,k,j}) = k \implies A^j \in \operatorname{Span}(A^1,\dots,A^k) \implies \operatorname{rg}(A) = k$.

Autovalori, diagonalizzabilità e triangolabilità

Sia $f \in \operatorname{End}(V)$. Si dice che $\lambda \in \mathbb{K}$ è un autovalore di f se e solo se $\exists\, \underline{v} \neq \underline{0}$, $\underline{v} \in V$ tale che $f(\underline{v}) = \lambda \underline{v}$, e in tal caso si dice che $\underline{v}$ è un autovettore relativo a λ. Un autovalore è tale se esiste una soluzione non nulla a $(f - \lambda \operatorname{Id}_V)\underline{v} = \underline{0}$, ossia se e solo se:

$$\det(f - \lambda \operatorname{Id}_V) = 0.$$

Questa relazione è ben definita dacché il determinante è invariante per qualsiasi cambio di base applicato ad una matrice associata di f. Si definisce allora $p_f(\lambda) = \det(f - \lambda \operatorname{Id}_V)$, detto polinomio caratteristico di f, ancora invariante per matrici associate a f. Si denota inoltre con spettro di f l'insieme $\operatorname{sp}(f)$ degli autovalori di f e con $V_\lambda = \operatorname{Ker}(f - \lambda \operatorname{Id}_V)$ lo spazio degli autovettori relativo a λ, detto autospazio di λ.

Si definisce la molteplicità algebrica $\mu_{a,f}(\lambda)$ di un autovalore λ come la molteplicità che assume come radice del polinomio $p_f(\lambda)$. Si definisce la molteplicità geometrica $\mu_{g,f}(\lambda)$ di un autovalore λ come la dimensione del suo autospazio V_λ. Quando è noto l'endomorfismo che si sta considerando si omette la dicitura f nel pedice delle molteplicità.

- $p_f(\lambda)$ ha sempre grado $n = \dim V$,
- $p_f(\lambda)$ è sempre monico a meno del segno,
- il coefficiente di λ^n è sempre $(-1)^n$,
- il coefficiente di λ^{n-1} è $(-1)^{n+1} \operatorname{tr}(f)$,
- il termine noto di $p_f(\lambda)$ è $\det(f - 0 \cdot \operatorname{Id}_V) = \det(f)$,
- il termine noto di $p_f(\lambda)$ è $\det(f - 0 \cdot \operatorname{Id}_V) = \det(f)$,
- $p_f(\lambda) = \sum_{i=0}^{n} (-\lambda)^i (\sum \det(M_{n-i}))$ dove i M_j sono i minori principali di taglia j,
- poiché $p_f(\lambda)$ appartiene all'anello euclideo $\mathbb{K}[\lambda]$, che è dunque un UFD, esso ammette al più n radici,
- $\operatorname{sp}(f)$ ha al più n elementi, ossia esistono al massimo n autovalori (dalla precedente considerazione),
- se $\mathbb{K} = \mathbb{C}$ e $p_f \in \mathbb{R}[\lambda]$, $\lambda \in \operatorname{sp}(f) \iff \overline{\lambda} \in \operatorname{sp}(f)$ (infatti λ è soluzione di p_f, e quindi anche $\overline{\lambda}$ deve esserne radice, dacché i coefficienti di p_f sono in $\mathbb{R}$),
- se $\mathbb{K}$ è un campo algebricamente chiuso, $p_f(\lambda)$ ammette sempre almeno un autovalore distinto (o esattamente n se contati con molteplicità),
- $0 \in \operatorname{sp}(f) \iff \dim \operatorname{Ker} f > 0 \iff \operatorname{rg} f < 0 \iff \det(f) = 0$,
- autovettori relativi ad autovalori distinti sono sempre linearmente indipendenti,
- dati λ_1, ..., λ_k autovalori di f, gli spazi V_{λ_1}, ..., V_{λ_k} sono sempre in somma diretta,
- $\sum_{i=1}^{k} \mu_a(\lambda_i)$ corrisponde al numero di fattori lineari di $p_f(\lambda)$,
- $\sum_{i=1}^{k} \mu_a(\lambda_i) = n \iff p_f(\lambda)$ è completamente fattorizzabile in $\mathbb{K}[\lambda]$,
- vale sempre la disuguaglianza $n \geq \mu_a(\lambda) \geq \mu_g(\lambda) \geq 1$ (è sufficiente considerare una base di V_λ estesa a base di V e calcolarne il polinomio caratteristico sfruttando i blocchi della matrice associata, notando che $\mu_g(\lambda)$ deve forzatamente essere minore di $\mu_a(\lambda)$),
- vale sempre la disuguaglianza $n \geq \sum_{i=1}^{k} \mu_a(\lambda_i) \geq \sum_{i=1}^{k} \mu_g(\lambda_i)$,
- se $W \subseteq V$ è un sottospazio f-invariante, allora $p_{f|_W}(\lambda) \mid p_f(\lambda)$[1] (è sufficiente prendere una base di W ed estenderla a base di V, considerando poi la matrice associata in tale base, che è a blocchi),
- se $W \subseteq V$ è un sottospazio f-invariante, ed estesa una base $\mathcal{B}_W$ di W ad una $\mathcal{B}$ di V, detto $U = \operatorname{Span}(\mathcal{B} \setminus \mathcal{B}_W)$ il supplementare di W che si ottiene da tale base $\mathcal{B}$, vale che $p_f = p_{f|_W}(\lambda) \cdot p_{\hat{f}}$, dove $\hat{f} : V/W \to V/W$ è tale che $\hat{f}(\underline{u} + W) = f(\underline{u}) + W$ (come prima, è sufficiente considerare una matrice a blocchi),
- se $V = W \oplus U$, dove sia W che U sono f-invarianti, allora $p_f = p_{f|_W}(\lambda) \cdot p_{f|_U}(\lambda)$ (la matrice associata in un'unione di basi di W e U è infatti diagonale a blocchi),
- se sia W che U sono f-invarianti, allora f è diagonalizzabile se e solo se sia $f|_W$ che $f|_U$ lo sono,
- se f è nilpotente, $p_f(\lambda) = \lambda^n$ (è sufficiente considerare un eventuale altro autovalore diverso da zero e mostrare che se tale autovalore esistesse, f non sarebbe nilpotente),
- un endomorfismo è nilpotente se e solo se $f^n = 0$ (discende direttamente dal teorema di Hamilton-Cayley e dalla forma di p_f),
- l'unico endomorfismo diagonalizzabile e nilpotente è quello nullo,
- in un campo algebricamente chiuso, un endomorfismo è diagonalizzabile se e solo se è semisemplice (i.e. se ogni sottospazio f-invariante ammette un supplementare f-invariante).

[1] Quando si lavora su degli endomorfismi, la notazione $f|_W$ è impiegata per considerare f ristretta a W sia sul dominio che sul codominio.

Si dice che f è diagonalizzabile se V ammette una base per cui la matrice associata di f è diagonale, o equivalentemente se, dati λ_1, ..., λ_k autovalori di f, si verifica che:

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}.$$

Ancora in modo equivalente si può dire che f è diagonalizzabile se e solo se:

$$\begin{cases} \sum_{i=1}^{k} \mu_a(\lambda_i) = n, \\ \mu_g(\lambda_i) = \mu_a(\lambda_i) \ \forall\, 1 \leq i \leq k, \end{cases}$$

ossia se il polinomio caratteristico è completamente fattorizzabile in $\mathbb{K}[\lambda]$ (se non lo fosse, la somma diretta $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ avrebbe forzatamente dimensione minore di V, ed esisterebbero altri autovalori in un qualsiasi campo di spezzamento di $p_f(\lambda)$) e se $\sum_{i=1}^{k} \mu_g(\lambda_i) = n$. Tale condizione, in un campo algebricamente chiuso, si riduce a $\mu_g(\lambda_i) = \mu_a(\lambda_i)$, $\forall\, 1 \leq i \leq k$.

Considerando la forma canonica di Jordan di f, si osserva anche che f è diagonalizzabile se e solo se per ogni autovalore la massima taglia di un blocco di Jordan è esattamente 1, ossia se il polinomio minimo di f è un prodotto di fattori lineari distinti (i.e. se $\varphi_f(t) = \prod_i (t - \lambda_i)$). Si può fare la stessa considerazione guardando al teorema di decomposizione primaria (gli indici di Fitting del sottospazio generalizzato sono esattamente le moltiplicità algebriche degli autovalori nel polinomio minimo).

Data f diagonalizzabile, la matrice diagonale J a cui f è associata è, dati gli autovalori λ_1, ..., λ_k, una matrice diagonale dove λ_i compare sulla diagonale esattamente $\mu_g(\lambda_i)$ volte.

Data $A \in M(n, \mathbb{K})$, A è diagonalizzabile se e solo se f_A, l'applicazione indotta dalla matrice A, è diagonalizzabile, ossia se A è simile ad una matrice diagonale J, computabile come prima. Si scrive in particolare $p_A(\lambda)$ per indicare $p_{f_A}(\lambda)$.

Una matrice $P \in \mathrm{GL}(M(n, \mathbb{K}))$ tale che $A = PJP^{-1}$, è tale che $AP = PJ$: presa la i-esima colonna, allora, $AP^{(i)} = PJ^{(i)} = P^{(i)}$; ossia è sufficiente costruire una matrice P dove l'i-esima colonna è un autovettore relativo all'autovalore presente in J_{ii} linearmente indipendente con gli altri autovettori presenti in P relativi allo stesso autovalore (esattamente nello stesso modo in cui si costruisce in generale tale P con la forma canonica di Jordan).

Se A e B sono diagonalizzabili, allora $A \sim B \iff p_A(\lambda) = p_B(\lambda)$ (infatti due matrici diagonali hanno lo stesso polinomio caratteristico se e solo se compaiono gli stessi identici autovalori).

Se f è diagonalizzabile, allora ogni spazio W f-invariante di V è tale che:

$$W = (W \cap V_{\lambda_1}) \oplus \cdots \oplus (W \cap V_{\lambda_k}),$$

dove λ_1, ..., λ_k sono gli autovalori distinti di f, e dunque $f|_W$ è sempre diagonalizzabile, se f lo è. In generale, dato un sottospazio W di V che è f-invariante, si può facilmente costruire un suo supplementare f-invariante. È infatti sufficiente prendere una base di W ed estenderla a base di V completandola tramite una base di autovettori di V.

Se f è diagonalizzabile, anche f^k lo è, per ogni $k \in \mathbb{N}$. Se ogni vettore di V è un autovettore di f, allora $f = \lambda \mathrm{Id}$, con $\lambda \in \mathbb{K}$ (è sufficiente considerare l'eventuale esistenza di più autospazi e due vettori $\underline{v}$ e $\underline{w}$ di due autospazi distinti e considerare le due scritture possibili di $f(\underline{v} + \underline{w})$).

Si dice infine che f è triangolabile (o triangolarizzabile) se V ammette una base per cui la matrice associata di f è triangolare superiore (o inferiore, dal momento che è sufficiente riordinare dal basso la base per ottenere una matrice associata triangolare superiore). Vale in particolare che f è triangolabile se e soltanto se $p_f(\lambda)$ è completamente riducibile in fattori lineari in $\mathbb{K}$ (dunque, nel caso di $\mathbb{K}$ algebricamente chiuso, f è sempre triangolabile). Infatti, se f è triangolabile, il polinomio caratteristico ha come radici esattamente gli elementi sulla diagonale della matrice associata di f nella base $\mathcal{B}$ in cui tale matrice è triangolare superiore (e dunque $p_f(\lambda)$ è riducibile in fattori lineari). Se invece $p_f(\lambda)$ è riducibile in fattori lineari, si può applicare il seguente algoritmo (su cui si fonda induttivamente la dimostrazione della proposizione):

1. Si calcolino le basi degli autospazi di f,
2. Si estenda l'unione $\mathcal{B}_A$ di queste basi a una base $\mathcal{B}$ di V,
3. Si consideri la matrice associata di f nella base $\mathcal{B}$, della forma:
$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right),$$
dove A è una matrice diagonale contenente gli autovalori di $\mathrm{sp}(f)$,
4. Se $M_{\mathcal{B}}(f)$ è triangolare superiore, l'algoritmo termina. Altrimenti si ripeta l'algoritmo su C (ossia sull'endomorfismo $p_W \circ f|_W \in \mathrm{End}(W)$, dove W è il sottospazio generato dai vettori aggiunti alla base $\mathcal{B}_A$ per costruire la base $\mathcal{B}$).

Inoltre, se W è un sottospazio f-invariante di V, e f è triangolabile, anche $f|_W$ lo è (infatti, in tal caso, il polinomio caratteristico di f si riduce in fattori lineari).

Diagonalizzabilità e triangolabilità simultanea

Due endomorfismi f, $g \in \mathrm{End}(V)$ diagonalizzabili si dicono simultaneamente diagonalizzabili se esiste una base $\mathcal{B}$ di V tale per cui sia la matrice associata di f in $\mathcal{B}$ che quella di g sono diagonali. Vale in particolare che f e g sono simultaneamente diagonalizzabili se e solo se $f \circ g = g \circ f$. Per trovare tale base è sufficiente, dati λ_1, ..., λ_k autovalori di f, considerare $g|_{V_{\lambda_i}}$ $\forall\, 1 \leq i \leq k$ (V_{λ_i} è infatti g-invariante, dacché, per $\underline{v} \in V_{\lambda_i}$, $f(g(\underline{v})) = g(f(\underline{v})) = g(\lambda_i \underline{v}) = \lambda_i g(\underline{v}) \implies g(\underline{v}) \in V_{\lambda_i}$), che, essendo una restrizione di un endomorfismo diagonalizzabile su un sottospazio invariante, è diagonalizzabile: presa allora una base di autovettori di $g|_{V_{\lambda_i}}$, questi sono anche base di autovettori di V_{λ_i}; unendo tutti questi autovettori in un'unica base $\mathcal{B}$ di V, si otterrà dunque che una base in cui le matrici associate di f e g sono diagonali.

Analogamente due endomorfismi f, $g \in \mathrm{End}(V)$ triangolabili si dicono simultaneamente triangolabili se esiste una base $\mathcal{B}$ in cui $M_{\mathcal{B}}(f)$ e $M_{\mathcal{B}}(g)$ sono due matrici triangolari superiori. Non è generalmente vero che due endomorfismi simultaneamente triangolabili commutano; è tuttavia vero il viceversa. Se infatti f e g sono due endomorfismi triangolabili tali che $f \circ g = g \circ f$, allora si può riapplicare, con le dovute modifiche, il precedente algoritmo di triangolarizzazione (anche questa volta dimostrabile per induzione):

1. Si calcolino le basi degli autospazi di f e si consideri $f|_U$, dove $U = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$,
2. Si cerchi una base $\mathcal{B}_U$ in cui $f|_U$ e $g|_U$ sono simultaneamente diagonalizzabili (osservando che g è U-invariante),
3. Si estenda tale base $\mathcal{B}_U$ ad una base $\mathcal{B}$ di V e si chiami W il sottospazio $\mathrm{Span}(\mathcal{B}_W)$, dove $\mathcal{B}_W := \mathcal{B} \setminus \mathcal{B}_U$,
4. Si considerino la matrice associata di f e di g nella base $\mathcal{B}$, della forma:
$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right),$$
$$M_{\mathcal{B}}(g) = \left(\begin{array}{c|c} A' & B' \\ \hline 0 & C' \end{array}\right),$$
dove A e A' sono matrici diagonali contenente gli autovalori dei rispettivi endomorfismi,
5. Se le due matrici sono triangolari superiori, l'algoritmo termina. Altrimenti si ripeta l'algoritmo su C e C' (ossia sugli endomorfismi $p_W \circ f|_W$, $p_W \circ g|_W \in \mathrm{End}(W)$, i quali commutano, dal momento che vale l'identità $CC' = C'C$, dedotta moltiplicando le due matrici associate di sopra).

