



UNIVERSITÁ DI PISA – FACOLTÁ DI MATEMATICA

APPUNTI DI
Geometria Analitica e Algebra Lineare

Giacomo Mezzedimi

frutto della rielaborazione delle lezioni tenute dai professori

E. Fortuna

R. Frigerio

a.a. 2013-2014

09/05/2017

INTRODUZIONE

Questi appunti nascono dall'esigenza mia (ma credo anche di altri) di un supporto per lo studio del corso "Geometria Analitica e Algebra Lineare" al primo anno; a differenza infatti di molti altri corsi, non è facile trovare del materiale adatto da affiancare durante lo studio.

Sostanzialmente queste pagine contengono gli argomenti svolti dalla professoressa E. Fortuna e dal professore R. Frigerio durante l'anno accademico 2013-2014, anno in cui io ho seguito il corso; molte parti sono prese dai lucidi della professoressa Fortuna, ma alcune sono state riadattate/modificate/completate per dare una continuità al testo.

I paragrafi 3.6 (Basi cicliche per endomorfismi) e 5.3 (Geometria affine euclidea) sono stati aggiunti nel giugno del 2015, in quanto svolti nell'a.a. 2014-2015; voglio ringraziare a tal proposito Dario Balboni, che ha realizzato questi due paragrafi, oltre ad avermi aiutato nell'opera di correzione del testo.

Voglio infine ringraziare tutti quelli che hanno contribuito o contribuiranno a migliorare questi appunti: è impossibile rendere un testo completamente privo di errori, ma l'obiettivo è quello di ripulirlo più possibile; invito dunque tutti a segnalarmi qualunque tipo di errore/imprecisione presente in queste pagine (la mia e-mail è mezzedimi@mail.dm.unipi.it).

Nella speranza che questi appunti vi siano utili, vi auguro un buono studio.

Giacomo Mezzedimi (con l'accento sulla seconda e)

SOMMARIO

CAPITOLO 1: Prime definizioni e proprietà	3
• 1.1 Prime definizioni	3
• 1.2 Strutture algebriche	6
CAPITOLO 2: Spazi vettoriali e applicazioni lineari	13
• 2.1 Spazi vettoriali	13
• 2.2 Spazi di matrici	14
• 2.3 Sottospazi e combinazioni lineari	15
• 2.4 Applicazioni lineari	18
• 2.5 Sistemi lineari	25
• 2.6 Basi e dimensione	30
• 2.7 Rango	38
• 2.8 SD-equivalenza	42
• 2.9 Spazio duale	47
CAPITOLO 3: Endomorfismi	50
• 3.0 Alcune nozioni sulle permutazioni	50
• 3.1 Determinante	51
• 3.2 Endomorfismi simili	60
• 3.3 Diagonalizzabilità	68
• 3.4 Triangolabilità	72
• 3.5 Forma canonica di Jordan	75
• 3.6 Basi cicliche per endomorfismi	90
CAPITOLO 4: Forme bilineari	92
• 4.1 Forme bilineari e forme quadratiche	92
• 4.2 Congruenza e decomposizione di Witt	97
• 4.3 Isometrie	111
• 4.4 Aggiunto	113
• 4.5 Spazi euclidei	117
• 4.6 Il teorema spettrale reale	123
CAPITOLO 5: Spazi affini	127
• 5.1 Isometrie affini	127
• 5.2 Spazi e sottospazi affini	134
• 5.3 Geometria affine euclidea	143
• 5.4 Affinità di \mathbb{K}^n	145
• 5.5 Quadriche	148

1 PRIME DEFINIZIONI E PROPRIETÀ

1.1 PRIME DEFINIZIONI

DEFINIZIONE 1.1.1: Siano A, B insiemi. Diciamo che:

- A è **sottoinsieme** di B ($A \subset B$ oppure $A \subseteq B$) se $\forall a \in A, a \in B$;
- A è **uguale** a B ($A = B$) se $A \subset B \wedge B \subset A$.

Se un insieme è finito, si può definire elencando tutti i suoi elementi:

$$A = \{a_1, \dots, a_n\}$$

Se un insieme è infinito, si definisce enunciando la proprietà che caratterizza tutti i suoi elementi:

$$A = \{x | P(x)\}$$

Esempio: $P = \{a \in \mathbb{N} | a \equiv 0 (2)\}$ è l'insieme dei numeri pari.

DEFINIZIONE 1.1.2: Dati A, B insiemi, definiamo:

- **unione** di due insiemi $A \cup B = \{x | x \in A \vee x \in B\}$;
- **intersezione** di due insiemi $A \cap B = \{x | x \in A \wedge x \in B\}$;
- **differenza** di due insiemi $A - B = A \setminus B = \{x | x \in A \wedge x \notin B\}$;
- **prodotto cartesiano** di due insiemi $A \times B = \{(a, b) | a \in A, b \in B\}$.

DEFINIZIONE 1.1.3: Una applicazione è una terna $f: A \rightarrow B$, dove A e B sono insiemi, chiamati rispettivamente **dominio** e **codominio**, e f è una legge che associa ad ogni elemento $x \in A$ **uno e un solo elemento** $f(x)$ di B .

$id_A: A \rightarrow A$ è l'applicazione identica, tale che $\forall x \in A, id_A(x) = x$.

DEFINIZIONE 1.1.4: Data un'applicazione $f: A \rightarrow B$, si definisce **immagine** di f l'insieme $Im(f) = \{y \in B | \exists x \in A \text{ tale che } f(x) = y\}$.

Vale sempre $Im(f) \subset B$.

Esempio: $f: \mathbb{N} \rightarrow \mathbb{N} | f(x) = 2x; Im(f) = \{\text{pari}\}$.

In generale, se $W \subset A$, allora $f(W) = \{y \in B | \exists x \in W \text{ tale che } f(x) = y\}$ perciò $f(W) \subset B$.

DEFINIZIONE 1.1.5: Si definisce **restrizione** di f a W , dove $f: A \rightarrow B$ e $W \subset A$, come la funzione $f|_W: W \rightarrow B$ tale che $\forall x \in W (f|_W)(x) \stackrel{\text{def}}{=} f(x)$.

In parole povere, $f|_W$ agisce come f ma in un dominio ristretto; si parla infatti di una **restrizione del dominio**.

Perciò: $Im(f|_W) = f(W)$.

DEFINIZIONE 1.1.6: Sia $f: A \rightarrow B$ e $Z \subset B$. Si indica con $f^{-1}(Z)$ il sottoinsieme del dominio che contiene tutti gli elementi che hanno immagine in Z , cioè:

$$f^{-1}(Z) = \{x \in A | f(x) \in Z\}$$

$f^{-1}(Z)$ viene chiamata **controimmagine** di Z .

DEFINIZIONE 1.1.7: Una applicazione $f: A \rightarrow B$ si dice:

- **surgettiva** se $Im(f) = B$ oppure equivalentemente se $\forall y \in B, \exists x \in A \mid f(x) = y$;
- **iniettiva** se $\forall x, y \in A, x \neq y \Rightarrow f(x) \neq f(y)$ oppure equivalentemente se $f(x) = f(y) \Rightarrow x = y$;
- **bigettiva** (o biunivoca) se è sia iniettiva che surgettiva.

DEFINIZIONE 1.1.8: Data un'applicazione $f: A \rightarrow B$ bigettiva, si definisce **funzione inversa** $f^{-1}: B \rightarrow A$ tale che $\forall y \in B, f^{-1}(y) = x$, dove $x \mid f(x) = y$.

L'unicità della x viene garantita dalla bigettività di f .

DEFINIZIONE 1.1.9: Date $f: A \rightarrow B$ e $g: B \rightarrow C$, si definisce **composizione di funzioni** la funzione $g \circ f: A \rightarrow C$ tale che $\forall x \in A (g \circ f)(x) \stackrel{\text{def}}{=} g(f(x))$.

PROPOSIZIONE 1.1.1: $Im(g \circ f) = Im(g|_{Im(f)})$

Dimostrazione:

Entrambi i contenimenti derivano direttamente dalla definizione di composizione.

DEFINIZIONE 1.1.10: Dato un insieme $E \neq \emptyset$, si definisce **relazione** \mathcal{R} su E come un sottoinsieme di $E \times E$ tale che $(x, y) \in \mathcal{R}$ per alcuni $x, y \in E$.

$(x, y) \in \mathcal{R}$ viene comunemente scritto $x\mathcal{R}y$ (x è in relazione con y).

DEFINIZIONE 1.1.11: Una relazione \mathcal{R} si dice **di equivalenza** se:

- è riflessiva, cioè $\forall x \in E, x\mathcal{R}x$;
- è simmetrica, cioè $x\mathcal{R}y \Rightarrow y\mathcal{R}x$;
- è transitiva, cioè $x\mathcal{R}y \wedge y\mathcal{R}z \Rightarrow x\mathcal{R}z$.

DEFINIZIONE 1.1.12: Sia \mathcal{R} una relazione di equivalenza. $\forall x \in E$, si definisce **classe di equivalenza** di x l'insieme $[x] = \{y \in E \mid x\mathcal{R}y\}$, cioè l'insieme degli elementi di E in relazione con x .

Evidentemente $\forall x \in E, [x] \neq \emptyset$, poiché $x \in [x]$.

LEMMA 1.1.2: Siano $x, y \in E$ e sia \mathcal{R} una relazione di equivalenza su E . Allora $[x] = [y] \Leftrightarrow x\mathcal{R}y$.

Dimostrazione:

\Rightarrow $x \in [x] \Rightarrow x \in [y] \Rightarrow x\mathcal{R}y$.

\Leftarrow Sia $z \in [x]$; allora $z\mathcal{R}x$.

Ma $x\mathcal{R}y$ per ipotesi, quindi per transitività $z\mathcal{R}y \Rightarrow z \in [y]$.

Perciò $[x] \subset [y]$.

Analogamente si prova che $[y] \subset [x]$, da cui la tesi.

PROPOSIZIONE 1.1.3: Le classi di equivalenza formano una **partizione**, cioè:

- 1) ogni classe è non vuota;
- 2) $\bigcup_{x \in E} [x] = E$;

3) $[x] \cap [y] \neq \emptyset \Rightarrow [x] = [y]$.

Dimostrazione:

1) già fatta.

2) $\forall x \in E, [x] \subset E \Rightarrow \bigcup_{x \in E} [x] \subset E$;

$\forall x \in E, x \in [x] \Rightarrow E \subset \bigcup_{x \in E} [x]$, da cui segue la tesi.

3) $[x] \cap [y] \neq \emptyset \Rightarrow \exists z \in [x] \cap [y] \Rightarrow z \mathcal{R} x \wedge z \mathcal{R} y \Rightarrow x \mathcal{R} y$ e per il lemma precedente ho che $[x] = [y]$, cioè la tesi.

Esempio: $E = \mathbb{R}, (x, y) \in \mathcal{R} \Leftrightarrow x - y \in \mathbb{Z}$.

Questa relazione è di equivalenza, in quanto:

1) è riflessiva, poiché $x - x = 0 \in \mathbb{Z}$;

2) è simmetrica, poiché se $x - y = k \in \mathbb{Z} \Rightarrow y - x = -k \in \mathbb{Z}$;

3) è transitiva, poiché se $x - y = k_1 \in \mathbb{Z}$ e $y - z = k_2 \in \mathbb{Z} \Rightarrow x - z = k_1 + k_2 \in \mathbb{Z}$.

DEFINIZIONE 1.1.13: Si definiscono **rappresentanti di una classe di equivalenza** tutti gli elementi di una certa classe.

DEFINIZIONE 1.1.14: Sia $E \neq \emptyset$ e \mathcal{R} una relazione di equivalenza su E . Si definisce **insieme quoziente** $E/\mathcal{R} \stackrel{\text{def}}{=} \{[x] \mid x \in E\}$ (si legge E modulo \mathcal{R}).

DEFINIZIONE 1.1.15: Si definisce **proiezione naturale al quoziente** l'applicazione:

$$\pi_{\mathcal{R}}: E \rightarrow E/\mathcal{R} \mid \pi_{\mathcal{R}}(x) = [x]$$

$\pi_{\mathcal{R}}$ è surgettiva poiché ogni classe di E/\mathcal{R} è immagine di tutti i suoi rappresentanti. Per lo stesso motivo non è iniettiva.

PROPOSIZIONE 1.1.4 (Leggi di De Morgan): Sia X un insieme e $A, B \subset X$. Allora, se $\bar{A} = X \setminus A$:

1) $\overline{(A \cup B)} = \bar{A} \cap \bar{B}$

2) $\overline{(A \cap B)} = \bar{A} \cup \bar{B}$

Dimostrazione:

1) $x \in \overline{(A \cup B)} \Leftrightarrow x \notin A \cup B \Leftrightarrow x \notin A \wedge x \notin B \Leftrightarrow x \in \bar{A} \wedge x \in \bar{B} \Leftrightarrow x \in \bar{A} \cap \bar{B}$.

2) Analoga.

PROPOSIZIONE 1.1.5: Sia $f: X \rightarrow Y$. Allora:

1) $\exists g: Y \rightarrow X \mid f \circ g = id_Y$ (**inversa destra**) $\Leftrightarrow f$ è surgettiva;

2) $\exists g: Y \rightarrow X \mid g \circ f = id_X$ (**inversa sinistra**) $\Leftrightarrow f$ è iniettiva;

3) g è unica sia in 1) che in 2).

Dimostrazione:

1) $\Rightarrow \forall y \in Y, y = f(g(y)) \Rightarrow$ ogni $y \in Y$ appartiene a $Im(f) \Rightarrow f$ è surgettiva

$\Leftarrow f$ surgettiva $\Rightarrow \forall y_0 \in Y \exists x_0 \in X \mid f(x_0) = y_0$.

Scelgo $g \mid g(y_0) = x_0 \forall y_0 \in Y$.

Allora $f(g(y_0)) = y_0 \forall y_0 \in Y \Rightarrow f \circ g = id_Y$.

2) Analoga.

3) g è fissata $\forall y_0 \in Y$, dunque è sicuramente unica.

1.2 STRUTTURE ALGEBRICHE

DEFINIZIONE 1.2.1: Dato un insieme A , si definisce **operazione** su A un'applicazione:

$$*: A \times A \rightarrow A$$

Esempio: la somma su \mathbb{Z} è definita come $+: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z} \mid + (x, y) = x + y$.

DEFINIZIONE 1.2.2: La coppia $(A, *)$, con A insieme e $*$ operazione su A , si chiama **gruppo** se valgono le seguenti proprietà:

- 1) **associativa**: $\forall a, b, c \in A, (a * b) * c = a * (b * c)$;
- 2) **dell'elemento neutro**: $\forall a \in A \exists e \in A \mid a * e = e * a = a$;
- 3) **dell'inverso**: $\forall a \in A \exists b \in A \mid a * b = b * a = e$.

Se vale anche la proprietà commutativa (cioè $\forall a, b \in A, a * b = b * a$), allora $(A, *)$ si dice **gruppo abeliano**.

Esempi: $(\mathbb{N}, +)$ non è un gruppo (non vale la proprietà dell'inverso).

$(\mathbb{Z}, +)$ è un gruppo abeliano.

(\mathbb{R}, \cdot) non è un gruppo perché non esiste l'inverso di 0.

$(\mathbb{R} \setminus \{0\}, \cdot)$ è un gruppo abeliano.

TEOREMA 1.2.1: Dato un gruppo $(A, *)$:

- 1) l'elemento neutro è unico
- 2) l'inverso di un elemento è unico
- 3) se $a, b, c \in A$ e $a * b = a * c$, allora $b = c$ (legge di cancellazione).

Dimostrazione:

1) Siano e_1, e_2 elementi neutri. Allora: $e_1 = e_1 * e_2 = e_2$, da cui $e_1 = e_2$.

2) Sia $a \in A$. Se a_1 e a_2 sono inversi di a allora:

$$a_1 = e * a_1 = (a_2 * a) * a_1 = a_2 * (a * a_1) = a_2 * e = a_2$$

da cui $a_1 = a_2$.

3) Se a^{-1} è l'inverso di a , allora $a * b = a * c \Rightarrow a^{-1} * a * b = a^{-1} * a * c \Rightarrow e * b = e * c \Rightarrow b = c$.

DEFINIZIONE 1.2.3: Siano $f, g: A \rightarrow B$. Si dice che $f = g \Leftrightarrow \forall x \in A, f(x) = g(x)$.

DEFINIZIONE 1.2.4: Siano $+, \cdot$ operazioni in A , con A insieme. La terna $(A, +, \cdot)$ si dice **anello** se:

- 1) $(A, +)$ è un gruppo abeliano;
- 2) (associativa di \cdot) $\forall a, b, c \in A, (ab)c = a(bc)$;
- 3) (elemento neutro per \cdot) $\exists 1 \mid \forall a \in A, a \cdot 1 = 1 \cdot a = a$;
- 4) (distributiva) $\forall a, b, c \in A, (a + b) \cdot c = ac + bc; a \cdot (b + c) = ab + ac$.

Se inoltre l'operazione \cdot è commutativa, cioè $\forall a, b \in A, a \cdot b = b \cdot a$, l'anello si dice **commutativo**.

Esempio: $(\mathbb{Z}, +, \cdot)$ è un anello commutativo.

DEFINIZIONE 1.2.5: $(A, +, \cdot)$ è un **campo** se:

- 1) $(A, +, \cdot)$ è un anello commutativo;
- 2) $\forall a \in A, a \neq 0$ (dove lo 0 rappresenta l'elemento neutro per la somma) $\exists b \in A \mid ab = ba = 1$.

Notazione: l'inverso rispetto alla somma a^{-1} si denota con $-a$.

Esempi: $(\mathbb{Q}, +, \cdot)$ è un campo.

$(\mathbb{R}, +, \cdot)$ è un campo.

$(\mathbb{Z}, +, \cdot)$ non è un campo perché $\forall a \neq 1 \nexists b \in \mathbb{Z} \mid ab = 1$.

PROPOSIZIONE 1.2.2: Sia $(A, +, \cdot)$ un anello. Allora:

- 1) $\forall a \in A, a \cdot 0 = 0 \cdot a = 0$;
- 2) $\forall a \in A, (-1) \cdot a = -a$ (-1 rappresenta l'inverso rispetto alla somma dell'elemento neutro per il prodotto).

Dimostrazione:

1) $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$.