Polinomio minimo

Sia $f \in \mathrm{End}(V)$. Si può allora definire l'applicazione $\sigma_f : \mathbb{K}[x] \to \mathrm{End}(V)$ tale per cui $\sigma_f(p) = p(f)$, dove per $p(f)$ s'intende la riscrittura di p a cui si sostituisce all'usuale somma e all'usuale prodotto, la somma di applicazioni e la composizione (intendendo, in particolare, i termini noti come multipli dell'identità $f^0 := \mathrm{Id}_V$). In particolare σ_f è un omomorfismo di anelli, ed è dunque anche un'applicazione lineare. σ_f non è mai iniettiva, ed esiste dunque sempre un polinomio p tale per cui $\sigma_f(p) = 0$, l'applicazione nulla (è sufficiente prendere $n^2 + 1$ potenze di f e osservare che devono essere linearmente dipendenti). Poiché $\mathbb{K}[x]$ è un PID, $\mathrm{Ker}\, \sigma_f$ è un ideale principale, e quindi esiste un polinomio monico φ_f, detto polinomio minimo di f, tale per cui $\mathrm{Ker}\, \sigma_f = (\varphi_f)$.

- $\varphi_f \mid p_f$ (teorema di Hamilton-Cayley),
- $\deg \varphi_f = d$ se e solo se Id_V, f, ..., f^{d-1} sono linearmente indipendenti e $f^d \in \mathrm{Span}(\mathrm{Id}_V, f, \ldots, f^{d-1})$,

- $\dim \mathbb{K}[f] = \deg \varphi_f$ (infatti, per il primo teorema di omomorfismo $\mathbb{K}[f] \cong \mathbb{K}[x]/(\varphi_f)$, da cui si ricava facilmente la dimensione dello spazio),
- Id_V, f, ..., f^{d-1} formano una base di $\mathbb{K}[f]$ (per i precedenti risultati), se $d = \deg \varphi_f$,
- φ_f e p_f condividono gli stessi fattori primi (se infatti non comparisse un autovalore come radice di φ_f, $\varphi_f(f)$ non sarebbe nullo),
- gli esponenti dei fattori lineari di φ_f sono esattamente gli indici di Fitting degli autospazi generalizzati di f,
- gli autovalori hanno moltiplicità algebrica 1 in φ_f se e solo se f è diagonalizzabile (è sufficiente utilizzare il precedente risultato, o considerare la forma canonica di Jordan),
- se f è nilpotente, $\varphi_f(t) = t^k$, dove k è l'indice di Fitting di $\operatorname{Ker} f$ (discende direttamente dalla forma di p_f se f è nilpotente),
- se $p \in \mathbb{K}[x]$ è tale per cui $p = p_1 \cdots p_k$ con p_1, ..., $p_k \in \mathbb{K}[x]$ coprimi, allora $\operatorname{Ker} p(f) = \operatorname{Ker} p_1(f) \oplus \cdots \oplus \operatorname{Ker} p_k(f)$ (teorema di decomposizione primaria; si dimostra facilmente attraverso il teorema di Bezout),
- $V = \widetilde{V_{\lambda_1}} \oplus \cdots \oplus \widetilde{V_{\lambda_k}}$, se λ_1, ..., λ_k sono tutti gli autovalori di f (deriva direttamente dal teorema di Hamilton-Cayley e dal teorema di decomposizione primaria, o, alternativamente, dalla decomposizione di Fitting).

Sia $\underline{v} \in V$. Si definisce allora l'applicazione $\mathrm{val}_{f,\underline{v}} : \mathbb{K}[x] \to V$ in modo tale che $\mathrm{val}_{f,\underline{v}}(p) = p(f)(\underline{v})$. Come prima, $\mathrm{val}_{f,\underline{v}}$ è un'applicazione lineare. Si osserva ancora che $\operatorname{Ker} \mathrm{val}_{f,\underline{v}}$ è un'ideale, e quindi che esiste un polinomio $\varphi_{f,\underline{v}}$ tale per cui $\operatorname{Ker} \mathrm{val}_{f,\underline{v}} = (\varphi_{f,\underline{v}})$. Tale polinomio viene denotato come polinomio minimo relativo al vettore $\underline{v}$. Si definisce in particolare $\mathbb{K}[f](\underline{v}) := \operatorname{Im} \mathrm{val}_{f,\underline{v}}$.

- $\varphi_{f,\underline{v}} \mid \varphi_f$ (infatti $\varphi_f(f) = 0$, e dunque $\varphi_f(f)$ annulla $\underline{v}$),
- $\deg \varphi_{f,\underline{v}} = d$ se e solo se $\underline{v}$, $f(\underline{v})$, ..., $f^{d-1}(\underline{v})$ sono linearmente indipendenti e $f^d(\underline{v}) \in \mathrm{Span}(\underline{v}, \dots, f^{d-1}(\underline{v}))$,
- $\dim \mathbb{K}[f](\underline{v}) = \deg \varphi_{f,\underline{v}}$ (si dimostra allo stesso modo in cui si è dimostrata la proposizione analoga per φ_f),
- $\underline{v}$, ..., $f^{d-1}(\underline{v})$ formano una base di $\mathbb{K}[f](\underline{v})$, se $d = \deg \varphi_{f,\underline{v}}$.
- se $\underline{v_1}$, ..., $\underline{v_k}$ sono generatori di V, allora $\varphi_f = \mathrm{mcm}(\varphi_{f,\underline{v_1}}, \dots, \varphi_{f,\underline{v_k}})$ (è sufficiente mostrare che φ_f annulla una base e che il grado è minimale).
- se $\underline{v}$, ..., $f^k(\underline{v})$ sono linearmente indipendenti per qualche $\underline{v} \in V$, allora $\deg \varphi_f \geq \varphi_{f,\underline{v}} \geq k+1$.
- esiste sempre un vettore $\underline{v}$ tale per cui $\varphi_f = \varphi_{f,\underline{v}}$ (se $\mathbb{K}$ è infinito).
- $p(f)$ è invertibile $\iff \operatorname{Ker} p(f) = \{\underline{0}\}$ $\iff \mathrm{MCD}(\varphi_f, p) \in \mathbb{K}^*$, se $p \in \mathbb{K}[x]$ (è sufficiente applicare il teorema di Bezout).

Un vettore $\underline{v}$ si dice ciclico rispetto a f se gli n vettori $\underline{v}$, ..., $f^{n-1}(\underline{v})$ formano una base di V, in tal caso detta base ciclica di V.

Se $\mathbb{K}$ è infinito, V ammette una base ciclica se e solo se $p_f = \pm\varphi_f$ (infatti esiste sempre un vettore $\underline{v}$ tale per cui $\varphi_f = \varphi_{f,\underline{v}}$). In una base ciclica $\mathcal{B}$ la matrice associata si scrive nel seguente modo:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & & & -a_0 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix},$$

dove $\varphi_f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Tale matrice viene detta matrice compagna del polinomio $p := \varphi_f$ (e dunque ogni polinomio monico è in particolare il polinomio minimo di un qualche endomorfismo; analogamente ogni polinomio monico è, a meno del segno, un polinomio caratteristico).

La forma canonica di Jordan

Si definisce blocco di Jordan di taglia k relativo all'autovalore λ la seguente matrice:

$$J_{\lambda,k} := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix},$$

ossia la matrice che ha solo λ sulla diagonale, 1 sulla sopradiagonale e 0 nelle altre posizioni. Si può sempre restringere un blocco di Jordan a un blocco nilpotente considerando $J = J_{\lambda,k} - \lambda I_k$. Tale blocco ha come polinomio minimo $\varphi_J(t) = t^k$, e dunque $\varphi_{J_{\lambda,k}}(t) = (t-\lambda)^k$. Allo stesso modo si calcola $p_{J_{\lambda,k}}(t) = (t-\lambda)^k$. Si osserva dunque che $\mu_{a,J_{\lambda,k}}(\lambda) = \mu_{a,J}(0)$.

Poiché il polinomio caratteristico ed il polinomio minimo coincidono a meno del segno, esiste sempre una base ciclica per la quale $J_{\lambda,k}$ si scrive come matrice compagna di $\varphi_{J_{\lambda,k}}$.

Si definisce forma canonica di Jordan di un endomorfismo f una sua matrice associata in una base $\mathcal{B}$ tale per cui:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix},$$

dove J_1, ..., J_s sono blocchi di Jordan. La forma canonica di Jordan esiste sempre ed è unica a meno di permutazione dei blocchi, se tutti gli autovalori di f sono in $\mathbb{K}$ (teorema di Jordan; se gli autovalori di f non sono tutti in $\mathbb{K}$, si può sempre considerare un'estensione di campo in cui esistono).

Si definisce autospazio generalizzato relativo all'autovalore λ di $f \in \mathrm{End}(V)$ lo spazio:

$$\widetilde{V_\lambda} = \operatorname{Ker}(f - \lambda \mathrm{Id}_V)^n.$$

Una definizione alternativa, ma equivalente di $\widetilde{V_\lambda}$ è la seguente:

$$\widetilde{V_\lambda} = \{\underline{v} \in V \mid \exists\, k \in \mathbb{N} \mid (f - \lambda \mathrm{Id}_V)^k = \underline{0}\},$$

ossia $\widetilde{V_\lambda}$ è lo spazio dei vettori $\underline{v} \in V$ tali per cui, applicando ripetutamente $f - \lambda \mathrm{Id}_V$, si ottiene un autovettore relativo a λ (per dimostrare l'equivalenza delle due dimostrazioni è sufficiente considerare la decomposizione di Fitting). In generale, dalla catena della decomposizione di Fitting, si deduce in realtà che:

$$\widetilde{V_\lambda} = \operatorname{Ker}(f - \lambda \mathrm{Id}_V)^q \ \ \forall\, q \geq k,$$

dove k è la molteplicità algebrica di λ in φ_f (in particolare si ottiene sempre l'autospazio generalizzato sostituendo $\mu_a(\lambda)$ a q, dacché $\mu_a(\lambda) \geq k$).

In generale vale che:

$$V = \widetilde{V_{\lambda_1}} \oplus \cdots \oplus \widetilde{V_{\lambda_k}},$$

se λ_1, ..., λ_k sono tutti gli autovalori di f (vd. polinomio minimo). Inoltre, $f|_{\widetilde{V_\lambda}}$ ammette come autovalore soltanto λ (pertanto $\dim \widetilde{V_\lambda} = \mu_{a,f}(\lambda)$, confrontando i polinomi caratteristici). Si osserva inoltre che $\widetilde{V_\lambda}$ è sempre f-invariante. Infatti ogni f induce due catene di inclusione:

$$\operatorname{Ker} f^0 = \{\underline{0}\} \subsetneq \operatorname{Ker} f^1 \subsetneq \cdots \subsetneq \operatorname{Ker} f^k = \operatorname{Ker} f^{k+1} = \cdots,$$
$$\operatorname{Im} f^0 = V \supsetneq \operatorname{Im} f^1 \supsetneq \cdots \supsetneq \operatorname{Im} f^k = \operatorname{Im} f^{k+1} = \cdots,$$

dove k è detto indice di Fitting di f. Vale in particolare la decomposizione di Fitting:

$$V = \operatorname{Ker} f^k \oplus \operatorname{Im} f^k,$$

dove $f|_{\operatorname{Ker} f^k}$ è nilpotente (e dunque ammette solo 0 come autovalore; infatti $(f|_{\operatorname{Ker} f^k})^k = f^k|_{\operatorname{Ker} f^k} = 0$), mentre $f|_{\operatorname{Im} f^k}$ è invertibile (e dunque non ammette 0 come autovalore; infatti tale endomorfismo mantiene le dimensioni delle immagini).

- esiste sempre almeno un blocco di Jordan relativo a λ di ordine k, dove k è la molteplicità algebrica di λ in φ_f,
- la successione di $\operatorname{Ker}(f - \lambda \mathrm{Id}_V)^t - \operatorname{Ker}(f - \lambda \mathrm{Id}_V)^{t-1}$ all'aumentare di t è decrescente ed è definitivamente 0,
- il numero di blocchi di Jordan di taglia maggiore o uguale a t relativi a λ è esattamente $\operatorname{Ker}(f - \lambda \mathrm{Id}_V)^t - \operatorname{Ker}(f - \lambda \mathrm{Id}_V)^{t-1}$,

- il numero di blocchi di Jordan di taglia t relativi a λ è esattamente:
$$2\dim\operatorname{Ker}(f-\lambda\operatorname{Id}_V)^t-\dim\operatorname{Ker}(f-\lambda\operatorname{Id}_V)^{t+1}-\dim\operatorname{Ker}(f-\lambda\operatorname{Id}_V)^{t-1},$$
riscrivibile anche come:
$$\operatorname{rg}(f-\lambda\operatorname{Id}_V)^{t+1}+\operatorname{rg}(f-\lambda\operatorname{Id}_V)^{t-1}-2\operatorname{rg}(f-\lambda\operatorname{Id}_V),$$
(da queste due identità risulta evidente l'unicità della forma canonica di Jordan),
- esistono esattamente $\mu_g(\lambda)=\dim\operatorname{Ker}(f-\lambda\operatorname{Id}_V)$ blocchi relativi all'autovalore λ,
- $\mu_g(\lambda)=1\ \forall\,\lambda\in\operatorname{sp}(f)$ implica che vi sia un solo blocco relativo ad ogni $\lambda\in\operatorname{sp}(f)$; dal momento che ne deve esiste uno di ordine massimo, tale blocco ha taglia k, dove k è la molteplicità algebrica di λ in φ_f,
- $\mu_g(\lambda)=1\ \forall\,\lambda\in\operatorname{sp}(f)$ implica che $p_f=\pm\varphi_f$ (e dunque che f ammette una base ciclica; segue direttamente dal precedente risultato),
- una base di $\operatorname{Ker}(f-\lambda\operatorname{Id}_V)^t$ è data dai primi t vettori di ogni blocco relativo a λ,
- due matrici A, B sono simili se e solo se condividono la stessa forma canonica di Jordan (a meno di permutazione di blocchi; dunque la forma canonica di Jordan è un invariante completto della similitudine),
- Se $\mathbb{K}=\mathbb{C}$, vale l'identità:
$$\overline{\operatorname{Ker}(f-\lambda\operatorname{Id}_V)^k}=\operatorname{Ker}(f-\overline{\lambda}\operatorname{Id}_V)^k,$$
da cui è possibile ottenere una base dell'autospazio generalizzato relativo a $\overline{\lambda}$ coniugando una base dell'autospazio generalizzato relativo a λ (in particolare i due spazi hanno la stessa dimensione),
- Se $\mathbb{K}=\mathbb{C}$, la forma canonica di Jordan contiene tanti blocchi di taglia t relativi a λ quanti ve ne sono di relativi a $\overline{\lambda}$,
- Esistono e sono unici i due endomorfismi μ, $\delta\in\operatorname{End}(V)$ tale che μ sia diagonalizzabile, δ sia nilpotente e che $f=\mu+\delta$ (se esiste la forma canonica di Jordan; decomposizione di Jordan-Chevalley),
- Se $\forall\,\lambda\in\operatorname{sp}(f)$, $\mu_g(\lambda)=1$, allora esiste un numero finito di sottospazi invarianti e sono tutte le possibili somme dirette dei sottospazi degli autospazi generalizzati,
- Se $\mathbb{K}$ è infinito ed esiste $\lambda\in\operatorname{sp}(f)$ tale per cui $\mu_g(\lambda)>1$, allora esiste un numero infinito di sottospazi invarianti per ogni dimensione, da 1 a $\dim V-1$.