Sommando da entrambe le parti l'inverso dell'elemento $a \cdot 0$:

$$a \cdot 0 - a \cdot 0 = a \cdot 0 + a \cdot 0 - a \cdot 0 \Rightarrow a \cdot 0 = 0.$$

2) Dobbiamo verificare che $(-1) \cdot a + a = 0$.

$$(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = ((-1) + 1) \cdot a = 0 \cdot a = 0.$$

PROPOSIZIONE 1.2.3: Sia $(\mathbb{K}, +, \cdot)$ un campo. Allora $ab = 0 \wedge a \neq 0 \Rightarrow b = 0$.

Dimostrazione:

$$\exists a^{-1}, \text{ dunque: } ab = 0 \Rightarrow a^{-1}ab = a^{-1} \cdot 0 = 0 \Rightarrow 1 \cdot b = 0 \Rightarrow b = 0.$$

Questo significa che in un campo non esistono **divisori di 0**, cioè, dato un $a \in \mathbb{K} \setminus \{0\}, \nexists b \in \mathbb{K} \setminus \{0\} \mid ab = 0$.

DEFINIZIONE 1.2.6: Definiamo l'insieme dei **numeri complessi** $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$, dove i è l'unità immaginaria tale che $i^2 = -1$.

DEFINIZIONE 1.2.7: Definiamo su \mathbb{C} una somma e un prodotto:

$$+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \mid (a + ib, c + id) \rightarrow (a + ib) + (c + id) \stackrel{\text{def}}{=} (a + c) + i(b + d);$$

$$\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C} \mid (a + ib, c + id) \rightarrow (a + ib) \cdot (c + id) \stackrel{\text{def}}{=} (ac - bd) + i(ad + bc).$$

DEFINIZIONE 1.2.8: $a + ib, c + id \in \mathbb{C}, a + ib = c + id \Leftrightarrow a = c \wedge b = d$.

PROPOSIZIONE 1.2.4: $(\mathbb{C}, +, \cdot)$ è un campo.

Dimostrazione:

- $(\mathbb{C}, +)$ è evidentemente un gruppo abeliano;
- \cdot è associativa;
- \cdot ha un elemento neutro, il numero $1 = 1 + 0i$;
- gode della proprietà distributiva:

$$((a + bi) + (c + di)) \cdot (e + if) = ((a + c) + i(b + d)) \cdot (e + if) = (ae + ce - bf - df) + i(af + cf + be + de),$$

$$(a + bi) \cdot (e + if) + (c + di) \cdot (e + if) = (ae - bf) + i(af + be) + (ce - df) + i(cf + de) = (ae - bf + ce - df) + i(af + be + cf + de);$$

- $(a + bi) \cdot (c + di) = (c + di) \cdot (a + bi) = (ac - bd) + i(ad + bc)$, dunque \cdot gode della proprietà commutativa;
- $\forall z = a + bi \in \mathbb{C}, \exists w \in \mathbb{C} \mid wz = zw = 1$.

Infatti poniamo $w = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}$. Allora:

$$zw = (a + bi) \left(\frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2} \right) = \frac{a^2 + b^2}{a^2 + b^2} + i \left(\frac{ab}{a^2 + b^2} - \frac{ab}{a^2 + b^2} \right) = 1$$

DEFINIZIONE 1.2.9: Sia \mathbb{K} un campo. Si definisce **polinomio nell'indeterminata x a coefficienti in \mathbb{K}** $p(x) = \sum_{i=0}^n a_i x^i$, con $a_i \in \mathbb{K} \forall 0 \leq i \leq n$.

DEFINIZIONE 1.2.10: Sia $\mathbb{K}[x]$ l'insieme dei polinomi in x a coefficienti in \mathbb{K} .

$$\mathbb{K}[x] = \left\{ p(x) = \sum_{i=0}^n a_i x^i \mid a_i \in \mathbb{K} \forall i \right\}$$

DEFINIZIONE 1.2.11: Due polinomi $p(x) = \sum_{i=0}^n a_i x^i, q(x) = \sum_{i=0}^n b_i x^i \in \mathbb{K}[x]$ si dicono uguali se $a_i = b_i \forall 0 \leq i \leq n$.

Notiamo che se gli esponenti massimi di $p(x)$ e $q(x)$ sono diversi, è sufficiente aggiungere termini del tipo $0 \cdot x^k$ per renderli uguali.

Notazione: Si denota con $0 \in \mathbb{K}[x]$ il polinomio con tutti i coefficienti nulli.

DEFINIZIONE 1.2.12: Dato $p(x) = \sum_{i=0}^n a_i x^i \in \mathbb{K}[x] \setminus \{0\}$, si definisce **grado** del polinomio $\deg(p(x)) = \max\{i \in \mathbb{N} \mid a_i \neq 0\}$.

DEFINIZIONE 1.2.13: Dati $p(x) = \sum_{i=0}^n a_i x^i$ e $q(x) = \sum_{i=0}^n b_i x^i \in \mathbb{K}[x]$, definiamo:

- $(p + q)(x) \stackrel{\text{def}}{=} \sum_{i=0}^n (a_i + b_i) x^i$;
- $(pq)(x) \stackrel{\text{def}}{=} \sum_{i=0}^{2n} c_i x^i$, dove $c_i = \sum_{j=0}^i a_j b_{i-j}$.

PROPOSIZIONE 1.2.5: $(\mathbb{K}[x], +, \cdot)$ è un anello commutativo ma non un campo.

Dimostrazione:

$(\mathbb{K}[x], +)$ è evidentemente un gruppo abeliano; inoltre valgono le proprietà associativa, distributiva e commutativa di \cdot perché valgono in \mathbb{K} ; \cdot ha l'elemento neutro $p(x) \equiv 1$.

Dunque è un anello commutativo.

Poiché $\deg(pq(x)) = \deg(p(x)) + \deg(q(x))$ (basta vedere che il coefficiente di grado massimo è il prodotto di due termini $\neq 0$), allora se esistesse $p^{-1}(x), 0 = \deg 1 = \deg(p(x)) + \deg(p^{-1}(x)) \Rightarrow \forall p(x) \mid \deg p > 0 \nexists p^{-1}(x) \Rightarrow (\mathbb{K}[x], +, \cdot)$ non è un campo.

TEOREMA 1.2.6 (di divisione in \mathbb{Z}): $\forall a, b \in \mathbb{Z} \wedge b \neq 0 \exists$ unici $q, r \in \mathbb{Z}$:

- $a = bq + r$;
- $0 \leq r < |b|$.

TEOREMA 1.2.7 (di divisione in $\mathbb{K}[x]$): $\forall a(x), b(x) \in \mathbb{K}[x] \setminus \{0\} \exists$ *unici* $q(x), r(x) \in \mathbb{K}[x]$:

- $a(x) = b(x)q(x) + r(x)$;
- $\deg(r(x)) < \deg(b(x))$.

Notiamo una similitudine fra la divisione in \mathbb{Z} e in $\mathbb{K}[x]$, poiché c'è una similitudine stretta fra la funzione valore assoluto e la funzione \deg .

Osservazione: se $r(x) = 0 \Rightarrow b(x)|a(x)$.

DEFINIZIONE 1.2.14: $a \in \mathbb{K}$ si dice **radice** di $p(x)$ se $p(a) = 0$.

TEOREMA 1.2.8 (di Ruffini): se a è radice di $p(x)$, allora $(x - a)|p(x)$.

Dimostrazione:

Applico il teorema di divisione; $\exists q(x), r(x) \in \mathbb{K}[x]$ tali che:

$$\begin{cases} p(x) = (x - a)q(x) + r(x) \\ \deg(r(x)) < \deg(x - a) = 1 \end{cases}, \text{ dunque } r(x) = \text{costante}.$$

Valuto in a :

$$0 = p(a) = (a - a)q(a) + r(a) = r(a), \text{ da cui segue la tesi.}$$

DEFINIZIONE 1.2.15: Sia a una radice di $p(x)$. Si definisce **molteplicità algebrica** della radice a il massimo numero naturale m tale che $(x - a)^m | p(x)$.

TEOREMA 1.2.9 (fondamentale dell'algebra): Ogni polinomio $p(x) \in \mathbb{C}[x]$ di grado $n \geq 1$ ha almeno una radice.

COROLLARIO 1.2.10: Ogni polinomio $p(x) \in \mathbb{C}[x]$ di grado $n \geq 1$ ha esattamente n radici (contate con molteplicità).

Dimostrazione:

Per induzione su n :

Passo base): $n = 1$, ovvio.

Passo induttivo): Per il teorema, so che $\exists a$ radice di $p(x)$, quindi per Ruffini $p(x) = (x - a)p_1(x)$, con $\deg(p_1(x)) = n - 1$, dunque per ipotesi induttiva $p_1(x)$ ha esattamente $n - 1$ radici contate con molteplicità. Dunque $p(x)$ ne ha n , da cui la tesi.

DEFINIZIONE 1.2.16: Un polinomio $p(x) \in \mathbb{K}[x]$ si dice **irriducibile** su $\mathbb{K}[x]$ se non può essere scritto come $p(x) = a(x)b(x)$, con $a(x), b(x) \in \mathbb{K}[x]$ non costanti.

Esempi: I polinomi di grado 1 sono sempre irriducibili.

$ax^2 + bx + c$ è riducibile su $\mathbb{R}[x]$ se ha radici (poiché la fattorizzazione di un polinomio di secondo grado può avvenire solo per mezzo di due polinomi di grado 1), quindi è riducibile $\Leftrightarrow \Delta \geq 0 \Leftrightarrow b^2 - 4ac \geq 0$.