Calcolo di una base di Jordan

Si dice base di Jordan una qualsiasi base $\mathcal{B}$ tale per cui $M_{\mathcal{B}}(f)$ è una forma canonica di Jordan, se $f\in\operatorname{End}(V)$. Per calcolare una base di Jordan si può seguire il seguente algoritmo:

1. Si calcoli il polinomio caratteristico p_f di f e se ne estragga lo spettro $\operatorname{sp}(f)$,
2. Si consideri una base $\mathcal{B}$ di V e si ponga $A:=M_{\mathcal{B}}(f)$,
3. Si consideri ogni autovalore $\lambda\in\operatorname{sp}(f)$:
   a. Si consideri $B:=A-\lambda I_n$. Si calcoli il rango di B per ricavare $\mu_g(\lambda)$, indicante il numero di blocchi relativi a λ,
   b. Se possibile, si facciano considerazioni riguardo a come deve essere la forma canonica di Jordan. Altrimenti si calcoli il numero di blocchi tramite la formula presentata precedentemente,
   c. Si calcolino le matrici della forma B^i con $2\le i\le k-1$, dove k è la taglia del blocco più grande,
   d. Si calcolino le basi dei sottospazi U_i tali per cui:
   $$\begin{aligned}&\operatorname{Ker}B^k=\operatorname{Ker}B^{k-1}\oplus U_1,\\&\operatorname{Ker}B^{k-1}=\operatorname{Ker}B^{k-2}\oplus B(U_1)\oplus U_2,\\&\vdots\\&\operatorname{Ker}B=B^{k-1}(U_1)\oplus B^{k-2}(U_2)\oplus\cdots\oplus U_k;\end{aligned}$$
   e. Si scelgano da queste basi i vettori che generano ogni blocco relativo a λ (in particolare ogni vettore di base di U_i genera un blocco di taglia $k-1+i$),
   f. Per ogni blocco, generato dal vettore $\underline{v}$, si costruisca una base ordinata nel seguente modo:
   $$\mathcal{B}'=\{B^{t-1}\underline{v},\dots,B\underline{v},\underline{v}\},$$
   dove t è l'indice minimo per cui $B^t\underline{v}=0$;
4. Si uniscano ordinatamente a catena le basi ottenute in una base $\mathcal{B}_J$. La base $[]_{\mathcal{B}}^{-1}\mathcal{B}_J$ è allora base di Jordan. In particolare, se $P=\begin{pmatrix}\underline{v}_1\cdots\underline{v}_n\end{pmatrix}$, dove $\mathcal{B}_J=\{\underline{v}_1,\dots,\underline{v}_n\}$, vale che $J=P^{-1}AP$ è esattamente la forma canonica di Jordan individuata da tale base.

Se f è nilpotente, l'algoritmo può essere velocizzato notevolmente considerando solamente $B:=A$. Se f ha un solo autovalore λ e ammette una base ciclica (ossia esiste un solo blocco di Jordan), considerando $B:=A-\lambda I_n$, quasi ogni vettore è un vettore ciclico (è pertanto consigliato cercare un vettore in modo casuale, piuttosto che estendere tutte le basi dei kernel).

La forma canonica di Jordan reale

Sia $A\in M(n,\mathbb{R})$. Allora la forma canonica di Jordan reale è una variante reale della forma canonica di Jordan che esiste sempre (infatti gli autovalori di A non sono forzatamente in $\mathbb{R}$, e potrebbero dunque essere in $\mathbb{C}\setminus\mathbb{R}$). La forma canonica di Jordan reale si costruisce a partire da una forma canonica di Jordan J e una sua base di Jordan $\mathcal{B}$ associata. Tale forma canonica si costruisce mediante il seguente algoritmo:

1. Si scelga un autovalore z, se non si è già considerato il suo coniugato $\overline{z}$:
   a. Si prenda la base $\mathcal{B}_z=\{\underline{v_1},\dots,\underline{v_k},\overline{\underline{v_1}},\dots,\overline{\underline{v_k}}\}$ che genera i blocchi di z e $\overline{z}$ e si consideri la nuova base $\mathcal{B}'_z=\{\Re(\underline{v_1}),\Im(\underline{v_1}),\dots,\Re(\underline{v_k}),\Im(\underline{v_k})\}$,
   b. In tale base la forma canonica di Jordan varia eliminando i blocchi di $\overline{z}$, sostituendo all'autovalore $z=a+bi$ il seguente blocco:
   $$\begin{pmatrix}a&-b\\b&a\end{pmatrix},$$
   ed ingrandendo gli eventuali 1 mediante l'identità I_2 (tale processo prende il nome di complessificazione).
2. La matrice ottenuta dopo aver considerato tutti gli eventuali autovalori complessi è una forma canonica di Jordan reale, e la base ottenuta mediante tutti i processi di complessificazione è una base di Jordan reale.

Prodotto scalare e congruenza

Si consideri una mappa $\varphi:V\times V\to\mathbb{K}$. Si dice che φ è un prodotto scalare (e quindi che $\varphi\in\operatorname{PS}(V)$, lo spazio dei prodotti scalari) se è una forma bilineare simmetrica. In particolare vale la seguente identità:
$$\varphi\left(\sum_{i=1}^s a_i\underline{v_i},\sum_{j=1}^t b_j\underline{w_j}\right)=\sum_{i=1}^s\sum_{j=1}^t a_ib_j\varphi(\underline{v_i},\underline{w_j}).$$
Se $\mathcal{B}=\{\underline{v_1},\dots,\underline{v_n}\}$ è una base di V, si definisce $M_{\mathcal{B}}(\varphi)=(\varphi(\underline{v_i},\underline{v_j}))_{i,j=1-n}$ come la matrice associata al prodotto scalare φ. In particolare, se $a_\varphi:V\to V^*$ è la mappa lineare che associa a $\underline{v}$ il funzionale $\varphi(\underline{v},\cdot)\in V^*$ tale che $\varphi(\underline{v},\cdot)(\underline{w})=\varphi(\underline{v},\underline{w})$. Si scrive (V,φ) per indicare uno spazio vettoriale V dotato del prodotto scalare φ.

Si definisce prodotto scalare *standard* il prodotto φ tale che $\varphi(\underline{v},\underline{w})=[\underline{v}]_{\mathcal{B}}^\top[\underline{w}]_{\mathcal{B}}$.

Si dice che due vettori $\underline{v}$, $\underline{w}\in V$ sono ortogonali tra loro, scritto come $\underline{v}\perp\underline{w}$, se $\varphi(\underline{v},\underline{w})=0$. Dato W sottospazio di V, si definisce $W^\perp$ come il sottospazio di V dei vettori ortogonali a tutti i vettori di W. Si dice che φ è non degenere se $V^\perp=\{\underline{0}\}$. Si scrive in particolare che $V^\perp=\operatorname{Rad}(\varphi)$.

Si dice che $V=U\oplus^\perp W$ (ossia che U e W sono in somma diretta ortogonale) se $V=U\oplus W$ e $U\subseteq W^\perp$. Sia $i:W\to V$ tale che $\underline{w}\mapsto\underline{w}$. Si scrive $\varphi|_W$ intendendo $\varphi|_{W\times W}$.

Ad ogni prodotto scalare si può associare una forma quadratica (e viceversa) $q:V\to\mathbb{K}$ tale che $q(\underline{v})=\varphi(\underline{v},\underline{v})$. Un vettore $\underline{v}\in V$ si dice isotropo se $q(\underline{v})=0$ (altrimenti si dice anisotropo). Si definisce il cono isotropo $\operatorname{CI}(\varphi)$ come l'insieme dei vettori isotropi di V.

Se $\mathbb{K}=\mathbb{R}$, si dice che φ è semidefinito positivo ($\varphi\ge0$) se $q(\underline{v})\ge0\ \forall\,\underline{v}\in V$, e che è semidefinito negativo ($\varphi\le0$) se $q(\underline{v})\le0\ \forall\,\underline{v}\in V$. Si dice che φ è definito positivo ($\varphi>0$) se $\varphi\ge0$ e se $q(\underline{v})=0\iff\underline{v}=\underline{0}$, e che è definito negativo ($\varphi<0$) se $\varphi\le0$ e se $q(\underline{v})=0\iff\underline{v}=\underline{0}$.

Si dice che φ è definito se è definito positivo o definito negativo. Analogamente φ è semidefinito se è semidefinito positivo o semidefinito negativo.

Si scrive $\underline{v}^\perp$ per indicare tutti i vettori ortogonali a $\underline{v}$ (e quindi $\underline{v}^\perp = \mathrm{Span}(\underline{v})^\perp$). Si definisce $\iota : W \to V$ come l'applicazione tale per cui $\iota(\underline{w}) = \underline{w}$. Si scrive $\varphi|_U$ con U sottospazio di V per indicare il prodotto scalare $\varphi|_{U\times U}$.

Sia ora V di dimensione finita.

- $M_{\mathcal{B}}(\varphi)$ è simmetrica,
- $\varphi(\underline{v},\underline{w}) = [\underline{v}]_{\mathcal{B}}^\top M_{\mathcal{B}}(\varphi)[\underline{w}]_{\mathcal{B}}$,
- $M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}^*}^{\mathcal{B}}(a_\varphi)$,
- $\mathrm{Ker}\, a_\varphi = V^\perp$,
- φ è non degenere se e solo se $M_{\mathcal{B}}(\varphi)$ è invertibile,
- $W^\perp = \mathrm{Ker}\, i^\top \circ a_\varphi$,
- $a_\varphi(W^\perp) = \mathrm{Ann}(W) \cap \mathrm{Im}\, a_\varphi$,
- $\dim W + \dim W^\perp = \dim V + \dim(W \cap V^\perp)$ (da sopra),
- $V = W \oplus^\perp W^\perp$ se e solo se $\varphi|_W$ è non degenere ($\iff W \cap W^\perp = \mathrm{Rad}(\varphi|_W) = \{\underline{0}\}$),
- $V = W + W^\perp$ se e solo se $\mathrm{Rad}(\varphi|_W) \subseteq \mathrm{Rad}(\varphi)$,
- se $V = U \oplus^\perp W$, allora $\mathrm{Rad}(\varphi) = \mathrm{Rad}(\varphi|_U) \oplus \mathrm{Rad}(\varphi|_W)$,
- $V = \mathrm{Span}(\underline{w}) \oplus^\perp \mathrm{Span}(\underline{w})^\perp \iff q(\underline{w}) \neq 0 \iff \underline{w} \notin \mathrm{CI}(\varphi)$,
- $\mathrm{Span}(\underline{w}) \subseteq \mathrm{Span}(\underline{w})^\perp \iff \underline{w} \in \mathrm{CI}(\varphi)$,
- $(W^\perp)^\perp = W^{\perp\perp} = W + \mathrm{Rad}(\varphi) = W + V^\perp$,
- $\underline{v} \in V^\perp \iff \mathrm{Span}(\underline{v})^\perp = V$,
- $W^\perp = (\mathrm{Span}(W))^\perp$,
- $W^\perp = \bigcap_{\underline{w}\in U} \underline{w}^\perp$,
- se $\underline{w_1}$, ..., $\underline{w_k}$ generano W, allora $W^\perp = \bigcap_{i=1}^k \underline{w_i}^\perp$,
- $W^\perp = \alpha_\varphi^{-1}(\mathrm{Ann}(W))$,
- $\alpha_\varphi(W^\perp) = \mathrm{Ann}(W) \cap \mathrm{Im}\, \alpha_\varphi$,
- $V \subseteq W \implies W^\perp \subseteq V^\perp$,
- $(U+W)^\perp = U^\perp \cap W^\perp$,
- $((W^\perp)^\perp)^\perp = (W + V^\perp)^\perp = W^\perp \cap (V^\perp)^\perp = W^\perp \cap V = W^\perp$,
- $(U \cap W)^\perp \supseteq U^\perp + W^\perp$,
- $(U \cap W)^\perp = U^\perp + W^\perp$, se φ è non degenere,
- φ è definito $\iff \mathrm{CI}(\varphi) = \{\underline{0}\}$,
- φ è (semi)definito $\implies$ ogni sua restrizione è (semi)definita,
- φ è semidefinito $\iff \mathrm{CI}(\varphi) = V^\perp = \mathrm{Rad}(\varphi)$ (considera l'esistenza di due vettori $\underline{v}$, $\underline{w} \in V$ con forme quadratiche discordi, osserva che sono linearmente indipendenti e trova un $\lambda \in \mathbb{K}$ tale per cui $\underline{v} + \lambda\underline{w}$ crea un assurdo),
- $\mathrm{Im}(\alpha_\varphi) \subseteq \mathrm{Ann}(V^\perp)$ (se V è di dimensione infinita),
- $\mathrm{Im}(\alpha_\varphi) = \mathrm{Ann}(V^\perp)$ (se V è di dimensione finita),
- $\mathrm{Rad}(\varphi|_U) = U^\perp \cap U$,
- $\mathrm{CI}(\varphi|_U) = \mathrm{CI}(\varphi) \cap U$,

Se U è un sottospazio di V, φ induce un prodotto scalare $\hat\varphi : V/U \times V/U \to \mathbb{K}$ tale che $\hat\varphi([\underline{v_1}]_U, [\underline{v_2}]_U) = \varphi(\underline{v_1}, \underline{v_2})$ se e solo se $U \subseteq V^\perp$. In particolare, se $U = V^\perp$, $\hat\varphi$ è anche non degenere.

Due esempi classici di prodotto scalare sono $\varphi(A,B) = \mathrm{tr}(AB)$ e $\psi(A,B) = \mathrm{tr}(AB^\top)$, entrambi su $M(n,\mathbb{K})$. I due prodotti sono entrambi non degeneri, e vale che:

- $\varphi|_{\mathrm{Sym}(n,\mathbb{K})} = \psi|_{\mathrm{Sym}(n,\mathbb{K})}$,
- $\varphi|_{\Lambda(n,\mathbb{K})} = -\psi|_{\Lambda(n,\mathbb{K})}$,
- se $\mathrm{char}\,\mathbb{K} \neq 2$, $V = \mathrm{Sym}(n,\mathbb{K}) \oplus^\perp \Lambda(n,\mathbb{K})$, per ambo i prodotti scalari.

Se $\mathcal{B}'$ è un'altra base di V, vale il seguente *teorema di cambiamento di base*:

$$M_{\mathcal{B}'}(\varphi) = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{Id}_V)^\top\, M_{\mathcal{B}}(\varphi)\, M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{Id}_V).$$

Si definisce relazione di congruenza la relazione di equivalenza $\cong$ (o $\equiv$) definita su $\mathrm{Sym}(n,\mathbb{K})$ nel seguente modo:

$$A \cong B \iff \exists\, P \in \mathrm{GL}(n,\mathbb{K}) \mid A = P^\top B P.$$

- $A \cong B \implies \mathrm{rg}(A) = \mathrm{rg}(B)$ (il rango è invariante per congruenza; e dunque si può definire $\mathrm{rg}(\varphi)$ come il rango di una qualsiasi matrice associata a φ),
- $A \cong B \implies \det(A)\det(B) \geq 0$ (in $\mathbb{K} = \mathbb{R}$ il segno del determinante è invariante per congruenza),
- Due matrici associate a φ in basi diverse sono congruenti per la formula di cambiamento di base.

Si definiscono i seguenti tre indici per $\mathbb{K} = \mathbb{R}$:

- $\iota_+ = \max\{\dim W \mid W \subseteq V \text{ e } \varphi|_W > 0\}$,
- $\iota_- = \max\{\dim W \mid W \subseteq V \text{ e } \varphi|_W < 0\}$,
- $\iota_0 = \dim V^\perp$,

e si definisce segnatura di φ la terna $\sigma = (\iota_+, \iota_-, \iota_0)$.