$x^2 - 2$ è riducibile su $\mathbb{R}[x]$, ma non è riducibile su $\mathbb{Q}[x]$.

DEFINIZIONE 1.2.17: Si definisce l'operazione **prodotto per scalari** in $\mathbb{K}[x]$:

$\cdot : \mathbb{K} \times \mathbb{K}[x] \rightarrow \mathbb{K}[x] \mid (\alpha, p(x)) \rightarrow \alpha p(x)$, con $\alpha \in \mathbb{K}$ e $p(x) \in \mathbb{K}[x]$.

Se $p(x) = \sum_{i=0}^n a_i x^i$, allora $\alpha p(x) = \sum_{i=0}^n (\alpha a_i) x^i$.

PROPOSIZIONE 1.2.11: $(\mathbb{K}[x], +, \cdot)$, dove \cdot è il prodotto per scalari, è un anello (nel senso che $+$ e $\cdot \mid_{\mathbb{K} \times \mathbb{K}[x]}$ soddisfano le proprietà di anello)

Dimostrazione:

Poiché $(\mathbb{K}[x], +)$ è un gruppo abeliano, le restanti verifiche sono immediate.

Riprendiamo la notazione $S(X) = \{f: X \rightarrow X \mid f \text{ bigettiva}\}$, dove $X \neq \emptyset$ è un insieme.

PROPOSIZIONE 1.2.12: La funzione inversa f^{-1} di una funzione $f: X \rightarrow Y$ bigettiva è bigettiva.

Dimostrazione:

f^{-1} è iniettiva, poiché se non lo fosse due elementi del dominio sarebbero immagine di un solo elemento del codominio, quindi $f^{-1} \circ f \neq id_X$.

Poiché $f \circ f^{-1} = id$, allora $\forall x \in X, (f \circ f^{-1})(f(x)) = f(x)$, ma f è bigettiva, dunque $(f^{-1} \circ f)(x) = x \Rightarrow f$ è un'inversa destra, quindi f^{-1} è surgettiva.

PROPOSIZIONE 1.2.13: $(S(X), \circ)$ è un gruppo (in generale non abeliano).

Dimostrazione:

L'elemento neutro è evidentemente id_X , e poiché abbiamo visto che l'inversa di una funzione bigettiva è bigettiva, resta la banale verifica dell'associatività.

Abbiamo inoltre visto che in generale le funzioni non commutano, dunque ho la tesi.

PROPOSIZIONE 1.2.14: $(S(X), \circ)$ è un gruppo abeliano $\Leftrightarrow |X| \leq 2$.

Dimostrazione:

Sicuramente se $|X| = 1 \Rightarrow S(X) = \{id\} \Rightarrow (S(X), \circ)$ è gruppo abeliano.

Se $|X| = 2 \Rightarrow S(X) = \{id, f\}$, dove $f \circ f = id$, dunque $(S(X), \circ)$ è gruppo abeliano.

Se $|X| = 3$, supponiamo $X = \{a, b, c\}$. Sia $f: X \rightarrow X \mid f(a) = b, f(b) = a, f(c) = c$ e sia $g: X \rightarrow X \mid g(a) = b, g(b) = c, g(c) = a$.

Allora $(f \circ g)(a) = f(g(a)) = f(b) = a$; $(g \circ f)(a) = g(b) = c$, quindi per $|X| = 3$, $(S(X), \circ)$ non è un gruppo abeliano.

Per $|X| > 3$ il ragionamento è analogo, basta scegliere come controesempio il precedente esteso con l'identità agli altri elementi di X .

Osservazione: Ogni gruppo G di due elementi è abeliano. La dimostrazione è analoga alla precedente.

Osservazione: Esistono campi con un numero finito di elementi.

Prendiamo $\mathbb{F} = \{[0]_3, [1]_3, [2]_3\}$, dove $[a]_3$ è la classe di resto a nella divisione per 3.

Definendo $[a]_3 \cdot [b]_3 = [ab]_3$ e $[a]_3 + [b]_3 = [a + b]_3$, non è difficile mostrare che $(\mathbb{F}, +, \cdot)$ è un campo.

Notazione: Sia \mathbb{K} un campo e $n \in \mathbb{N}, n \geq 1$. Si denota con:

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \dots \times \mathbb{K}}_{n \text{ volte}}$$

il prodotto cartesiano di \mathbb{K} per se stesso n volte.

Perciò $\mathbb{K}^n = \{(x_1, \dots, x_n) | x_i \in \mathbb{K} \forall i\}$.

DEFINIZIONE 1.2.18: Definiamo una somma e un prodotto per scalari su \mathbb{K}^n :

$+$: $\mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n | (x_1, \dots, x_n) + (y_1, \dots, y_n) \stackrel{\text{def}}{=} (x_1 + y_1, \dots, x_n + y_n)$;

\cdot : $\mathbb{K} \times \mathbb{K}^n \rightarrow \mathbb{K}^n | \alpha \cdot (x_1, \dots, x_n) \stackrel{\text{def}}{=} (\alpha x_1, \dots, \alpha x_n)$.

PROPOSIZIONE 1.2.15: $(\mathbb{K}^n, +, \cdot)$ è un anello (nello stesso senso della PROPOSIZIONE 1.2.11).

Dimostrazione:

Le semplici verifiche sono lasciate al lettore.

PROPOSIZIONE 1.2.16: Siano $f(t), g(t)$ polinomi in $\mathbb{R}[t]$, con $f(t) \neq 0$. Sia $h(t)$ un polinomio in $\mathbb{C}[t] | f(t) = g(t) \cdot h(t)$ in $\mathbb{C}[t]$, allora $h(t) \in \mathbb{R}[t]$.

Dimostrazione 1:

Siano $f(t) = a_n t^n + \dots + a_0, g(t) = b_m t^m + \dots + b_0, h(t) = c_l t^l + \dots + c$, $a_n \neq 0, b_m \neq 0, c_l \neq 0$.

La nostra ipotesi è che $a_i, b_j \in \mathbb{R} \forall i, j$.

Mostriamo con l'induzione II su i che $c_{l-i} \in \mathbb{R}$.

Passo base): Poiché $f(t) = g(t) \cdot h(t)$ e $b_m \neq 0 \Rightarrow c_l = \frac{a_n}{b_m} \in \mathbb{R}$.

Passo induttivo): Mostriamo che se $c_{l-i} \in \mathbb{R} \forall 0 \leq i \leq k \Rightarrow c_{l-k-1} \in \mathbb{R}$.

$a_{n-k-1} = b_m c_{l-k-1} + b_{m-1} c_{l-k} + \dots + b_{m-k-1} c_l$, perciò $c_{l-k-1} = \frac{a_{n-k-1} - b_{m-1} c_{l-k} - \dots - b_{m-k-1} c_l}{b_m}$.

Ma tutti i termini della frazione $\in \mathbb{R}$, dunque ho la tesi.

Dimostrazione 2: Poiché in $\mathbb{C}[t]$ ho che $f(t) = g(t) \cdot h(t)$ e $f(t) \neq 0$, allora $g(t) \neq 0$.

Dunque divido $f(t)$ per $g(t)$ in $\mathbb{R}[t]$:

$f(t) = g(t) \cdot q(t) + r(t)$ in $\mathbb{R}[t]$ e dunque in $\mathbb{C}[t]$.

$\deg r(t) < \deg g(t)$

Allora $g(t)h(t) = g(t)q(t) + r(t)$ in $\mathbb{C}[t] \Rightarrow g(t)(h(t) - q(t)) = r(t)$ in $\mathbb{C}[t]$, ma

$\deg r(t) < \deg g(t)$ e poiché il grado è additivo, allora $h(t) - q(t) = 0 \Rightarrow h(t) = q(t) \in \mathbb{R}[t]$.

DEFINIZIONE 1.2.18: Sia $c: \mathbb{C} \rightarrow \mathbb{C}$ l'applicazione definita in modo tale che associ ad ogni numero complesso $z = a + bi$ il suo coniugato $\bar{z} = a - bi$. c prende il nome di **coniugio** ed è evidentemente biunivoca.

PROPOSIZIONE 1.2.17: 1) $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$

2) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$

3) $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$

4) $\overline{\bar{z}} = z$

5) $z + \bar{z} \in \mathbb{R}, z \cdot \bar{z} \in \mathbb{R}$.

Dimostrazione:

Queste semplici verifiche sono lasciate al lettore per esercizio.

COROLLARIO 1.2.18: 1) $\overline{\sum_{i=1}^n z_i} = \sum_{i=1}^n \overline{z_i}$

2) $\overline{\prod_{i=1}^n z_i} = \prod_{i=1}^n \overline{z_i}$

3) $\overline{z^n} = \overline{z}^n$

Dimostrazione:

1) Per induzione su $n \geq 2$:

Passo base): $n = 2$, già visto;

Passo induttivo): $\overline{\sum_{i=1}^n z_i} = \overline{\sum_{i=1}^{n-1} z_i + z_n} = (\text{ip. ind.}) = \sum_{i=1}^{n-1} \overline{z_i} + \overline{z_n} = \sum_{i=1}^n \overline{z_i}$.

2) Analoga.

3) È un caso particolare del 2) con $z_1 = \dots = z_n = z$.

PROPOSIZIONE 1.2.18: Sia $f(t) \in \mathbb{R}[t] \setminus \{0\}$ e sia α un numero complesso non reale. Allora, se α è radice di $f(t)$:

1) anche $\overline{\alpha}$ è radice di $f(t)$;

2) la molteplicità algebrica di α è uguale a quella di $\overline{\alpha}$.

Dimostrazione:

1) Poiché α è radice di $f(t)$, allora $f(\alpha) = \sum_{i=0}^n a_i \alpha^i = 0$. Allora:

$$0 = a_n \alpha^n + \dots + a_0 = \overline{a_n \alpha^n + \dots + a_0} = \overline{a_n \alpha^n + \dots + a_0} = (\text{poichè } a_i \in \mathbb{R}) = a_n \overline{\alpha}^n + \dots + a_0 = a_n \overline{\alpha}^n + \dots + a_0 = f(\overline{\alpha}).$$

2) Dimostriamolo per induzione II su $n = \deg f(t)$:

Passo base): $n = 0 \Rightarrow f(t) = \beta \Rightarrow f(t)$ non ha radici, quindi molteplicità algebrica di $\alpha =$ molteplicità algebrica di $\overline{\alpha} = 0$.

Passo induttivo): Supponiamo che l'enunciato sia vero $\forall k \leq n$ e dimostriamo che è vero per $n + 1$.

Se né α né $\overline{\alpha}$ sono radice, ho la tesi.

Altrimenti $f(\alpha) = 0$ e $f(\overline{\alpha}) = 0$.

Dunque $f(t) = (t - \alpha)g(t)$ per Ruffini in $\mathbb{C}[t]$.

Valutando in $\overline{\alpha}$:

$$0 = f(\overline{\alpha}) = (\overline{\alpha} - \alpha)g(\overline{\alpha}), \text{ ma } \alpha \neq \overline{\alpha}, \text{ dunque } g(\overline{\alpha}) = 0 \text{ e quindi } (t - \overline{\alpha})|g(t)|f(t).$$

$$\text{Per cui } f(t) = (t - \alpha)(t - \overline{\alpha})h(t) = (t^2 - (\alpha + \overline{\alpha})t + \alpha\overline{\alpha})h(t).$$

Sappiamo che $\alpha + \overline{\alpha}, \alpha\overline{\alpha} \in \mathbb{R}$, dunque per la proposizione 1.2.16 so che $h(t) \in \mathbb{R}[t]$.

Applicando l'ipotesi induttiva a $h(t)$, vediamo che la molteplicità algebrica μ di α e $\overline{\mu}$ di $\overline{\alpha}$ in $h(t)$ coincidono, ma le loro molteplicità in $f(t)$ sono semplicemente $\mu + 1$ e $\overline{\mu} + 1$, che quindi coincidono.

2 SPAZI VETTORIALI E APPLICAZIONI LINEARI

2.1 SPAZI VETTORIALI

DEFINIZIONE 2.1.1: Un \mathbb{K} -spazio vettoriale è una quaterna $(V, +, \cdot, \mathbb{K})$, dove \mathbb{K} è un campo e V insieme $\neq \emptyset$.

$+$: $V \times V \rightarrow V$ \cdot : $\mathbb{K} \times V \rightarrow V$ tali che:

- 1) $(V, +)$ è un gruppo abeliano;
- 2) $\forall \alpha, \beta \in \mathbb{K}, \forall x \in V, (\alpha\beta)x = \alpha(\beta x)$;
- 3) $\forall \alpha, \beta \in \mathbb{K}, \forall x \in V, (\alpha + \beta)x = \alpha x + \beta x$;
- 4) $\forall \alpha \in \mathbb{K}, \forall x, y \in V, \alpha(x + y) = \alpha x + \alpha y$;
- 5) $\forall x \in V, 1 \cdot x = x$.

PROPOSIZIONE 2.1.1: Sia $(V, +, \cdot, \mathbb{K})$ un \mathbb{K} -spazio vettoriale. Allora:

- 1) 0 è unico;
- 2) $\forall x \in V, -x$ è unico;
- 3) $\forall x \in V, 0 \cdot x = 0$;
- 4) $\forall \alpha \in \mathbb{K}, \alpha \cdot 0 = 0$
- 5) $\forall \alpha \in \mathbb{K}, \forall x \in V, \alpha x = 0 \Rightarrow \alpha = 0 \vee x = 0$;
- 6) $\forall x \in V, (-1) \cdot x = -x$.

Dimostrazione:

1), 2) derivano dal fatto che $(V, +)$ è un gruppo abeliano.

3) $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x \Rightarrow 0 \cdot x = 0$

4) analoga alla 3)

5) Se $\alpha = 0$, abbiamo subito la tesi.

Se $\alpha \neq 0 \Rightarrow 0 = \alpha^{-1} \cdot 0 = \alpha^{-1} \cdot \alpha \cdot x = 1 \cdot x = x$

6) $x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0$

DEFINIZIONE 2.1.2: Ogni elemento di uno spazio vettoriale si definisce **vettore**.

Notazione: Al posto di $x + (-y)$ scriveremo $x - y$.

PROPOSIZIONE 2.1.2: \mathbb{K}^n è un \mathbb{K} -spazio vettoriale $\forall n \geq 1$.

Dimostrazione:

La verifica è lasciata al lettore.

Osservazioni: • Se $V = \mathbb{K}$, il prodotto per scalari è definito $\cdot : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$, dove il primo \mathbb{K} rappresenta il campo degli scalari, mentre il secondo lo spazio vettoriale.

- Per l'osservazione precedente \mathbb{C} è un \mathbb{C} -spazio vettoriale. Consideriamo ora una restrizione dell'operazione prodotto per scalari su \mathbb{R} , cioè $\cdot : \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$.

Poiché $\mathbb{R} \subset \mathbb{C}$, le definizioni che valgono su \mathbb{C} valgono anche su \mathbb{R} ; perciò \mathbb{C} è un \mathbb{R} -spazio vettoriale.

In generale possiamo effettuare una restrizione del campo degli scalari, cioè se \mathbb{K}' è sottocampo di \mathbb{K} ($\mathbb{K}' \subset \mathbb{K}$ e \mathbb{K}' è chiuso rispetto a $+$ e \cdot), ogni \mathbb{K} -spazio vettoriale è anche un \mathbb{K}' -spazio vettoriale.

- Fissiamo nel piano due assi cartesiani. Allora la funzione:

$Piano \rightarrow \mathbb{R}^2$
 $P \rightarrow (x_P, y_P)$ è biunivoca.

Inoltre la funzione:

$Piano \rightarrow$ *vettori uscenti da O*
 $P \rightarrow \overrightarrow{OP}$ è biunivoca, dunque possiamo identificare P con (x_P, y_P) e con \overrightarrow{OP} . Quindi la somma in \mathbb{R}^2 corrisponde alla regola del parallelogramma.

2.2 SPAZI DI MATRICI

DEFINIZIONE 2.2.1: Definiamo l'insieme $\mathcal{M}(p, n, \mathbb{K})$ come l'insieme delle **matrici** $p \times n$, cioè con p righe e n colonne, a coefficienti in \mathbb{K} .

DEFINIZIONE 2.2.2: Una matrice si dice **quadrata** se $p = n$ e l'insieme delle matrici $n \times n$ si indica con $\mathcal{M}(n, \mathbb{K})$ (o più semplicemente con $\mathcal{M}(n)$).

Notazioni: • Per indicare l'elemento di posto i, j della matrice A si usa il simbolo $[A]_{ij}$;

- A_i indica l' i -esima riga di A ;
- A^j indica la j -esima colonna di A ;
- 0 rappresenta la matrice nulla, cioè $[0]_{ij} = 0 \forall i, j$.

DEFINIZIONE 2.2.3: Siano $A, B \in \mathcal{M}(p, n, \mathbb{K})$ matrici. Allora si dice che A e B sono uguali se $[A]_{ij} = [B]_{ij} \forall i, j$.

DEFINIZIONE 2.2.4: $\forall A, B \in \mathcal{M}(p, n, \mathbb{K})$, poniamo:

$$[A + B]_{ij} \stackrel{\text{def}}{=} [A]_{ij} + [B]_{ij} \quad \forall i, j;$$

$$[\alpha A]_{ij} \stackrel{\text{def}}{=} \alpha [A]_{ij} \quad \forall i, j.$$

PROPOSIZIONE 2.2.1: $\mathcal{M}(p, n, \mathbb{K})$ è un \mathbb{K} -spazio vettoriale.

Dimostrazione:

Le verifiche sono immediate.

Osservazione: L'applicazione:

$$\begin{aligned} \mathbb{K}^n &\rightarrow \mathcal{M}(n, 1, \mathbb{K}) \\ (x_1, \dots, x_n) &\rightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

è bigettiva, quindi scriveremo indifferentemente un vettore di \mathbb{K}^n come n -upla o come colonna.

DEFINIZIONE 2.2.5: $A \in \mathcal{M}(n)$ si dice:

- **diagonale** se $[A]_{ij} = 0 \quad \forall i \neq j$;
- **simmetrica** se $[A]_{ij} = [A]_{ji} \quad \forall i, j$;
- **antisimmetrica** se $[A]_{ij} = -[A]_{ji} \quad \forall i, j$ (dunque $[A]_{ii} = 0 \quad \forall i$);
- **triangolare superiore** se $[A]_{ij} = 0 \quad \forall i > j$.