Si dice che una base $\mathcal{B}$ di V è ortogonale se i suoi vettori sono a due a due ortogonali (e quindi la matrice associata in tale base è diagonale). Se $\mathrm{char}\,\mathbb{K} \neq 2$, valgono i seguenti risultati:

- $\varphi(\underline{v},\underline{w}) = \frac{q(\underline{v}+\underline{w}) - q(\underline{v}) - q(\underline{w})}{2}$ (formula di polarizzazione; φ è completamente determinata dalla sua forma quadratica),
- Esiste sempre una base ortogonale $\mathcal{B}$ di V (teorema di Lagrange; è sufficiente considerare l'esistenza di un vettore anisotropo $\underline{w} \in V$ ed osservare che $V = W \oplus^\perp W^\perp$, dove $W = \mathrm{Span}(\underline{w})$, concludendo per induzione; o in caso di non esistenza di tale $\underline{w}$, concludere per il risultato precedente),
- (se $\mathbb{K} = \mathbb{C}$) Esiste sempre una base ortogonale $\mathcal{B}$ di V tale che:

$$M_{\mathcal{B}}(\varphi) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

  dove $r = \mathrm{rg}(\varphi)$ (teorema di Sylvester, caso complesso; si consideri una base ortogonale e se ne normalizzino i vettori anisotropi),
- Due matrici simmetriche ad elementi complessi con stesso rango allora non solo sono SD-equivalenti, ma sono anche congruenti,
- (se $\mathbb{K} = \mathbb{R}$) Esiste sempre una base ortogonale $\mathcal{B}$ di V tale che:

$$M_{\mathcal{B}}(\varphi) = \left(\begin{array}{c|c|c} I_{\iota_+} & 0 & 0 \\ \hline 0 & -I_{\iota_-} & 0 \\ \hline 0 & 0 & 0 \cdot I_{\iota_0} \end{array}\right).$$

  Inoltre σ è un invariante completo per la congruenza, e vale che, su una qualsiasi base ortogonale $\mathcal{B}'$ di V, ι_+ è esattamente il numero di vettori anisotropi di base con forma quadratica positiva, che ι_- è il numero di vettori con forma negativa e che ι_0 è il numero di vettori isotropi (teorema di Sylvester, caso reale; si consideri una base ortogonale e se ne normalizzino i vettori anisotropi, facendo infine eventuali considerazioni dimensionali per dimostrare la seconda parte dell'enunciato),
- $\varphi > 0 \iff \sigma = (n,0,0)$,
- $\varphi < 0 \iff \sigma = (0,n,0)$,
- $\varphi \geq 0 \iff \sigma = (n-k,0,k)$,
- $\varphi \leq 0 \iff \sigma = (0,n-k,k)$, con $0 \leq k \leq n$ tale che $k = \dim V^\perp$,
- I vettori isotropi di una base ortogonale sono una base di $V^\perp$,
- $\mathrm{rg}(\varphi) = \iota_+ + \iota_-$,
- $n = \iota_+ + \iota_- + \iota_0$,
- Se W è un sottospazio di V, $\iota_+(\varphi) \geq \iota_+(\varphi|_W)$ e $\iota_-(\varphi) \geq \iota_-(\varphi|_W)$,
- Se $V = U \oplus^\perp W$, $\sigma(\varphi) = \sigma(\varphi|_U) + \sigma(\varphi|_W)$,
- Se $\mathbb{K} = \mathbb{R}$ e $A = M_{\mathcal{B}}(\varphi)$, allora:

$$\sigma = \left(\sum_{\substack{\lambda\in\mathrm{sp}(\varphi)\\ \lambda>0}} \mu_a(\lambda),\ \sum_{\substack{\lambda\in\mathrm{sp}(A)\\ \lambda<0}} \mu_a(\lambda),\ \mu_0(\lambda)\right),$$

  come conseguenza del teorema spettrale reale.

Si chiama matrice di Sylvester una matrice della forma vista nell'enunciato del teorema di Sylvester reale, e si dice che una base $\mathcal{B}$ è una base di Sylvester se la matrice ad essa associata è di Sylvester. Per il teorema di Sylvester, tale base esiste sempre, e la matrice di Sylvester è unica per ogni prodotto scalare φ.

Se $M \in \mathrm{Sym}(2,\mathbb{R})$, $\det(M) < 0 \iff \sigma(M) = (1,1,0)$ (e dunque se e solo se M rappresenta un piano iperbolico). Al contrario $\det(M) > 0$ se e solo se M è definita (e in tal caso è definita positiva se il suo primo elemento è positivo, e negativa se è negativo). Se $\mathbb{K} = \mathbb{R}$, $q(\underline{v}) > 0$ e $q(\underline{w}) < 0$, allora $\underline{v}$ e $\underline{w}$ sono linearmente indipendenti; in particolare $\mathrm{Span}(\underline{v},\underline{w})$ è un piano iperbolico ed esistono λ_1, $\lambda_2 \in \mathbb{R}$ tali per cui $\lambda_1\underline{v} + \lambda_2\underline{w}$ è isotropo.

Prodotto hermitiano

Sia V un $\mathbb{C}$-spazio. Allora una mappa $\varphi : V \times V \to \mathbb{C}$ si dice prodotto hermitiano (e quindi si dice che $\varphi \in \mathrm{PH}(V)$, l'$\mathbb{R}$-spazio dei prodotti hermitiani[2]) se è una forma sesquilineare, ossia se è antilineare nel primo argomento ed è lineare nel secondo[3], e se il coniugio applicato a φ ne inverte gli argomenti. In particolare φ è un prodotto hermitiano se:

(i) $\varphi(\underline{v}, \lambda\underline{u} + \underline{w}) = \lambda\varphi(\underline{v},\underline{u}) + \varphi(\underline{v},\underline{w})$, $\forall\, \underline{v},\, \underline{u},\, \underline{w} \in V,\ \lambda \in \mathbb{K}$,

(ii) $\overline{\varphi(\underline{v},\underline{w})} = \varphi(\underline{w},\underline{v})$.

Un prodotto hermitiano φ si comporta pressoché come un prodotto scalare su $\mathbb{R}$ (le definizioni principali sono infatti le medesime). Se $\mathcal{B}$ è una base di V, la matrice associata $M_{\mathcal{B}}(\varphi)$ è definita in modo tale che $M_{\mathcal{B}}(\varphi)_{ij} = \varphi(\underline{v_i},\underline{v_j})$. Infatti tale prodotto soddisfa le seguenti proprietà:

- $\varphi(\lambda\underline{v} + \underline{w}, \underline{u}) = \overline{\lambda}\varphi(\underline{v},\underline{u}) + \varphi(\underline{w},\underline{u})$, $\forall\, \underline{v},\, \underline{w},\, \underline{u} \in V,\ \lambda \in \mathbb{C}$,
- $V^\perp = []_{\mathcal{B}}^{-1}(\mathrm{Ker}\, M_{\mathcal{B}}(\varphi))$,
- φ è non degenere se e solo se $\mathrm{Ker}\, M_{\mathcal{B}}(\varphi) = \{\underline{0}\}$,
- $\dim W + \dim W^\perp = \dim V + \dim(W \cap V^\perp)$ (formula delle dimensioni),
- $\varphi(\underline{v},\underline{w}) = [\underline{v}]_{\mathcal{B}}^* M_{\mathcal{B}}(\varphi)[\underline{w}]_{\mathcal{B}}$,
- $M_{\mathcal{B}'}(\varphi) = \left(M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{Id}_V)\right)^* M_{\mathcal{B}}(\varphi)\, M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{Id}_V)$ (formula del cambiamento di base),
- si può definire una relazione di equivalenza analoga alla congruenza:
  $A \sim_* B \overset{\text{def}}{\iff} \exists\, M \in \mathrm{GL}(n,\mathbb{C}) \mid A = M^* B M$,
- φ è completamente determinato dalla sua forma quadratica q secondo le seguenti due formule di polarizzazione:
  - $q(\underline{v} + \underline{w}) - q(\underline{v}) - q(\underline{w}) = 2\Re(\varphi(\underline{v},\underline{w}))$,
  - $q(\underline{v} + i\underline{w}) - q(\underline{v}) - q(\underline{w}) = 2i\Im(\varphi(\underline{v},\underline{w}))$,
- esiste sempre una base ortogonale per φ (*teorema di Lagrange*),
- vale il teorema di Sylvester reale e la segnatura in senso hermitiano è un invariante per la relazione $\sim_*$,
- $\varphi > 0 \iff \sigma(\varphi) = (n,0,0)$,
- $\varphi < 0 \iff \sigma(\varphi) = (0,n,0)$,
- $\varphi \geq 0 \iff \sigma(\varphi) = (n-k,0,k)$, dove $k = \dim V^\perp = \dim \mathrm{Ker}\, M_{\mathcal{B}}(\varphi)$,
- $\varphi \leq 0 \iff \sigma(\varphi) = (0,n-k,k)$, dove $k = \dim V^\perp = \dim \mathrm{Ker}\, M_{\mathcal{B}}(\varphi)$.

Esiste un unico modo per complessificare un prodotto scalare φ, ossia esiste un unico prodotto hermitiano $\varphi_{\mathbb{C}}$ tale per cui $\varphi_{\mathbb{C}}(\underline{v},\underline{w}) = \varphi(\underline{v},\underline{w})$ se $\underline{v}$, $\underline{w}$ sono vettori della parte reale dello spazio complessificato. In particolare $\varphi_{\mathbb{C}}$ è determinato dalla seguente formula:

$$\varphi_{\mathbb{C}}(\underline{v_1} + i\underline{v_2}, \underline{w_1} + i\underline{w_2}) = \varphi(\underline{v_1},\underline{w_1}) + \varphi(\underline{v_2},\underline{w_2}) + i(\varphi(\underline{v_1},\underline{w_2}) - \varphi(\underline{v_2},\underline{w_1})).$$

Funzionali rappresentabili

Un funzionale $f \in V^*$ si dice rappresentabile tramite φ se $f \in \mathrm{Im}\,\alpha_\varphi$, ossia se $\exists\, \underline{v} \in V \mid f = \varphi(\underline{v},\cdot)$. Dal momento che $\mathrm{Im}(\alpha_\varphi) = \mathrm{Ann}(V^\perp)$ (l'inclusione verso destra è facile da dimostrare e l'uguaglianza è data dall'uguaglianza dimensione), f è rappresentabile se e solo se $V^\perp \subseteq \mathrm{Ker}\, f$.

Se φ è non degenere, ogni funzionale f è rappresentabile in modo unico (*teorema di rappresentazione di Riesz*; infatti α_φ sarebbe in tal caso un isomorfismo). In particolare, se φ è un prodotto scalare, tale vettore $\underline{v}$, data una base ortogonale $\mathcal{B} = \{\underline{v_1},\dots,\underline{v_n}\}$, è determinato nel seguente modo:

$$\underline{v} = \sum_{i=1}^{n} \frac{f(\underline{v_i})}{\varphi(\underline{v_i},\underline{v_i})}\, \underline{v_i}.$$

Se invece φ è un prodotto hermitiano, tale vettore $\underline{v}$ si determina nel seguente altro modo:

$$\underline{v} = \sum_{i=1}^{n} \overline{\left(\frac{f(\underline{v_i})}{\varphi(\underline{v_i},\underline{v_i})}\right)}\, \underline{v_i}.$$

In generale, f è rappresentabile se e solo se, scelta una base $\mathcal{B}$ di V, il sistema $M_{\mathcal{B}}(\varphi)\underline{x} = [f]_{\mathcal{B}^*}$ è risolvibile.

Se W è un supplementare di $V^\perp$, e dunque $V = W \oplus V^\perp$, allora $\varphi|_W$ è non degenere, e dunque $\alpha_\varphi|_W : W \to \mathrm{Im}\,\alpha_\varphi$ è un isomorfismo da W a $\mathrm{Im}\,\alpha_\varphi$ (quindi se f è rappresentabile, lo è tramite un unico vettore di W).

In particolare, se $\underline{v}$ rappresenta f, allora $\mathrm{Ker}\, f = \underline{v}^\perp$; da cui segue che $(\mathrm{Ker}\, f)^\perp = \underline{v}^{\perp\perp} = \mathrm{Span}(\underline{v}) + V^\perp$. Se f non è l'applicazione nulla, $\underline{v} \notin V^\perp$, e quindi $\mathrm{Span}(\underline{v}) \cap V^\perp = \{\underline{0}\} \implies (\mathrm{Ker}\, f)^\perp = \mathrm{Span}(\underline{v}) \oplus^\perp V^\perp$. Quindi, per computare un vettore $\underline{v_0}$ che rappresenti f è sufficiente prendere un supplementare $\mathrm{Span}(\underline{u})$ di $V^\perp$ in $(\mathrm{Ker}\, f)^\perp$ (infatti l'aggiunta di un vettore di $V^\perp$ non varierebbe l'immagine secondo α_φ) e computare $\lambda \in \mathbb{K} \mid \underline{v_0} = \lambda\underline{u}$ nel seguente modo:

$$\lambda = \frac{\varphi(\lambda\underline{u},\underline{w})}{\varphi(\underline{u},\underline{w})} = \frac{f(\underline{w})}{\varphi(\underline{u},\underline{w})},$$

dove $\underline{w} \notin \mathrm{Ker}\, f$.

Algoritmo di ortogonalizzazione di Gram-Schmidt

Data una base $\mathcal{B}$ di V, se $|\mathrm{CI}(\varphi) \cap \mathcal{B}| \leq 1$ (ossia se ogni vettore di $\mathcal{B}$ è anisotropo o al più vi è un vettore isotropo, posto in fondo come $\underline{v_n}$), si può trovare una base ortogonale $\mathcal{B}' = \{\underline{v_1}',\dots,\underline{v_n}'\}$ a partire da $\mathcal{B}$ tale che ne mantenga la stessa bandiera, ossia tale che:

$$\mathrm{Span}(\underline{v_1}',\dots,\underline{v_i}') = \mathrm{Span}(\underline{v_1},\dots,\underline{v_i})\ \forall\, 1 \leq i \leq n.$$

Si definisce $C(\underline{w},\underline{v}) = \frac{\varphi(\underline{v},\underline{w})}{\varphi(\underline{w},\underline{w})}$ come il coefficiente di Fourier di $\underline{v}$ rispetto a $\underline{w}$. L'algoritmo allora funziona nel seguente modo:

1. Si prenda in considerazione $\underline{v_1}$ e si sottragga ad ogni altro vettore $\underline{v_i}$ della base il vettore $C(\underline{v_1},\underline{v_i})\,\underline{v_1}$,
2. Si ripeta il processo considerando come $\mathcal{B}$ tutti i vettori di $\mathcal{B}$ con $\underline{v_1}$ escluso, o si termini l'algoritmo una volta che è rimasto un solo vettore.

Dal momento che l'algoritmo mantiene invariata la bandiera della base, una matrice triangolabile è anche ortogonalmente triangolabile se la base non contiene alcun (o contiene al più un) vettore isotropo secondo un certo prodotto scalare.

Metodo di Jacobi per il calcolo della segnatura

Sia $A = M_{\mathcal{B}}(\varphi)$ una matrice associata a φ nella base $\mathcal{B}$. Sia $d_0 := 1$. Se $d_i = \det(A^{1,\dots,i}_{1,\dots,i})$ (è possibile anche prendere un'altra sequenza di minori, a patto che essi siano principali e che siano crescenti per inclusione) è diverso da zero per ogni $1 \leq i \leq n-1$, allora ι_+ è il numero di permanenze di segno di d_i (zero escluso), ι_- è il numero di variazioni di segno (zero escluso), e ι_0 è 1 se $d_n = 0$ o 0 altrimenti.