Notazione: Denoteremo:

- $\mathcal{D}(n) = \{A \in \mathcal{M}(n) | A \text{ è diagonale}\}$;
- $\mathcal{S}(n) = \{A \in \mathcal{M}(n) | A \text{ è simmetrica}\}$;
- $\mathcal{A}(n) = \{A \in \mathcal{M}(n) | A \text{ è antisimmetrica}\}$;
- $\mathcal{T}(n) = \{A \in \mathcal{M}(n) | A \text{ è triangolare superiore}\}$.

PROPOSIZIONE 2.2.2: Ogni spazio di polinomi $\mathbb{K}[x]$ è un \mathbb{K} -spazio vettoriale.

DEFINIZIONE 2.2.6: Sia A un insieme e V un \mathbb{K} -spazio vettoriale. Definiamo

$\mathcal{F}(A, V) = \{f: A \rightarrow V\}$ tale che $\forall f, g \in \mathcal{F}(A, V), \forall \alpha \in \mathbb{K}$:

$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x) \quad \forall x \in A$;

$(\alpha f)(x) \stackrel{\text{def}}{=} \alpha f(x) \quad \forall x \in A$.

PROPOSIZIONE 2.2.3: $\mathcal{F}(A, V)$ è uno spazio vettoriale di funzioni.

Osservazioni: • $\mathcal{F}(\mathbb{N}, \mathbb{K}) = \{\text{successioni a valori in } \mathbb{K}\}$

- $\mathcal{F}(\{1, \dots, n\}, \mathbb{K}) = \mathbb{K}^n$
- $\mathcal{F}(\{1, \dots, p\} \times \{1, \dots, q\}, \mathbb{K}) = \mathcal{M}(p, q, \mathbb{K})$.

2.3 SOTTOSPAZI E COMBINAZIONI LINEARI

DEFINIZIONE 2.3.1: Dato V \mathbb{K} -spazio vettoriale, $W \subset V$ si dice **sottospazio vettoriale** di V se:

- 1) $0_V \in W$;
- 2) $\forall x, y \in W, x + y \in W$ (cioè W è chiuso rispetto alla somma);
- 3) $\forall \alpha \in \mathbb{K}, \forall x \in W, \alpha x \in W$ (cioè W è chiuso rispetto al prodotto per scalari).

Quindi, poiché se le 8 proprietà di spazio vettoriale valgono per V , allora valgono anche per W e poiché $+$ e \cdot sono chiusi rispetto a W , allora W è uno spazio vettoriale.

PROPOSIZIONE 2.3.1: $\mathcal{D}(n), \mathcal{S}(n), \mathcal{A}(n), \mathcal{T}(n)$ sono sottospazi vettoriali di $\mathcal{M}(n)$.

Dimostrazione:

Dimostriamolo per $\mathcal{D}(n)$, per gli altri il procedimento è analogo.

- 1) $0 \in \mathcal{D}(n)$, poiché $[0]_{ij} = 0 \quad \forall i \neq j$;
- 2), 3) Evidentemente, se $A, B \in \mathcal{D}(n)$, allora $A + B \in \mathcal{D}(n)$ e $\alpha A \in \mathcal{D}(n)$.

Notazione: Fissato $m \in \mathbb{N}$, si denoti con $\mathbb{K}_m[x] = \{p(x) \in \mathbb{K}[x] | \deg p(x) \leq m\}$ l'insieme dei polinomi di $\mathbb{K}[x]$ di grado $\leq m$.

PROPOSIZIONE 2.3.2: $\forall m \in \mathbb{N}$, $\mathbb{K}_m[x]$ è sottospazio vettoriale di $\mathbb{K}[x]$.

Dimostrazione:

La verifica è lasciata al lettore.

PROPOSIZIONE 2.3.3: Le rette in \mathbb{R}^2 per l'origine sono sottospazi vettoriali di \mathbb{R}^2 .

Le rette e i piani per l'origine in \mathbb{R}^3 sono sottospazi vettoriali di \mathbb{R}^3 .

Dimostrazione:

Semplice verifica.

PROPOSIZIONE 2.3.4: Se $\{W_i\}_{i \in I}$ è una famiglia arbitraria di sottospazi vettoriali di V , allora $\bigcap_{i \in I} W_i$ è sottospazio vettoriale di V .

Dimostrazione:

1) $0_V \in W_i \forall i \in I$, perciò $0_V \in \bigcap_{i \in I} W_i$.

2), 3) Se $+$ e \cdot sono chiusi in $W_i \forall i \in I$, a maggior ragione saranno chiusi in $\bigcap_{i \in I} W_i$.

DEFINIZIONE 2.3.2: Dati $v_1, \dots, v_n \in V$ e $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, si definisce **combinazione lineare** dei v_1, \dots, v_n il vettore $\alpha_1 v_1 + \dots + \alpha_n v_n \in V$.

DEFINIZIONE 2.3.3: Dato $S \subset V$, denotiamo con:

$Span(S) \stackrel{\text{def}}{=} \{v \in V \mid \exists v_1, \dots, v_n \in S, \exists \alpha_1, \dots, \alpha_n \in \mathbb{K} \text{ per cui } v = \alpha_1 v_1 + \dots + \alpha_n v_n\}$.

Esempio: $V = \mathbb{R}^2$, $S = \{(1,1)\}$.

$Span(S) = \{a(1,1) \mid a \in \mathbb{R}\}$, cioè $Span(S)$ è semplicemente la retta passante per l'origine e per $(1,1)$ (che è quindi uno spazio vettoriale).

PROPOSIZIONE 2.3.5: 1) $Span(S)$ è sottospazio vettoriale di $V \quad \forall S \subset V$.

2) $S \subset Span(S)$

3) Se W è sottospazio vettoriale di $V \mid S \subset W \subset Span(S) \Rightarrow W = Span(S)$.

Dimostrazione:

1) Semplice verifica.

2) Ovvio.

3) Sappiamo che $W \subset Span(S)$ per ipotesi, dunque basta dimostrare che $Span(S) \subset W$.

Se $v \in Span(S) \Rightarrow v = \alpha_1 v_1 + \dots + \alpha_n v_n$ per certi $v_1, \dots, v_n \in S, \alpha_1, \dots, \alpha_n \in \mathbb{K}$.

Ma $v_1, \dots, v_n \in S \subset W$, dunque, poiché W è sottospazio vettoriale,

$v = \alpha_1 v_1 + \dots + \alpha_n v_n \in W$, da cui segue immediatamente la tesi.

PROPOSIZIONE 2.3.6: $Span(S) = \bigcap_{\substack{W \text{ ssv di } V \\ W \supseteq S}} W$.

Dimostrazione:

Evidentemente $S \subset \bigcap_{\substack{W \text{ ssv di } V \\ W \supseteq S}} W$, poiché interseco insiemi che contengono S .

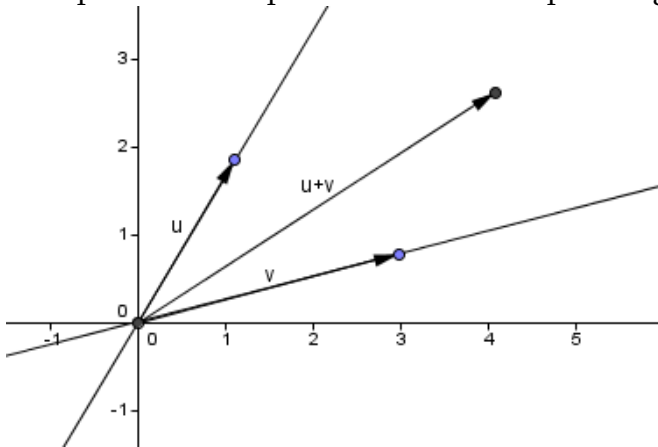
Inoltre $\bigcap_{\substack{W \text{ ssv di } V \\ W \supseteq S}} W \subset Span(S)$, poiché fra i W che interseco c'è anche $Span(S)$, dunque

l'intersezione sarà sicuramente "più piccola" di $Span(S)$.

Grazie alla proposizione precedente, ho la tesi.

Osservazione: In generale l'unione di sottospazi vettoriali non è un sottospazio vettoriale.

Esempio: $V = \mathbb{R}^2$ e prendiamo due rette per l'origine distinte come sottospazi.



Vediamo che evidentemente la somma non è chiusa, dunque $u \cup v$ non è sottospazio di \mathbb{R}^2 .

DEFINIZIONE 2.3.4: Siano U, W sottospazi vettoriali di V . Definiamo l'**insieme somma**:

$U + W \stackrel{\text{def}}{=} \{x \in V \mid \exists u \in U, \exists w \in W \text{ t.c. } x = u + w\}$.

Osservazione: Se ad esempio U e W sono due rette per l'origine distinte, con combinazioni lineari di vettori su di esse posso individuare qualsiasi altro vettore di \mathbb{R}^2 , semplicemente scomponendolo nelle due componenti. Dunque $U + W = \mathbb{R}^2$.

PROPOSIZIONE 2.3.7: $U + W$ è sottospazio vettoriale di V ed è il più piccolo sottospazio contenente U e W .

Dimostrazione:

Evidentemente $U + W$ è sottospazio vettoriale.

Ovviamente $U \subset U + W$, poiché $\forall u \in U, u = u + 0 \in U + W$; analogamente $W \subset U + W$.

Prendiamo Z sottospazio di V tale che $U \subset Z$ e $W \subset Z$ e mostriamo che $U + W \subset Z$.

Infatti $\forall u \in U, \forall w \in W, u, w \in Z$, ma Z è sottospazio vettoriale $\Rightarrow u + w \in Z \Rightarrow$ tesi.

Osservazioni: • In \mathbb{R}^2 , $\text{retta} + \text{retta} = \begin{cases} \text{se stessa se collineari;} \\ \mathbb{R}^2 \text{ altrimenti} \end{cases}$;

• In \mathbb{R}^3 , $\text{retta} + \text{retta} = \text{piano che le contiene}$;

• In \mathbb{R}^3 , $\text{piano} + \text{retta} = \mathbb{R}^3$.

DEFINIZIONE 2.3.5: Se $U \cap W = \{0\}$, la somma $U + W$ si denota con $U \oplus W$ e prende il nome di **somma diretta**.

PROPOSIZIONE 2.3.8: Ogni vettore in $U \oplus W$ si scrive in modo unico come $u + w$, con $u \in U$ e $w \in W$.

Dimostrazione:

Supponiamo che il vettore v si scriva $v = u_1 + w_1$ e $v = u_2 + w_2$. Allora:

$u_1 + w_1 = u_2 + w_2 \Rightarrow u_1 - u_2 = w_2 - w_1$, ma $u_1 - u_2 \in U$ e $w_2 - w_1 \in W$, perciò $u_1 - u_2 = w_2 - w_1 \in U \cap W = \{0\} \Rightarrow u_1 = u_2$ e $w_1 = w_2$.

DEFINIZIONE 2.3.6: Se $V = U \oplus W$, con U e W sottospazi vettoriali di V , sono ben definite le applicazioni:

$$\pi_U: \begin{array}{ccc} V & \rightarrow & U \\ v = u + w & \rightarrow & u \end{array}; \quad \pi_W: \begin{array}{ccc} V & \rightarrow & W \\ v = u + w & \rightarrow & w \end{array}$$

dette **proiezioni** di V su U e su W .

Esempio: In \mathbb{R}^2 siano U, W gli assi cartesiani. Allora se $v = (x, y)$, semplicemente $\pi_U(v) = x$ e $\pi_W(v) = y$.

DEFINIZIONE 2.3.7: Sia U un sottospazio vettoriale di V . Si definisce **supplementare** di U ogni sottospazio W di V tale che $V = U \oplus W$.

Osservazione: Il supplementare non è unico, ad esempio in \mathbb{R}^2 il supplementare di una retta per l'origine è una qualsiasi altra retta per l'origine di \mathbb{R}^2 .

2.4 APPLICAZIONI LINEARI

DEFINIZIONE 2.4.1: Siano V, W \mathbb{K} -spazi vettoriali. $f: V \rightarrow W$ si dice **\mathbb{K} -lineare** (o semplicemente **lineare**) se:

- 1) $\forall x, y \in V, f(x + y) = f(x) + f(y)$;
- 2) $\forall \alpha \in \mathbb{K}, \forall x \in V, f(\alpha x) = \alpha f(x)$.

Osservazione: Se f è lineare $\Rightarrow f(0) = 0$.

Infatti $f(0) = f(0 + 0) = f(0) + f(0) \Rightarrow f(0) = 0$.

DEFINIZIONE 2.4.2: Definiamo l'applicazione **trasposta**:

$${}^t: \mathcal{M}(p, q, \mathbb{K}) \rightarrow \mathcal{M}(q, p, \mathbb{K}) \mid [{}^t A]_{ij} = [A]_{ji} \quad \forall i, j$$

DEFINIZIONE 2.4.3: Definiamo l'applicazione **traccia**:

$$tr: \mathcal{M}(n) \rightarrow \mathbb{K} \mid tr(A) = \sum_{i=1}^n [A]_{ii}$$

DEFINIZIONE 2.4.4: Definiamo l'applicazione **valutazione** in $a \in \mathbb{K}$:

$$v_a: \mathbb{K}[x] \rightarrow \mathbb{K} \mid v_a(p(x)) = p(a)$$

PROPOSIZIONE 2.4.1: Le seguenti applicazioni sono lineari:

- 1) l'applicazione nulla $0: V \rightarrow W \mid f(v) = 0 \quad \forall v \in V$;
- 2) l'applicazione identica;
- 3) l'applicazione trasposta;

- 4) l'applicazione traccia;
- 5) la valutazione;
- 6) le proiezioni indotte dalla scomposizione $V = U \oplus W$.

Dimostrazione:

Mostriamo che la 3) è lineare, per le altre il ragionamento è analogo.

$$\forall A, B \in \mathcal{M}(p, q, \mathbb{K}), \quad [{}^t(A+B)]_{ij} = [A+B]_{ji} = [A]_{ji} + [B]_{ji} = [{}^tA]_{ij} + [{}^tB]_{ij};$$

$$\forall \alpha \in \mathbb{K}, \forall A \in \mathcal{M}(p, q, \mathbb{K}), \quad [{}^t(\alpha A)]_{ij} = [\alpha A]_{ji} = \alpha [A]_{ji} = \alpha [{}^tA]_{ij}.$$

PROPOSIZIONE 2.4.2: $\mathcal{M}(n, \mathbb{R}) = \mathcal{S}(n, \mathbb{R}) \oplus \mathcal{A}(n, \mathbb{R})$.

Dimostrazione:

Sia $C \in \mathcal{M}(n, \mathbb{R})$. Poniamo $S = \frac{C+{}^tC}{2}$ e $A = \frac{C-{}^tC}{2}$.

Vediamo che:

$${}^tS = {}^t\left(\frac{C+{}^tC}{2}\right) = \frac{1}{2}({}^tC + {}^t{}^tC) = \frac{1}{2}({}^tC + C) = S \Rightarrow S \in \mathcal{S}(n, \mathbb{R})$$

$${}^tA = {}^t\left(\frac{C-{}^tC}{2}\right) = \frac{1}{2}({}^tC - {}^t{}^tC) = \frac{1}{2}({}^tC - C) = -A \Rightarrow A \in \mathcal{A}(n, \mathbb{R})$$

$$S + A = \frac{C+{}^tC}{2} + \frac{C-{}^tC}{2} = C$$

Poiché evidentemente $\mathcal{S}(n, \mathbb{R}) \cap \mathcal{A}(n, \mathbb{R}) = \{0\}$, ho la tesi.

PROPOSIZIONE 2.4.3: L'applicazione coniugio è \mathbb{R} -lineare (ma non \mathbb{C} -lineare).

Dimostrazione:

Sicuramente $\forall z, w \in \mathbb{C}, \overline{z+w} = \overline{z} + \overline{w}$; inoltre $\forall \alpha \in \mathbb{R}, \forall z \in \mathbb{C}, \overline{\alpha z} = \alpha \overline{z}$, poiché $\alpha = \overline{\alpha}$.
Se invece prendiamo \mathbb{C} come campo di scalari, in generale $\overline{\alpha z} \neq \alpha \overline{z}$.

DEFINIZIONE 2.4.5: Sia \mathbb{K} un campo. Definiamo la **caratteristica** $char(\mathbb{K})$ del campo:

- se $\forall n \in \mathbb{N}, n \cdot 1 \neq 0 \Rightarrow char(\mathbb{K}) = 0$;
- se $\exists n \in \mathbb{N} | n \cdot 1 = 0 \Rightarrow char(\mathbb{K}) = \min\{p \in \mathbb{N} | p \cdot 1 = 0\}$.

Osservazione: È vero che $\mathcal{M}(n, \mathbb{K}) = \mathcal{S}(n, \mathbb{K}) \oplus \mathcal{A}(n, \mathbb{K})$ per qualsiasi campo \mathbb{K} ?

Vediamo che $A \in \mathcal{S}(n, \mathbb{R}) \cap \mathcal{A}(n, \mathbb{R}) \Leftrightarrow [A]_{ij} = -[A]_{ij} \quad \forall i, j \Leftrightarrow 2[A]_{ij} = 0 \quad \forall i, j$.

Questo implica $A = 0$ solamente se $char(\mathbb{K}) \neq 2$.

In questo caso vediamo anche che ha senso dividere per 2 nella prima parte della dimostrazione, dunque possiamo affermare che $\mathcal{M}(n, \mathbb{K}) = \mathcal{S}(n, \mathbb{K}) \oplus \mathcal{A}(n, \mathbb{K}) \Leftrightarrow char(\mathbb{K}) \neq 2$.

Infatti prendiamo un campo $\mathbb{F}_2 | char(\mathbb{F}_2) = 2$, ad esempio $\mathbb{F}_2 = \{[0]_2, [1]_2\}$, dove $[a]_2$ è la classe di resto a modulo 2:

$$\begin{pmatrix} [1]_2 & [1]_2 \\ [1]_2 & [1]_2 \end{pmatrix} + \begin{pmatrix} [1]_2 & [1]_2 \\ [1]_2 & [1]_2 \end{pmatrix} = \begin{pmatrix} [0]_2 & [0]_2 \\ [0]_2 & [0]_2 \end{pmatrix},$$

e $M = \begin{pmatrix} [1]_2 & [1]_2 \\ [1]_2 & [1]_2 \end{pmatrix} \in \mathcal{S}(2, \mathbb{F}_2) \cap \mathcal{A}(2, \mathbb{F}_2)$ (poiché $[1]_2 = -[1]_2$) e $M \neq 0$.