In generale, se W è un sottospazio di W', W ha codimensione 1 rispetto a W' e $\det(M_{\mathcal{B}_W}(\varphi|_W)) \neq 0$ per una base $\mathcal{B}_W$ di W, allora la segnatura di $\varphi|_{W'}$ è la stessa di $\varphi|_W$, dove si aggiunge 1 a ι_+, se i determinanti $\det(M_{\mathcal{B}_W}(\varphi|_W))$ e $\det(M_{\mathcal{B}_{W'}}(\varphi|_W))$ (dove $\mathcal{B}_{W'}$ è una base di W') concordano di segno, 1 a ι_-, se sono discordi, o 1 a ι_0 se l'ultimo di questi due determinanti è nullo.

Dal metodo di Jacobi si deduce il criterio di definitezza di Sylvester: A è definita positiva se e solo se $d_i > 0\ \forall\, 1 \leq i \leq n$; A è definita negativa se e solo se $(-1)^i d_i > 0\ \forall\, 1 \leq i \leq n$.

[2] Infatti, se $\lambda \in \mathbb{C} \setminus \mathbb{R}$ e $\varphi \in \mathrm{PH}(V)$, $\lambda\varphi$ non è un prodotto hermitiano, mancando della proprietà del coniugio.

[3] In realtà questa convenzione è spesso e volentieri implementata nelle ricerche di Fisica, mentre in Matematica si tende in realtà a mettere l'antilinearità nel secondo argomento. Il corso ha comunque implementato la prima delle due convenzioni, e così si è riportato in queste schede la convenzione scelta.

Sottospazi isotropi e indice di Witt

Si dice che un sottospazio W di V è isotropo se $\varphi|_W = 0$, o equivalentemente se $W \subseteq W^\perp$ (i.e. se $W \cap W^\perp = W$, e quindi se $\mathrm{Rad}(\varphi|_W) = W$). Si definisce allora l'indice di Witt $W(\varphi)$ come la dimensione massima di un sottospazio isotropo di V.

- $V^\perp$ è un sottospazio isotropo,
- Se W è isotropo, allora $\dim W \leq \lfloor \frac{\dim V + \dim \mathrm{Rad}(\varphi)}{2} \rfloor$,
- Se W è isotropo e φ è non degenere, allora $\dim W \leq \lfloor \frac{1}{2} \dim V \rfloor$,
- Se $\mathbb{K} = \mathbb{R}$, allora $W(\varphi) = \min\{i_+, i_-\} + i_0$ (è sufficiente considerare una base di Sylvester e costruire un nuovo insieme linearmente indipendente $\mathcal{B}_W$ i cui i vettori sono o isotropi o della forma $\underline{v_i} - \underline{w_i}$, dove $q(\underline{v_i}) = 1$ e $q(\underline{w_i}) = 1$, mostrando che $W = \mathrm{Span}(\mathcal{B}_W)$ è isotropo, concludendo con discussioni dimensionali),
- Se $\mathbb{K} = \mathbb{C}$, allora $W(\varphi) = \lfloor \frac{\dim V + \dim V^\perp}{2} \rfloor$ (è sufficiente considerare una base di Sylvester per φ, costruire un nuovo insieme linearmente indipendente $\mathcal{B}_W$ prendendo quante più coppie $(\underline{v_i}, \underline{v_j})$ possibili di vettori della base non isotropi poi associate al vettore $\underline{v_i} + i\underline{v_j}$, mostrando infine che $W = \mathrm{Span}(\mathcal{B}_W)$ è isotropo e che è sicuramente massimale perché realizza la dimensione massima possibile secondo le precedenti proposizioni),
- Se $\mathbb{K} = \mathbb{R}$ e φ è definito, allora $W(\varphi) = 0$,
- Se $\mathbb{K} = \mathbb{R}$ e φ è semidefinito, allora $W(\varphi) = i_0$ (e $W = V^\perp$ è un sottospazio isotropo di tale dimensione).

Isometrie tra spazi vettoriali

Due spazi vettoriali (V, φ) e (W, ψ) su $\mathbb{K}$ si dicono isometrici tra loro se esiste un isomorfismo $f : V \to W$ tale che $\varphi(\underline{v_1}, \underline{v_2}) = \psi(f(\underline{v_1}), f(\underline{v_2}))$.
Se f è un isomorfismo tra V e W, sono equivalenti le seguenti affermazioni:

(i) (V, φ) e (W, ψ) sono isometrici tra loro tramite f,
(ii) $\forall\, \mathcal{B}$ base di V, $M_{\mathcal{B}}(\varphi) = M_{f(\mathcal{B})}(\psi)$,
(iii) $\exists\, \mathcal{B}$ base di V, $M_{\mathcal{B}}(\varphi) = M_{f(\mathcal{B})}(\psi)$.

Inoltre, V e W sono isometrici se e solo se hanno la stessa dimensione e le matrici associate a φ e ψ in due basi di V e di W sono congruenti (infatti, in tal caso, esistono due basi di V e di W che condividono la stessa matrice associata, ed è possibile associare ad uno ad uno gli elementi di queste basi).

Pertanto, se $\mathcal{B}_V$ e $\mathcal{B}_W$ sono due basi di V e di W, $\mathbb{K} = \mathbb{R}$ e $M_{\mathcal{B}_V}(\varphi)$ e $M_{\mathcal{B}_W}(\psi)$ condividono la stessa segnatura, allora V e W sono isometrici tra loro (come conseguenza del teorema di Sylvester reale).

Analogamente, se $\mathbb{K} = \mathbb{C}$ e $M_{\mathcal{B}_V}(\varphi)$ e $M_{\mathcal{B}_W}(\psi)$ condividono lo stesso rango, allora V e W sono isometrici tra loro (come conseguenza stavolta del teorema di Sylvester complesso).

Trasposta e aggiunta di un'applicazione

Sia (V, φ) uno spazio dotato di un prodotto φ non degenere. Allora si definisce $f^* \in \mathrm{End}(V)$ (talvolta indicato come $f^\top$ se φ non è hermitiano, quando è chiaro che non ci stia riferendo alla trasposizione dell'operatore f) come l'unico operatore tale per cui $\varphi(f^*(\underline{v}), \underline{w}) = \varphi(\underline{v}, f(\underline{w}))$. In particolare, se φ non è hermitiano, tale operatore soddisfa la seguente relazione:

$$\alpha_\varphi \circ f^* = f^\top \circ \alpha_\varphi,$$

dove con $f^\top$ si indica l'applicazione trasposta di f. Scelta allora una base $\mathcal{B}$ di V, sempre se φ non è hermitiano, si può scrivere in relazione a $M_{\mathcal{B}}(f)$ la matrice associata a f^*:

$$M_{\mathcal{B}}(f^*) = M_{\mathcal{B}}(\varphi)^{-1} M_{\mathcal{B}}(f)^\top M_{\mathcal{B}}(\varphi).$$

Se invece φ è hermitiano, vale la seguente relazione:

$$M_{\mathcal{B}}(f^*) = M_{\mathcal{B}}(\varphi)^{-1} M_{\mathcal{B}}(f)^* M_{\mathcal{B}}(\varphi).$$

Se φ è un prodotto scalare, $f^* = f^\top$ si chiama *trasposto* di f, mentre se φ è hermitiano f^* si dice *aggiunto* di f.

D'ora in poi si intenderà con $f^\top$ il trasposto di f (con φ scalare) e con f^* l'aggiunto di f (con φ hermitiano).

Un'operatore f si dice *simmetrico* se $f = f^\top$ e quindi se $\varphi(\underline{v}, f(\underline{w})) = \varphi(f(\underline{v}), \underline{w})$ (analogamente un'operatore si dice *hermitiano* se $f = f^*$).
Un'operatore f si dice *ortogonale* se è un'isometria da (V, φ) in (V, φ), ossia se e solo se $\varphi(\underline{v}, \underline{w}) = \varphi(f(\underline{v}), f(\underline{w}))$ (analogamente un'operatore si dice *unitario* se è un'isometria con φ prodotto hermitiano).
Sia $\mathcal{B}$ una base di V.

- se $\mathcal{B}$ è una base ortonormale, $M_{\mathcal{B}}(f^\top) = M_{\mathcal{B}}(f)^\top$ (infatti in tal caso $M_{\mathcal{B}}(\varphi) = I_n$),
- se $\mathcal{B}$ è una base ortonormale, $M_{\mathcal{B}}(f^*) = M_{\mathcal{B}}(f)^*$ (come sopra),
- f è simmetrico $\iff f = f^\top \iff$ $M_{\mathcal{B}}(\varphi)^{-1} M_{\mathcal{B}}(f)^\top M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(f^\top) = M_{\mathcal{B}}(f) \iff$ $M_{\mathcal{B}}(\varphi) M_{\mathcal{B}}(f)$ è simmetrica,
- se $\mathcal{B}$ è una base ortonormale, f è simmetrico $\iff$ $f = f^\top \iff M_{\mathcal{B}}(f) = M_{\mathcal{B}}(f)^\top \iff M_{\mathcal{B}}(f)$ è simmetrica,
- f è hermitiano $\iff f = f^* \iff$ $M_{\mathcal{B}}(\varphi)^{-1} M_{\mathcal{B}}(f)^* M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(f^*) = M_{\mathcal{B}}(f) \iff$ $M_{\mathcal{B}}(\varphi) M_{\mathcal{B}}(f)$ è hermitiana,
- se $\mathcal{B}$ è una base ortonormale, f è hermitiana $\iff$ $f = f^* \iff M_{\mathcal{B}}(f) = M_{\mathcal{B}}(f)^* \iff M_{\mathcal{B}}(f)$ è hermitiana,
- f è ortogonale $\iff f \circ f^\top = f^\top \circ f = \mathrm{Id}_V$,
- se $\mathcal{B}$ è una base ortonormale, f è ortogonale $\iff$ $M_{\mathcal{B}}(f)$ è ortogonale,
- f è unitario $\iff f \circ f^* = f^* \circ f = \mathrm{Id}_V$,
- se $\mathcal{B}$ è una base ortonormale, f è unitaria $\iff M_{\mathcal{B}}(f)$ è unitaria,
- $(f^\top)^\top = f$,
- $(f^*)^* = f$,
- $(\lambda f)^\top = \lambda f^\top$,
- $(\lambda f)^* = \overline{\lambda} f^*$,
- $(f + g)^\top = f^\top + g^\top$,
- $(f + g)^* = f^* + g^*$,
- $(f \circ g)^\top = g^\top \circ f^\top$,
- $(f \circ g)^* = g^* \circ f^*$,
- se f è invertibile, $(f^\top)^{-1} = (f^{-1})^\top$ (è sufficiente mostrare che $\varphi((f^\top \circ (f^{-1})^\top)(\underline{v}), \underline{w}) = \varphi(\underline{v}, \underline{w})$ e dedurre, sottraendo i due membri, che deve valere $f^\top \circ (f^{-1})^\top = \mathrm{Id}_V$),
- se f è invertibile, $(f^*)^{-1} = (f^{-1})^*$ (come sopra),
- $\mathrm{Ker}\, f^\top = (\mathrm{Im}\, f)^\perp$,
- $\mathrm{Ker}\, f^* = (\mathrm{Im}\, f)^\perp$,
- $\mathrm{Im}\, f^\top = (\mathrm{Ker}\, f)^\perp$,
- $\mathrm{Im}\, f^* = (\mathrm{Ker}\, f)^\perp$,
- se W è un sottospazio di V, allora W è f-invariante se e solo se $W^\perp$ è $f^\top$-invariante (o f^*-invariante),
- se f è simmetrico (o hermitiano), allora W è f-invariante se e solo se $W^\perp$ è f-invariante,
- l'operatore $\top \in \mathrm{End}(\mathrm{End}(V))$ è sempre diagonalizzabile e ha spettro $\{\pm 1\}$, dal momento che il suo polinomio minimo divide $x^2 - 1$ (infatti $(f^\top)^\top = f$),
- l'autospazio V_1 di $\top$ raccoglie gli operatori simmetrici, mentre V_{-1} raccoglie gli operatori antisimmetrici.

Operatori normali

Un operatore f in uno spazio euclideo reale si dice normale se commuta col suo trasposto, ossia se $f \circ f^\top = f^\top \circ f$, mentre si dice normale in uno spazio euclideo complesso se commuta col suo aggiunto, ossia se $f \circ f^* = f^* \circ f$.

Analogamente una matrice si dice normale se commuta con la sua trasposta (se è a elementi reali) o con la sua aggiunta (se è a elementi complessi). Una matrice contemporaneamente normale e triangolare è necessariamente una matrice diagonale.

In uno spazio euclideo complesso, f è normale $\iff$ f è unitariamente diagonalizzabile (f è triangolarizzabile con una base ortonormale $\mathcal{B}$ tramite l'algoritmo di ortogonalizzazione di Gram-Schmidt, e quindi la matrice $M_{\mathcal{B}}(f)$ è sia normale che triangolare, e quindi diagonale). In uno spazio euclideo reale, se f è triangolarizzabile e normale, allora f è diagonalizzabile (come prima).

Spazi euclidei reali e complessi

Si dice che (V, φ) è uno spazio euclideo reale se V è un $\mathbb{R}$-spazio e se φ è un prodotto scalare definito positivo. Si dice che $(V_{\mathbb{C}}, \varphi_{\mathbb{C}})$ è uno spazio euclideo complesso se $V_{\mathbb{C}}$ è un $\mathbb{C}$-spazio e se $\varphi_{\mathbb{C}}$ è un prodotto hermitiano definito positivo.

Questi due tipi di spazi hanno in comune alcune proprietà particolari. Si definisce innanzitutto la norma euclidea per uno spazio euclideo (V, φ) come:

$$\|\underline{v}\| = \sqrt{q(\underline{v})} = \sqrt{\varphi(\underline{v}, \underline{v})}.$$

Tale norma soddisfa alcune proprietà:

- $\|\lambda \underline{v}\| = |\lambda| \, \|\underline{v}\|$,
- $\|\underline{v}\| \, \|\underline{w}\| \geq |\varphi(\underline{v}, \underline{w})|$ (disuguaglianza di Cauchy-Schwarz),
- $\|\underline{v} + \underline{w}\| \leq \|\underline{v}\| + \|\underline{w}\|$ (disuguaglianza triangolare).

Su questi due spazi possono essere definiti due particolare operatori: la proiezione ortogonale e l'inversione ortogonale.

Si definisce proiezione ortogonale su un sottospazio $W \neq \{\underline{0}\}$ l'operatore $\mathrm{pr}_W \in \mathrm{End}(V)$ tale che $\mathrm{pr}_W(\underline{v}) = \underline{w}$, dove $\underline{v} = \underline{w} + \underline{w}^\perp$, con $\underline{w} \in W$ e $\underline{w}^\perp \in W^\perp$. Tale decomposizione è ben definita e unica dacché $V = W \oplus^\perp W^\perp$ (infatti φ è definita positiva). Una proiezione ortogonale soddisfa la relazione $\mathrm{pr}_W^2 = \mathrm{pr}_W$, da cui si ricava che $\varphi_{\mathrm{pr}_W} \mid x(x-1)$ (implicandone la diagonalizzabilità). Infatti $V_1 = \mathrm{Ker}(\mathrm{pr}_W - \mathrm{Id}_V) = W$ e $V_0 = \mathrm{Ker}(\mathrm{pr}_W) = W^\perp$ (per cui $\varphi_{\mathrm{pr}_W}(x) = x(x-1)$). La proiezione ortogonale è un operatore simmetrico (se lo spazio è euclideo reale) o hermitiano (se lo spazio è euclideo complesso); infatti vale che $\varphi(\mathrm{pr}_W(\underline{v}), \underline{w}) = \varphi(\mathrm{pr}_W(\underline{v}), \mathrm{pr}_W(\underline{w}) + \mathrm{pr}_{W^\perp}(\underline{w})) = \varphi(\mathrm{pr}_W(\underline{v}), \mathrm{pr}_W(\underline{w})) = \varphi(\mathrm{pr}_W(\underline{v}) + \mathrm{pr}_{W^\perp}(\underline{v}), \mathrm{pr}_W(\underline{w})) = \varphi(\underline{v}, \mathrm{pr}_W(\underline{w}))$.

Si definisce inversione ortogonale su un sottospazio $W \neq \{\underline{0}\}$ l'operatore $\rho_W \in \mathrm{End}(V)$ tale che $\rho_W(\underline{v}) = \underline{w} - \underline{w}^\perp$, dove $\underline{v} = \underline{w} + \underline{w}^\perp$, con $\underline{w} \in W$ e $\underline{w}^\perp \in W^\perp$. Come prima, tale decomposizione è unica e ben definita. Un'inversione ortogonale soddisfa la relazione $\rho_W^2 = \mathrm{Id}_V$, da cui si ricava che $\varphi_{\rho_W} \mid (x+1)(x-1)$ (implicandone la diagonalizzabilità). Infatti $V_1 = \mathrm{Ker}(\rho_W - \mathrm{Id}_V) = W$ e $V_{-1} = \mathrm{Ker}(\rho_W + \mathrm{Id}_V) = W^\perp$ (per cui $\varphi_{\rho_W}(x) = (x+1)(x-1)$). Se $\dim W = \dim V - 1$, allora si dice che l'inversione ortogonale è una riflessione ortogonale. L'inversione ortogonale è sempre un operatore ortogonale (se lo spazio euclideo è reale) o unitario (se lo spazio euclideo è complesso); infatti vale che $\varphi(\underline{v}, \underline{w}) = \varphi(\mathrm{pr}_W(\underline{v}) + \mathrm{pr}_{W^\perp}(\underline{v}), \mathrm{pr}_W(\underline{w}) + \mathrm{pr}_{W^\perp}(\underline{w})) = \varphi(\mathrm{pr}_W(\underline{v}), \mathrm{pr}_W(\underline{w})) + \varphi(\mathrm{pr}_{W^\perp}(\underline{v}), \mathrm{pr}_{W^\perp}(\underline{w})) = \varphi(\mathrm{pr}_W(\underline{v}), \mathrm{pr}_W(\underline{w})) + \varphi(-\mathrm{pr}_{W^\perp}(\underline{v}), -\mathrm{pr}_{W^\perp}(\underline{w})) = \varphi(\mathrm{pr}_W(\underline{v}) - \mathrm{pr}_{W^\perp}(\underline{v}), \mathrm{pr}_W(\underline{w}) - \mathrm{pr}_{W^\perp}(\underline{w})) = \varphi(\rho_W(\underline{v}), \rho_W(\underline{w}))$.

Teorema spettrale reale e complesso

Sia (V, φ) uno spazio euclideo reale. Se f è un operatore simmetrico, allora f ammette solo autovalori reali. Prendendo infatti il prodotto hermitiano complessificato di φ, allora, se λ è un autovalore in $\mathbb{C}$ di f, $\overline{\lambda}\varphi(\underline{v}, \underline{v}) = \varphi(\lambda \underline{v}, \underline{v})$ $= \varphi(f(\underline{v}), \underline{v}) = \varphi(\underline{v}, f(\underline{v})) = \varphi(\underline{v}, \lambda \underline{v}) = \lambda \varphi(\underline{v}, \underline{v})$, dove $\underline{v}$ è un autovettore non nullo relativo a λ; pertanto $\lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}$ dacché $\varphi(\underline{v}, \underline{v}) \neq 0$). Seguendo gli stessi passaggi algebrici si mostra che se f è un operatore hermitiano uno spazio euclideo complesso, questo ammette solo autovalori reali.

Se f è simmetrico o hermitiano, allora $V_\lambda \perp V_\mu$ se λ e μ sono due autovalori distinti. Infatti, se $\underline{v}$ è un autovettore relativo a λ e $\underline{w}$ è un autovettore relativo a μ, allora[4] $\lambda \varphi(\underline{v}, \underline{w}) = \varphi(\lambda \underline{v}, \underline{w}) = \varphi(\underline{v}, \mu \underline{w}) = \mu \varphi(\underline{v}, \underline{w})$; poiché $\lambda \neq \mu$ deve allora per forza valere $\varphi(\underline{v}, \underline{w}) = 0$.

Se f è simmetrico o hermitiano, esiste sempre una base ortonormale di autovettori (*teorema spettrale*). Se così non fosse, detto $W = V_{\lambda_1} \oplus^\perp \cdots \oplus^\perp V_{\lambda_k}$, $W^\perp$ sarebbe f-invariante e simmetrico/hermitiano, e dunque ammetterebbe un autovalore reale, contrariamente a quanto ipotizzato, $\lightning$. Alternativamente, poiché f è simmetrico (e in tal caso anche perché il polinomio caratteristico è completamente fattorizzabile in $\mathbb{R}$) o hermitiano, f è anche normale, ed è dunque diagonalizzabile; allora, poiché gli autospazi sono in somma diretta ortogonale, f è anche ortogonalmente o unitariamente diagonalizzabile.

In termini matriciali, se A è una matrice simmetrica a elementi reali (o hermitiana a elementi complessi), esiste una matrice $O \in O(n)$ (o $U \in U(n)$) tale per cui $O^\top A O$ (o $U \in U(n)$) è diagonale. Infatti f_A, l'operatore indotto da A nella base ortonormale di $\mathbb{R}^n$ (o $\mathbb{C}^n$), è un operatore simmetrico (o hermitiano) rispetto al prodotto standard dello spazio euclideo che si sta studiando.
Una matrice reale è simmetrica se e solo se è ortogonalmente diagonalizzabile. Una matrice complessa è hermitiana se e solo se è unitariamente diagonalizzabile con autovalori reali.

Radice quadrata di una matrice simmetrica, decomposizione polare e simultanea ortogonalizzabilità

Se $A \in \mathrm{Sym}(n, \mathbb{R})$ è semidefinita positiva, allora esiste sempre una matrice $S \in \mathrm{Sym}(n, \mathbb{R})$ tale per cui $S^2 = A$. Se si suppone anche che S è semidefinita positiva, tale matrice diventa unica e viene detta *radice quadrata* di A, indicata come $\sqrt{A}$.
Per costruire tale radice quadrata è sufficiente considerare $P \in O(n)$ tale per cui $P^\top A P = D$, dove $D \in M(n, \mathbb{R})$ è diagonale, secondo il teorema spettrale. Poiché A è semidefinita positiva, D si compone di soli elementi non negativi, ed è dunque possibile costruire la matrice $\sqrt{D} \in M(n, \mathbb{R})$ dove $\sqrt{D}_{ii} = \sqrt{D_{ii}}$ (da cui si deduce che $\sqrt{D}^2 = D$ e che $\sqrt{D}$ è esattamente la radice quadrata di D).
Si consideri dunque $S = P\sqrt{D}P^\top$; vale che $S^2 = PDP^\top = A$, e dunque S è la radice quadrata $\sqrt{A}$ di A (per dimostrare l'unicità di tale matrice è sufficiente ridursi all'uguaglianza negli autospazi). Si osserva che se A è definita positiva, anche S lo è.

Se $A \in M(n, \mathbb{R})$ esistono e sono uniche le matrici $P \in O(n)$, $S \in \mathrm{Sym}(n, \mathbb{R})$, con S semidefinita positiva, tali per cui $A = PS$. In particolare vale che $S = \sqrt{AA^\top}$; se dunque $A \in \mathrm{GL}(n, \mathbb{R})$, S è definita positiva (in tal caso $\mathrm{Ker}\, A^\top A = \mathrm{Ker}\, A = \{\underline{0}\}$ – come visto nella sezione sulle matrici –, e dunque $\underline{x}^\top A^\top A \underline{x} = \langle A\underline{x}, A\underline{x} \rangle > 0$ $\implies A^\top A > 0 \implies S > 0$).

Se $A \in \mathrm{GL}(n, \mathbb{R})$, esistono e sono unici $P \in O(n)$, $S \in \mathrm{Sym}(n, \mathbb{R})$ tali per cui $A = PS$ (in particolare $S = \sqrt{AA^\top}$).

Due prodotti φ, ψ si dicono simultaneamente ortogonalizzabili se esiste una base $\mathcal{B}$ tale per cui sia che $M_{\mathcal{B}}(\varphi)$ che $M_{\mathcal{B}}(\psi)$ sono diagonali (ossia se esiste una base ortogonale per entrambi i prodotti).

Se $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, φ è definito positivo, allora i due prodotti φ e ψ sono sempre simultaneamente ortogonalizzabili. È sufficiente infatti prendere una base $\mathcal{B}$ ortonormale di φ, e trovare la base ortonormale $\mathcal{B}'$ di autovettori che rende $M_{\mathcal{B}}(\psi)$ diagonale. In tale base $\mathcal{B}'$, $M_{\mathcal{B}'}(\varphi)$ è l'identità e $M_{\mathcal{B}}(\psi)$ è diagonale: dunque la base è ortogonale per ambo i prodotti.

Azioni di gruppo

Sia G un gruppo e X un insieme. Un'azione sinistra[5] di G su X a sinistra un'applicazione $\cdot : G \times X \to X$, per la quale si pone $g \cdot x := \cdot(g, x)$, tale che:

(i) $e \cdot x = x$, $\forall\, x \in X$, dove e è l'identità di G,
(ii) $g \cdot (h \cdot x) = (gh) \cdot x$, $\forall\, g, h \in G$, $\forall\, x \in X$.

Si definisce l'applicazione $f_g : X \to X$ indotta dalla relazione $f_g(x) = g \cdot x$; tale applicazione è bigettiva. Se $\cdot$ è un'azione sinistra di G su X, si dice che G opera a sinistra su X e che X è un G-insieme.

Si definisce *stabilizzatore* di $x \in X$ il sottogruppo di G $\mathrm{Stab}_G(x)$ tale che:

$$\mathrm{Stab}_G(x) = \{g \in G \mid g \cdot x = x\},$$

dove si scrive $\mathrm{Stab}(x)$ per indicare $\mathrm{Stab}_G(x)$ qualora non fosse ambigua l'azione a cui ci si riferisce.
Si può costruire un omomorfismo $\tau : G \to S_X$, dove $(S_x, \circ)$ è il gruppo delle bigezioni di X, dove $\tau(g) = f_g$. Si dice che l'azione di G su X è *fedele* se l'omomorfismo $g \to f_g$ è iniettivo, ossia se e solo se:

$$f_g = \mathrm{Id}_X \implies g = e,$$

ossia se e solo se:

$$\bigcap_{x \in X} \mathrm{Stab}(x) = \{e\}.$$

[4] λ non è stato coniugato come argomento del prodotto dal momento che per il risultato precedente è reale, e quindi $\overline{\lambda} = \lambda$.
[5] Un'azione sinistra induce sempre anche un'azione destra, ponendo $x \cdot g = g^{-1} \cdot x$.

Per esempio, S_X opera fedelmente su X tramite l'azione indotta dalla relazione $g \cdot x = g(x)$ (ed è in realtà anche un'azione transitiva). G stesso opera su G tramite l'azione banale indotta dalla relazione $g \cdot g' = gg'$.

Si definisce su X la relazione $x \sim_G y \iff \exists\, g \in G$ t.c. $y = g \cdot x$. La relazione $\sim_G$ è una relazione d'equivalenza: due elementi equivalenti tramite $\sim_G$ si dicono coniugati tramite G. Le classi di equivalenza si dicono orbite di G. In particolare si definisce $\mathrm{Orb}(x) = O_x$, con $x \in X$, come $[x]_{\sim_G}$, ossia come la classe di equivalenza di x rispetto a $\sim_G$.

Si presentano alcuni esempi di orbite:

1. $\mathrm{GL}(n, \mathbb{K})$ opera su $M(n, \mathbb{K})$ tramite la similitudine e le orbite sono le classi di matrici simili, rappresentate dalle forme canoniche di Jordan,
2. $\mathrm{GL}(n, \mathbb{K})$ opera su $\mathrm{Sym}(n, \mathbb{K})$ tramite la congruenza e le orbite sono le classi di matrici congruenti, rappresentate in $\mathbb{R}$ dalle matrici diagonali con 1, -1 e 0 come elementi, e in $\mathbb{C}$ dalle stesse matrici rappresentanti delle classi di equivalenza della SD-equivalenza, ossia le matrici del tipo $I_r^{m \times n}$,
3. O_n agisce naturalmente su $\mathbb{R}^n$ e l'orbita di $\underline{x} \in \mathbb{R}^n$ è la sfera di raggio $\|\underline{x}\|$ secondo il prodotto scalare standard di $\mathbb{R}^n$.

Vale il teorema di orbita-stabilizzatore: l'applicazione $f : G/\mathrm{Stab}_G(x) \to \mathrm{Orb}(x)$ tale che $g\,\mathrm{Stab}_G(x) \mapsto g \cdot x$ è una bigezione tra $G/\mathrm{Stab}_G(x)$ e $\mathrm{Orb}(x)$ (tale teorema è un analogo del primo teorema di omomorfismo per i gruppi). Se G è finito, vale allora che $|G| = |\mathrm{Stab}_G(x)| \cdot |\mathrm{Orb}(x)|$.

Si dice che G opera *liberamente* su X se $\forall\, x \in X$ l'applicazione da G in X tale che $g \mapsto g \cdot x$ è iniettiva, ossia se e solo se $\mathrm{Stab}_G(x) = \{e\}$, $\forall\, x \in X$. Se G opera liberamente su X, G opera anche fedelmente su X.

Si dice che G opera *transitivamente* su X se $x \sim_G y$, $\forall\, x, y \in X$, cioè se esiste un'unica orbita, che coincide dunque con G. In tal caso si dice che X è *omogeneo* per l'azione di G, oppure che X è G-omogeneo.

Si presentano alcuni esempi di azioni transitive:

1. O_n opera transitivamente sulla sfera n-dimensionale di $\mathbb{R}^n$,
2. Sia $\mathrm{Gr}_k(\mathbb{R}^n) = \{W \text{ sottospazio di } \mathbb{R}^n \mid \dim W = k\}$ la Grassmanniana di ordine k su $\mathbb{R}^n$. Allora O_n opera transitivamente su $\mathrm{Gr}_k(\mathbb{R}^n)$.

Si dice che G opera in maniera *semplicemente transitiva* su X se opera transitivamente e liberamente su X; in tal caso si dice che X è un insieme G-omogeneo principale. Equivalentemente G opera in maniera semplicemente transitiva se $\exists\, x \in X$ t.c. $g \to g \cdot x$ è una bigezione.

Se X è un insieme G-omogeneo e G è abeliano, allora G agisce fedelmente su $X \iff X$ è G-insieme omogeneo principale (per dimostrare l'implicazione a destra è sufficiente mostrare che, se $x \in X$, $g \in \mathrm{Stab}(x) \implies f_g = \mathrm{Id}_X$, da cui si conclude che $g = e$ per la fedeltà dell'azione).

Proprietà generali di uno spazio affine

Si dice spazio affine E un qualunque insieme V-omogeneo principale, dove V è uno spazio vettoriale, inteso in tal senso come il gruppo abeliano $(V, +)$. Si scrive in tal caso l'azione $\underline{v} \cdot P$ come $P + \underline{v}$. Equivalentemente E è uno spazio affine se $\forall\, P, Q \in E$, $\exists!\, \underline{v} \in V$ t.c. $P + \underline{v} = Q$. In particolar modo, ci si riferisce a $\underline{v} \mid P + \underline{v} = Q$ come $Q - P$ o $\overrightarrow{PQ}$.

Valgono le seguenti proprietà generali:

- fissato $\underline{v} \in V$, l'applicazione da E in E tale che $P \mapsto P + \underline{v}$ è una bigezione,
- fissato $O \in E$, l'applicazione da V in E tale che $\underline{v} \mapsto O + \underline{v}$ è una bigezione,
- fissato $O \in E$, l'applicazione da E in V tale che $P \mapsto P - O$ è una bigezione ed è l'inversa della bigezione presentata nello scorso punto.

Siano P_1, ..., $P_k \in E$ e λ_1, ..., $\lambda_k \in \mathbb{K}$. Siano inoltre O, $O' \in E$. Allora se si pone $P = O + \sum_{i=1}^k \lambda_i (P_i - O)$ e $P' = O' + \sum_{i=1}^k \lambda_i (P_i - O')$, vale che:

$$P = P' \;\forall\, O, O' \in E \iff \sum_{i=1}^k \lambda_i = 1$$

Pertanto un punto $P \in E$ si dice *combinazione affine* dei punti P_1, ..., P_k se $\exists\, \lambda_1$, ..., $\lambda_k \in \mathbb{K}$ tali che $\sum_{i=1}^k \lambda_i = 1$ e che $\forall\, O \in E$, $P = O + \sum_{i=1}^k \lambda_i (P_i - O)$. Si scrive in tal caso $P = \sum_{i=1}^k \lambda_i P_i$ (la notazione è ben definita dal momento che non dipende da O per l'asserzione precedente).

Un sottoinsieme $D \subseteq E$ si dice *sottospazio affine* se è chiuso per combinazioni affini. Il sottospazio affine $D \subseteq E$ generato da un sottoinsieme $S \subseteq E$ è l'insieme delle combinazioni affini (finite) dei punti di S; si denota tale sottospazio affine D come $\mathrm{Aff}(S)$. Vale inoltre che $\mathrm{Aff}(S)$ è il più piccolo sottospazio affine contenente S.

Ogni spazio vettoriale V su $\mathbb{K}$ induce uno spazio affine tramite l'azione banale che compie $(V, +)$ su $(V, +)$, ossia con $\underline{v} \cdot \underline{w} = \underline{v} + \underline{w} = \underline{w} + \underline{v}$, dove l'operazione $+$ coincide sia con la somma affine che con quella vettoriale. In questo caso una combinazione affine diventa un caso particolare di combinazione lineare. Lo spazio affine generato in questo modo su $\mathbb{K}^n$ viene detto *spazio affine standard* ed è indicato come $\mathcal{A}_\mathrm{n}(\mathbb{K})$.

Se E è uno spazio affine sul $\mathbb{K}$-spazio V, allora ogni scelta di un punto $O \in E$ e di una base $\mathcal{B}$ di V induce la bigezione naturale $\varphi_{O,\mathcal{B}} : E \to \mathcal{A}_\mathrm{n}(\mathbb{K})$ tale che $\varphi_{O,\mathcal{B}}(P) = [P - O]_\mathcal{B}$, dove $P \in E$.

Un sottoinsieme $D \subseteq E$ è un sottospazio affine $\iff \forall\, P_0 \in D$, $D_0 = \{P - P_0 \mid P \in D\} \subseteq V$ è un sottospazio vettoriale di V. Si può allora scrivere che $D = P_0 + D_0$, ossia si deduce che D è il traslato di D_0 per P_0, e quindi che ogni sottospazio affine è in particolare il traslato di un sottospazio vettoriale. L'insieme D_0, scritto anche come $\mathrm{Giac}(D)$, è detto *direzione* (o *giacitura*) del sottospazio affine D ed è invariante per la scelta del punto P_0; in particolare vale che $D_0 = \{Q - P \mid P, Q \in D\}$.

Si definisce la dimensione di un sottospazio affine D come la dimensione della sua direzione D_0. In particolare $\dim E = \dim V$. Quindi, così come accade per gli spazi vettoriali, i sottospazi affini di dimensione nulla corrispondono ai punti di E, quelli di dimensione unitaria corrispondono alle *rette* di E, quelli di dimensione 2 corrispondono ai *piani*, mentre quelli di codimensione unitaria (ossia di dimensione $\dim V - 1$) corrispondono agli *iperpiani affini*.

Due sottospazi affini con la stessa direzione si dicono *paralleli* se sono distinti, o *coincidenti* se sono uguali. Due sottospazi affini paralleli hanno sempre intersezione vuota e si ottengono l'uno dall'altro mediante traslazione.

Dei punti P_1, ..., $P_k \in E$ si dicono *affinemente indipendenti* se per $P \in \mathrm{Aff}(P_1, \ldots, P_k)$ esistono unici λ_1, ..., λ_k tali per cui $P = \sum_{i=1}^k \lambda_i P_i$ è una combinazione affine. Un sottoinsieme $S \subseteq E$ si dice affinemente indipendente se ogni suo sottoinsieme finito è affinemente indipendente.

I punti P_1, ..., P_k sono affinemente indipendenti se e solo se $\forall\, i = 1\text{—}k$ i vettori $P_j - P_i$ con $j \neq i$ sono linearmente indipendenti in $V \iff \forall\, i = 1\text{—}k$, $P_i \notin \mathrm{Aff}(S \setminus \{P_i\})$, dove $S = \{P_1, \ldots, P_k\}$. Pertanto, possono esistere al più dim $D_0 + 1$ punti affinemente indipendenti in D. In particolare, se si scelgono $n + 1$ punti P_0, ..., $P_n \in E$ affinemente indipendenti, vale che $\mathrm{Aff}(P_0, \ldots, P_n) = E$ (in tal caso infatti la direzione sarebbe tutto V). Esistono sempre P_0, ..., P_n punti di D tali che $\mathrm{Aff}(P_0, \ldots, P_n) = D$, se $\dim D = n$; in tal caso l'insieme di questi punti viene detto *riferimento affine*. Ogni riferimento affine ha lo stesso numero di elementi (in generale valgono le stesse proprietà di una base vettoriale, mediante cui se ne dimostra l'esistenza).

Sia $E = \mathcal{A}_\mathrm{n}(\mathbb{K})$ allora $\underline{w_1}$, ..., $\underline{w_n} \in E$ sono affinemente indipendenti se e solo se i vettori $\underline{\hat{w_1}}$, ..., $\underline{\hat{w_n}}$ con

$\underline{\hat{w_i}} = \begin{pmatrix} \underline{w_i} \\ \hline 1 \end{pmatrix} = \iota(\underline{w_i}) \in \mathbb{K}^{n+1}$ sono linearmente indipendenti.

Siano P_0, ..., P_k i punti di un riferimento affine per il sottospazio affine D. Allora ogni punto $P \in D$ è univocamente determinato dagli scalari λ_i in $\mathbb{K}$ tali per cui $P = \sum_{i=0}^k \lambda_i P_i$, eccetto per uno di questi scalari che è già determinato dagli altri (infatti vale sempre $\sum_{i=0}^k \lambda_i = 1$). Vi è dunque una bigezione tra D e $\mathcal{A}_k(\mathbb{K})$. L'immagine di P tramite questa bigezione è un vettore contenente le cosiddette *coordinate affini* di P.

Si dice *combinazione convessa* una qualsiasi combinazione affine finita in un insieme di punti affinemente indipendenti S in cui ogni coordinata affine è maggiore o uguale a zero. Si pone in particolare $\mathrm{IC}(S)$ come l'insieme di questo tipo di combinazioni (intuitivamente un inviluppo convesso è l'insieme dei punti contenuti "dentro" il riferimento affine scelto; per tre punti è il triangolo, per due punti è il segmento). Si scrive $\mathrm{IC}(P_1, \ldots, P_k)$ per indicare $\mathrm{IC}(\{P_1, \ldots, P_k\})$.

Si osserva che $\mathrm{IC}(S)$ è un insieme convesso (ossia $\forall\, P, Q \in \mathrm{IC}(S)$, $[P, Q] \subseteq \mathrm{IC}(S)$, dove $[P, Q] := \mathrm{IC}(\{P, Q\})$ è il segmento congiungente P e Q).

Si definisce il *baricentro geometrico* di P_1, ..., $P_n \in E$ come la seguente combinazione convessa:

$$G = \frac{1}{n}\sum_{i=1}^{n} P_i \in \mathrm{IC}(P_1, \dots, P_n).$$

Se $A \subseteq E$ è finito, si definisce G_A come il baricentro geometrico dei punti di A. Inoltre, se A è un'unione di insiemi disgiunti, G_A è una combinazione convessa dei baricentri di questi insiemi con peso la loro cardinalità divisa per la cardinalità di A; in altre parole se $A = B \sqcup C$ (i.e. $A = B \cup C \wedge B \cap C = \varnothing$), allora:

$$G_A = \frac{|B|}{|A|}G_B + \frac{|C|}{|A|}G_C.$$

In questo modo si dimostra facilmente che in un triangolo il baricentro geometrico giace sulle congiungenti dei punti medi con i vertici opposti.

Si osserva che se A e B sono due sottospazi affini, allora anche $A \cap B$ è un sottospazio affine se $A \cap B \neq \varnothing$. Inoltre, se $A \cap B \neq \varnothing$, allora $(A \cap B)_0 = A_0 \cap B_0$.

Si definisce *somma affine* $A + B$ di due sottospazi affini A e B di E il sottospazio affine $\mathrm{Aff}(A \cup B)$. In generale vale la seguente uguaglianza:

$$(A + B)_0 = A_0 + B_0 + \mathrm{Span}(P_0' - P_0), \quad P_0' \in A, P_0 \in B.$$

Inoltre $\mathrm{Span}(P_0' - P_0) \subseteq A_0 + B_0 \iff A \cap B \neq \varnothing$, altrimenti $(A + B)_0 = A_0 + B_0 \oplus \mathrm{Span}(P_0' - P_0)$. Pertanto, se $A \cap B \neq \varnothing$, continua a valere la formula di Grassmann:

$$\dim(A + B) = \dim A + \dim B - \dim(A \cap B) \quad \text{se } A \cap B \neq \varnothing,$$

altrimenti vale la formula di Grassmann modificata:

$$\dim(A + B) = \dim A + \dim B - \dim(A \cap B) + 1 \quad \text{se } A \cap B = \varnothing.$$

Applicazioni affini e affinità

Siano E spazio affine su V, E' spazio affine su V' sullo stesso campo $\mathbb{K}$.

Un'applicazione $f : E \to E'$ si dice *applicazione affine* se conserva le combinazioni affini, ossia se:

$$f\left(\sum_{i=1}^{k} \lambda_i P_i\right) = \sum_{i=1}^{k} \lambda_i f(P_i) \iff \sum_{i=1}^{k} \lambda_i = 1.$$

Se D è un sottospazio affine di E f-invariante, allora $f|_D$ è ancora un'applicazione affine. Esiste ed è unica l'applicazione lineare $g : V \to V'$ tale che $f(P) = f(O) + g(P - O)$ per ogni scelta di P, $O \in E$; tale applicazione lineare si denota con df e viene detta *differenziale* di g. Analogamente si può sempre costruire un'applicazione affine tale per cui $df = g$, data $g \in \mathcal{L}(V, V')$.

Nel caso in cui $E = \mathcal{A}_n(\mathbb{K})$, $E' = \mathcal{A}_m(\mathbb{K})$, un'applicazione affine f è della forma $f(\underline{x}) = f(\underline{0}) + g(\underline{x}) = A\underline{x} + \underline{b}$ con $A \in M(m, n, \mathbb{K})$ e $\underline{b} \in A_m(\mathbb{K})$.

Sia E'' un altro spazio affine associato a V''. Se $f' : E' \to E''$ è affine, allora $f' \circ f : E \to E''$ è affine e vale $d(f' \circ f) = df' \circ df$.

Si dice che $f : E \to E$ è un'*affinità* di E se f è un'applicazione affine bigettiva; si definisce $A(E)$ come il gruppo delle affinità di E rispetto alla composizione. Vale in particolare che f è un'affinità di $E \iff df$ è invertibile.

L'applicazione $\pi : A(E) \to \mathrm{GL}(V) : f \mapsto g$ è un omomorfismo surgettivo il cui nucleo è dato dalle traslazioni, che pertanto formano un sottogruppo normale.

Sia $f \in A(E)$. Allora $d(f^{-1}) = df^{-1}$; in particolare, se $E = \mathcal{A}_n(\mathbb{K})$, $f^{-1}(\underline{x}) = A^{-1}\underline{x} - A^{-1}\underline{b}$, dove $f(\underline{x}) = A\underline{x} + \underline{b}$.

Si definisce l'applicazione affine $\iota : \mathcal{A}_n(\mathbb{K}) \to \mathcal{A}_{n+1}(\mathbb{K})$ in modo tale che $\underline{x} \mapsto \hat{\underline{x}} = \begin{pmatrix} \underline{x} \\ 1 \end{pmatrix}$. Si osserva che ι è un isomorfismo affine tra $\mathcal{A}_n(\mathbb{K})$ e l'iperpiano $H_{n+1} = \{\underline{x} \in \mathcal{A}_{n+1}(\mathbb{K}) \mid x_{n+1} = 1\} \subset \mathcal{A}_{n+1}(\mathbb{K})$.

Sia $f \in A(\mathcal{A}_n(\mathbb{K}))$ data da $f(\underline{x}) = A\underline{x} + \underline{b}$. Allora esiste un'unica applicazione lineare $\hat{f} \in \mathrm{End}(\mathbb{K}^{n+1})$ che estende f in $\mathcal{A}_{n+1}(\mathbb{K})$ tale per cui $f(\iota(\underline{x})) = \iota(f(\underline{x}))$; in particolare tale applicazione è rappresentata dalla matrice $\hat{A}$, dove:

$$\hat{A} = \left(\begin{array}{c|c} A & \underline{b} \\ \hline 0 & 1 \end{array}\right).$$

In particolare $\hat{A}$ dipende da $n^2 + n$ parametri; se si fosse posto $f(D) = D$, sarebbe dipesa invece da $k(k + 1) + n(n - k)$ parametri. Le matrici di questa forma formano un sottogruppo di $\mathrm{GL}_{n+1}(\mathbb{K})$ isomorfo ad $A(\mathcal{A}_n(\mathbb{K}))$.

Sia E spazio affine di dimensione n.

(i) se $f \in A(E)$ e i punti P_0, ..., $P_n \in E$ sono affinemente indipendenti, allora anche i punti $f(P_0)$, ..., $f(P_n)$ sono affinemente indipendenti,

(ii) se $\dim E_0 = n$, i punti P_0, ..., P_n sono affinemente indipendenti e anche i punti Q_0, ... Q_n sono affinemente indipendenti, allora esiste ed è unica l'affinità $f : E \to E$ tale che $f(P_i) = Q_i \; \forall\, i = 1—n$,

(iii) se $f \in A(E)$, $D \subseteq E$ sottospazio affine $\implies f(D)$ è un sottospazio affine della stessa dimensione.

Siano (P_1, P_2, P_3), (Q_1, Q_2, Q_3) due terne di punti distinti di $\mathcal{A}_1(\mathbb{K})$. Allora esiste ed è unica l'applicazione affine $f \in A(\mathcal{A}_1(\mathbb{K}))$ tale che $f(P_i) = Q_i \; \forall\, i = 1, 2, 3$ $\iff \lambda(P_1, P_2, P_3) = \lambda(Q_1, Q_2, Q_3)$, dove $\lambda(P_1, P_2, P_3)$ è detto *rapporto semplice* ed è definito come:

$$\lambda(P_1, P_2, P_3) = \frac{P_3 - P_1}{P_2 - P_1}.$$

Infatti f è già unica ponendo $f(P_1) = Q_1$ e $f(P_2) = Q_2$; allora, poiché $\mathcal{A}_1(\mathbb{K})$ è di dimensione unitaria, P_3 deve scriversi come combinazione affine di P_1 e P_2 in modo tale che $P_3 = P_1 + \lambda(P_2 - P_1)$. In questo modo, poiché $f(P_3) = Q_3$, anche $Q_3 = Q_1 + \lambda(Q_2 - Q_1)$, da cui la motivazione dietro all'uguaglianza dei rapporti semplici.

- $A(\mathcal{A}_1(\mathbb{K}))$ agisce transitivamente su $\mathcal{A}_1(\mathbb{K})$, $\mathrm{Stab}(x_0) = \{f \mid f(x_0) = x_0\} \cong \mathrm{GL}_1(\mathbb{K})$ (infatti la matrice associata all'affinità dipende da un solo parametro),
- $|\mathrm{Fix}(f)| \leq 1$, e $|\mathrm{Fix}(f)| = 0 \iff f$ è una traslazione, dove $\mathrm{Fix}(f) = \{x \in \mathcal{A}_1(\mathbb{K}) \mid f(x) = x\}$,
- $A(\mathcal{A}_1(\mathbb{K}))$ agisce in maniera semplicemente transitiva sulle coppie di punti $(P_1, P_2) \in \mathcal{A}_1(\mathbb{K}) \times \mathcal{A}_1(\mathbb{K})$ con $P_1 \neq P_2$.

Sia $f(\underline{x}) = M\underline{x} + \underline{t}$ un'affinità di $A(\mathcal{A}_n(\mathbb{K}))$. Allora, se 1 non è un autovalore di M, f ha un unico punto fisso (in tal caso, infatti $(M - I)$ è invertibile, e quindi $(M - I)\underline{x} = -\underline{t}$ ammette un'unica soluzione).

Spazio proiettivo

Si definisce *spazio proiettivo* relativo a $\mathbb{K}^{n+1}$ l'insieme delle rette di $\mathbb{K}^{n+1}$. Tale spazio viene denotato come $\mathbb{P}(\mathbb{K}^{n+1}) = \mathbb{P}^n(\mathbb{K})$ (intuitivamente lo spazio proiettivo perde una dimensione rispetto allo spazio di partenza perché è la proiezione di tutte le rette in un unico punto, eccetto per i punti all'infinito).

Equivalentemente lo spazio proiettivo è l'insieme quoziente di $\mathbb{K}^{n+1}$ tramite la relazione di equivalenza $\sim$ dove $\underline{x} \sim \underline{y} \overset{\text{def}}{\iff} \exists\, \lambda \in \mathbb{K}, \lambda \neq 0 \mid \underline{x} = \lambda\underline{y}$.

Ogni punto $\underline{x} \in \mathbb{K}^n$ individua un unico sottospazio di dimensione unitaria in $\mathbb{K}^{n+1}$ tramite ι, ossia: $\mathrm{Span}(\iota(\underline{x})) = \mathrm{Span}(\begin{pmatrix} \underline{x} & | & 1 \end{pmatrix}^\top)$. L'insieme di rette non individuate tramite elementi di $\mathbb{K}^n$ è in particolare formato dalle rette appartenenti al piano $\{\underline{x} \in \mathbb{K}^{n+1} \mid x_{n+1} = 0\} \cong \mathbb{K}^n$; dal momento che queste rette si identificano come tutte le rette di $\mathbb{K}^n$, esse rappresentano in particolare lo spazio proiettivo di una dimensione ancora minore, $\mathbb{P}^{n-1}(\mathbb{K})$.

Le rette appartenenti al piano $\{\underline{x} \in \mathbb{K}^{n+1} \mid x_{n+1} = 0\}$ sono dette *punti all'infinito* di $\mathbb{P}^n(\mathbb{K})$ (intuitivamente un punto all'infinito indica la direzione dei vari infiniti del piano).

Si può ricoprire $\mathbb{P}^n(\mathbb{K})$ con gli iperpiani $H_i = \{\underline{x} \in \mathcal{A}_{n+1}(\mathbb{K}) \mid x_i = 1\}$ dal momento che ogni retta deve intersecare almeno uno di questi iperpiani in un punto.

Coniche e quadriche

Si definisce *quadrica* il luogo di zeri in $\mathcal{A}_n(\mathbb{K})$ di un polinomio di secondo grado $p(x_1, \dots, x_n) \in \mathbb{K}[x_1, \dots, x_n]$ in n variabili, dove si identifica con una n-upla $(x_1, \dots, x_n)$ di variabili che sono zeri del polinomio un elemento di $\mathcal{A}_n(\mathbb{K})$ con le stesse coordinate. Una *conica* è in particolare una quadrica in due variabili. Si osserva che una quadrica è invariante per la moltiplicazione per uno scalare diverso da 0.

D'ora in poi si intenderà con $\underline{x}$ l'n-upla $(x_1, \dots, x_n)$ e si supporrà $\mathrm{char}\,\mathbb{K} \neq 2$. Il polinomio $p(x_1, \dots, x_n)$ può essere descritto come:

$$p(x_1, \dots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij}x_ix_y + \sum_{i=1}^{n} 2b_ix_i + c.$$

Si definiscono le seguenti quantità:

- la *parte quadratica* $\mathcal{A}(p) \in \mathrm{Sym}(n, \mathbb{K})$ di p, dove:

$$\mathcal{A}(p)_{ij} = \begin{cases} a_{ij} & \text{se } i \leq j, \\ a_{ji} & \text{se } i > j. \end{cases}$$

- la *parte lineare* $\mathcal{L}(p) \in M(n, 1, \mathbb{K})$ di p, dove $\mathcal{L}(p)_{i1} = b_i$,
- il *termine noto* $c(p) \in \mathbb{K}$ di p, dove $c(p) = c$.

Allora il polinomio può essere riscritto come:

$$p(\underline{x}) = \underline{x}^\top \mathcal{A}(p)\, \underline{x} + 2\mathcal{L}(p)^\top \underline{x} + c(p).$$

Si definisce inoltre la matrice $\mathcal{M}(p)$, dove:

$$\mathcal{M}(p) = \left(\begin{array}{c|c} \mathcal{A}(p) & \mathcal{L}(p) \\ \hline \mathcal{L}(p) & c(p) \end{array}\right) \in \mathrm{Sym}(n+1, \mathbb{K}),$$

la quale viene detta *matrice associata alla quadrica* in p. Si dice che la quadrica relativa a p è *non degenere* se $\mathrm{rg}(\mathcal{M}(p)) = n+1$. Allora, tramite l'identificazione di $\mathcal{A}_n(\mathbb{K})$ in $H_{n+1} \subset \mathbb{K}^{n+1}$ mediante ι, vale che:

$$p(\underline{x}) = \hat{\underline{x}}^\top \mathcal{M}(p) \hat{x}, \quad \text{dove } \hat{\underline{x}} = \iota(\underline{x}) = \left(\underline{x} \;\middle|\; 1\right)^\top.$$

Si deduce allora che una quadrica altro non è che la controimmagine tramite ι dell'intersezione tra H_{n+1} e il cono isotropo $\mathrm{CI}(\varphi_{\mathcal{M}(p)})$, dove $\varphi_{\mathcal{M}(p)}$ è il prodotto scalare indotto da $\mathcal{M}(p)$ in $\mathbb{K}^{n+1}$ (i.e. la quadrica è esattamente $\iota^{-1}(H_{n+1} \cap \mathrm{CI}(\varphi_{\mathcal{M}(p)}))$).

Nel caso di una conica determinata dal polinomio $p(x, y) = ax^2 + by^2 + cxy + dx + ey + f$, vale che:

$$\mathcal{M}(p) = \left(\begin{array}{cc|c} a & c/2 & d/2 \\ c/2 & b & e/2 \\ \hline d/2 & e/2 & f \end{array}\right).$$

Sia $f \in A(\mathcal{A}_n(\mathbb{K}))$ tale per cui $f(\underline{x}) = M\underline{x} + \underline{t}$, con $M \in \mathrm{GL}(n, \mathbb{K}), \underline{t} \in \mathbb{K}^n$. Si definisce allora un'azione destra di $A(\mathcal{A}_n(\mathbb{K}))$ sulle quadriche di n variabili, dove $p \cdot f$ è indicato come $p \circ f$, che a sua volta indica il polinomio $p(f(\underline{x})) = p(M\underline{x} + \underline{t})$. Il luogo di zeri $Z(p)$ di una quadrica su cui agisce $A(\mathcal{A}_n(\mathbb{K}))$ varia a sua volta secondo l'affinità; in particolare vale che:

$$Z(p) = f(Z(p \circ f)).$$

Si definisce allora una relazione di equivalenza detta *equivalenza affine*, dove:

$$p \sim q \overset{\text{def}}{\iff} \exists\, \lambda \in \mathbb{K}^*, f \in A(\mathcal{A}_n(\mathbb{K})) \mid p = \lambda(q \circ f).$$

Vale inoltre la seguente identità:

$$p(f(\underline{x})) = \underline{x}^\top A' \underline{x} + 2\underline{b'}^\top \underline{x} + c',$$

dove $\mathcal{A}' = M^\top \mathcal{A} M$, $\underline{b'} = M^\top(\mathcal{A}\underline{t} + \underline{b})$ e $c' = p(\underline{t})$. Pertanto la matrice $\mathcal{M}(p \circ f)$ può essere scritta come:

$$\mathcal{M}(p \circ f) = \hat{M}^\top \mathcal{M}(p) \hat{M} = \left(\begin{array}{c|c} M^\top \mathcal{A}(p) M & M^\top(\mathcal{A}(p)\underline{t} + \mathcal{L}(p)) \\ \hline (M^\top(\mathcal{A}(p)\underline{t} + \mathcal{L}(p)))^\top & p(\underline{t}) \end{array}\right).$$

Se f è solo una traslazione (i.e. $M = I_n$), la formula si semplifica:

$$\mathcal{M}(p \circ f) = \hat{M}^\top \mathcal{M}(p) \hat{M} = \left(\begin{array}{c|c} \mathcal{A}(p) & \mathcal{A}(p)\underline{t} + \mathcal{L}(p) \\ \hline (\mathcal{A}(p)\underline{t} + \mathcal{L}(p))^\top & p(\underline{t}) \end{array}\right).$$

Una quadrica relativa al polinomio p si dice *a centro* se $\exists\, \underline{x_0} \in \mathcal{A}_n(\mathbb{K}) \mid p(\underline{x_0} + \underline{t}) = p(\underline{x_0} - \underline{t})$, dove $\underline{t} \in \mathcal{A}_n(\mathbb{K})$; in particolare tale $\underline{x_0}$ è detto *centro di simmetria*. Vale in particolare che $\underline{0}$ è un centro di simmetria di p se e solo se $\mathcal{L}(p) = \underline{0}$.

Inoltre vale che $\underline{x_0}$ è un centro di simmetria di p se e solo se $p \circ f$ ha centro di simmetria $\underline{0}$ tramite la traslazione $f(\underline{x}) = \underline{x} + \underline{x_0}$; pertanto p è a centro se e solo se è risolvibile in $\underline{x_0}$ il sistema:

$$\mathcal{L}(p \circ f) = \mathcal{A}(p)\underline{x_0} + \mathcal{L}(p) = \underline{0}.$$

Poiché allora i centri della quadriche sono soluzione di un sistema lineare, se esistono, essi formano un sottospazio affine di dimensione $n - \mathrm{rg}(\mathcal{A})$; in particolare, se $\underline{x_0}$ è un particolare centro, tale sottospazio affine C è esattamente dato da:

$$C = \underline{x_0} + \mathrm{Ker}\, \mathcal{A}(p).$$

Classificazione delle coniche in $\mathbb{C}$ ed $\mathbb{R}$

Esistono due tipi di classificazioni: una *affine*, dove per ogni quadrica si trova una forma canonica per equivalenza affine tramite la moltiplicazione per scalare λ e per applicazione delle affinità di $A(\mathcal{A}_n(\mathbb{K}))$, e una *isometrica*, dove si classificano le quadriche rispetto alle isometrie del gruppo delle isometrie $\mathrm{Iso}(\mathcal{A}_n(\mathbb{K}))$ (e.g. le ellissi in generale sono affinemente equivalenti, ma non sono isometricamente equivalenti), dove $\mathrm{Iso}(\mathcal{A}_n(\mathbb{K}))$ è composto dalla affinità di $A(\mathcal{A}_n(\mathbb{K}))$ che preservano la distanza tra punti, qualora definibile.

Sia $\mathbb{K} = \mathbb{C}$. Allora ogni conica è affinemente equivalente ad un'equazione canonica della seguente tabella, unicamente determinata dagli invarianti $\mathrm{rg}(\mathcal{M}(p))$ e $\mathrm{rg}(\mathcal{A}(p))$.

	$\mathrm{rg}(\mathcal{M}(p))$	$\mathrm{rg}(\mathcal{A}(p))$	Eq. canonica	A centro
$\mathcal{C}_1$	3	2	$x^2 + y^2 = 1$	Sì
$\mathcal{C}_2$	3	1	$x^2 = y$	No
$\mathcal{C}_3$	2	2	$x^2 + y^2 = 0$	Sì
$\mathcal{C}_4$	2	1	$x^2 = 1$	Sì
$\mathcal{C}_5$	1	1	$x^2 = 0$	Sì

Tra le uniche coniche non degeneri di $\mathbb{C}$, $\mathcal{C}_1$ prende il nome di *ellisse* e $\mathcal{C}_2$ di *parabola*. $\mathcal{C}_3$ rappresenta invece una *coppia di rette incidenti*, $\mathcal{C}_4$ una *coppia di rette parallele* e $\mathcal{C}_5$ un singolo *punto*.

Sia $\mathbb{K} = \mathbb{R}$. Allora ogni conica è affinemente equivalente ad un'equazione canonica della seguente tabella, unicamente determinata dagli invarianti $\mathrm{rg}(\mathcal{M}(p))$, $\mathrm{rg}(\mathcal{A}(p))$, $S(\mathcal{M}(p)) := |\iota_+(\mathcal{M}(p)) - \iota_-(\mathcal{M}(p))|$ e $S(\mathcal{A}(p)) := |\iota_+(\mathcal{A}(p)) - \iota_-(\mathcal{A}(p))|$.

	$\mathrm{rg}(\mathcal{M}(p))$	$\mathrm{rg}(\mathcal{A}(p))$	$S(\mathcal{M}(p))$	$S(\mathcal{A}(p))$	Eq. canonica
$\mathcal{C}_1$	3	2	1	2	$x^2 + y^2 - 1 = 0$
$\mathcal{C}_2$	3	2	1	0	$x^2 - y^2 - 1 = 0$
$\mathcal{C}_3$	3	1	1	1	$x^2 - y = 0$
$\mathcal{C}_4$	2	2	0	0	$x^2 - y^2 = 0$
$\mathcal{C}_5$	2	1	0	1	$x^2 - 1 = 0$
$\mathcal{C}_6$	1	1	1	1	$x^2 = 0$
$\mathcal{C}_7$	3	2	3	2	$x^2 + y^2 + 1 = 0$
$\mathcal{C}_8$	2	2	2	2	$x^2 + y^2 = 0$
$\mathcal{C}_9$	2	1	2	1	$x^2 + 1 = 0$

Per $\mathbb{K} = \mathbb{R}$, $\mathcal{C}_1$ prende il nome di *ellisse reale*, $\mathcal{C}_2$ di *iperbole* e $\mathcal{C}_3$ di *parabola*. $\mathcal{C}_4$ rappresenta invece una *coppia di rette reali incidenti*, $\mathcal{C}_5$ una *coppia di rette reali parallele* e $\mathcal{C}_6$ un singolo *punto*. $\mathcal{C}_7$ è una *ellisse immaginaria*, $\mathcal{C}_8$ una *coppia di rette immaginarie incidenti* e $\mathcal{C}_9$ una *coppia di rette immaginarie parallele*. Tutte queste coniche sono a centro eccetto per la parabola ($\mathcal{C}_3$).

Ad opera di Gabriel Antonio Videtta,
https://poisson.phc.dm.unipi.it/~videtta/.
Reperibile su https://g1.hearot.it.