DEFINIZIONE 2.4.6: Siano V, W \mathbb{K} -spazi vettoriali. Allora definiamo l'**insieme degli omomorfismi**:

$$\text{Hom}(V, W) \stackrel{\text{def}}{=} \{f: V \rightarrow W \mid f \text{ è lineare}\} \subset \mathcal{F}(V, W)$$

PROPOSIZIONE 2.4.4: $\text{Hom}(V, W)$ è sottospazio vettoriale di $\mathcal{F}(V, W)$.

DEFINIZIONE 2.4.7: Sia $f \in \text{Hom}(V, W)$. Si definisce **kernel** (o **nucleo**) di f :

$$\text{Ker}(f) \stackrel{\text{def}}{=} \{x \in V \mid f(x) = 0\}$$

PROPOSIZIONE 2.4.5: Sia $f \in \text{Hom}(V, W)$. Allora:

- 1) $\text{Ker}(f)$ è sottospazio vettoriale di V ;
- 2) $\text{Im}(f)$ è sottospazio vettoriale di W ;
- 3) f è iniettiva $\Leftrightarrow \text{Ker}(f) = \{0\}$.

Dimostrazione:

1), 2) ovvie.

3) \Rightarrow) Sia $x \in \text{Ker}(f)$. Allora $f(x) = 0 = f(0)$, in quanto f è lineare. Ma f è iniettiva $\Rightarrow x = 0$.

\Leftarrow) Per dimostrare che f è iniettiva, dobbiamo mostrare che se $f(x) = f(y) \Rightarrow x = y$.

Prendiamo $f(x) = f(y)$. Allora per linearità $f(x) - f(y) = f(x - y) = 0$, quindi $x - y \in \text{Ker}(f) = \{0\}$. Dunque $x - y = 0 \Rightarrow x = y$.

DEFINIZIONE 2.4.8: Siano $(a_1 \dots a_n) \in \mathcal{M}(1, n, \mathbb{K})$ e $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{M}(n, 1, \mathbb{K})$. Si definisce **prodotto fra la riga e la colonna**:

$$(a_1 \dots a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \stackrel{\text{def}}{=} a_1 b_1 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$$

DEFINIZIONE 2.4.9: Sia $A \in \mathcal{M}(p, n)$, $x \in \mathcal{M}(n, 1)$. Si definisce **prodotto fra la matrice e la colonna**:

$$A \cdot X = \begin{pmatrix} A_1 \\ \vdots \\ A_p \end{pmatrix} \cdot (X) \stackrel{\text{def}}{=} \begin{pmatrix} A_1 \cdot X \\ \vdots \\ A_p \cdot X \end{pmatrix} \in \mathcal{M}(p, 1)$$

Osservazione: Possiamo notare che fare il prodotto fra la matrice e la colonna nel modo illustrato sopra è equivalente a eseguire il prodotto:

$$X \cdot A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot (A^1 \quad \dots \quad A^n) \stackrel{\text{def}}{=} x_1 A^1 + \dots + x_n A^n$$

Esempio: $\begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 0 \cdot 0 + 2 \cdot 1 \\ 1 \cdot 5 - 1 \cdot 0 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$

$$\begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}.$$

Perciò $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 A^1 + \dots + x_n A^n \in \text{Span}(A^1, \dots, A^n)$.

DEFINIZIONE 2.4.10: Definiamo **spazio delle colonne** di una matrice A
 $\mathcal{C}(A) = \text{Span}(A^1, \dots, A^n)$.

PROPOSIZIONE 2.4.6: Sia $A \in \mathcal{M}(p, q, \mathbb{K})$. Allora l'applicazione:

$L_A: \mathbb{K}^q \rightarrow \mathbb{K}^p$
 $L_A: X \rightarrow A \cdot X$ è lineare.

Dimostrazione:

Sfruttando la definizione $A \cdot X = x_1 A^1 + \dots + x_q A^q$, la dimostrazione è immediata.

Notazione: Fissiamo il campo \mathbb{K}^q . Denotiamo:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^q, \dots, e_q = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{K}^q$$

Osservazione: $L_A(e_1) = A \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A^1, \dots, L_A(e_q) = A \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = A^q$, quindi:
 $A = (Ae_1 \quad \dots \quad Ae_q)$

Esempio: Prendiamo $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \mid L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$.

Dunque A rappresenta la riflessione nel piano rispetto alla bisettrice del 1°/3° quadrante.

TEOREMA 2.4.7: Ogni applicazione lineare $\mathbb{K}^q \rightarrow \mathbb{K}^p$ è indotta da una matrice, ossia
 $\forall g: \mathbb{K}^q \rightarrow \mathbb{K}^p$ lineare $\exists! A \in \mathcal{M}(p, q, \mathbb{K})$ tale che $g(X) = A \cdot X \quad \forall X \in \mathbb{K}^q$.

Dimostrazione:

Grazie all'osservazione precedente è chiaro che l'unica matrice di questo tipo può essere solo:

$$A = (g(e_1) \quad \dots \quad g(e_q))$$

poiché per imposizione nel teorema $g(e_1) = A \cdot e_1 = A^1 \dots$

Verifichiamo che con una tale scelta $g(X) = A \cdot X \quad \forall X \in \mathbb{K}^q$:

$$A \cdot X = x_1 g(e_1) + \dots + x_q g(e_q) = (\text{per linearità}) = g(x_1 e_1 + \dots + x_q e_q) = g \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} = g(X)$$

Esempio: L'applicazione $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3 \mid g(x, y) = (y, 2x - y, 5x)$ è indotta dalla matrice:

$$A = (g(e_1) \quad g(e_2)) = \begin{pmatrix} 0 & 1 \\ 2 & -1 \\ 5 & 0 \end{pmatrix}$$

DEFINIZIONE 2.4.11: $f: V \rightarrow W$ lineare si dice **isomorfismo** se è bigettiva.

PROPOSIZIONE 2.4.8: L'applicazione:

$f: \mathcal{M}(p, n, \mathbb{K}) \rightarrow \text{Hom}(\mathbb{K}^n, \mathbb{K}^p)$
 $A \rightarrow L_A$ è un isomorfismo.

Dimostrazione:

a. f è lineare:

$$\forall X \in \mathbb{K}^n \quad L_{A+B} = (A+B) \cdot X = x_1(A+B)^1 + \dots + x_n(A+B)^n = \\ = x_1(A^1 + B^1) + \dots + x_n(A^n + B^n) = A \cdot X + B \cdot X = L_A(X) + L_B(X)$$

Analogamente per il prodotto per scalari.

b. f è surgettiva per il teorema precedente

c. f è iniettiva:

$$\text{Sia } A \in \text{Ker}(f) \Rightarrow L_A(X) = 0 \quad \forall X \Rightarrow L_A(e_i) = A \cdot e_i = A^i = 0 \quad \forall i \Rightarrow A = 0 \Rightarrow \text{Ker}(f) = \{0\}.$$

PROPOSIZIONE 2.4.9: Se $f: V \rightarrow W$ è un isomorfismo, allora $f^{-1}: W \rightarrow V$ è un isomorfismo.

Dimostrazione:

Sappiamo già che l'inversa di una funzione bigettiva è bigettiva, dunque dobbiamo mostrare che f^{-1} è lineare.

Siano $w_1, w_2 \in W$ e sia $v_1 = f^{-1}(w_1), v_2 = f^{-1}(w_2)$. Allora:

$$f^{-1}(w_1 + w_2) = f^{-1}(f(v_1) + f(v_2)) = f^{-1}(f(v_1 + v_2)) = v_1 + v_2 = f^{-1}(w_1) + f^{-1}(w_2)$$

Per il prodotto per scalari il ragionamento è analogo.

PROPOSIZIONE 2.4.10: Dati V, W, Z \mathbb{K} -spazi vettoriali. Siano $f: V \rightarrow W$ e $g: W \rightarrow Z$ lineari.

Allora $g \circ f: V \rightarrow Z$ è lineare.

Dimostrazione:

La verifica è immediata.

COROLLARIO 2.4.11: La composizione di isomorfismi è un isomorfismo.

DEFINIZIONE 2.4.12: Definiamo $GL(V) = \{f: V \rightarrow V \mid f \text{ è isomorfismo}\}$.

COROLLARIO 2.4.12: $(GL(V), \circ)$ è un gruppo, detto **gruppo lineare generale**.

DEFINIZIONE 2.4.13: Siano V, W \mathbb{K} -spazi vettoriali. V e W si dicono **isomorfi** (si scrive $V \cong W$) se $\exists f: V \rightarrow W$ isomorfismo.

Osservazione: $\mathcal{M}(p, q, \mathbb{K}) \cong \text{Hom}(\mathbb{K}^q, \mathbb{K}^p)$

Osservazione: L'“essere isomorfi” è una relazione di equivalenza:

- 1) è riflessiva, poiché sicuramente $V \cong V$ tramite $f = id_V$;
- 2) è simmetrica, poiché se $V \cong W$ tramite f , allora $W \cong V$ tramite f^{-1} , che sappiamo essere un isomorfismo;
- 3) è transitiva, poiché se $V \cong W$ tramite f e $W \cong Z$ tramite g , $V \cong Z$ tramite $g \circ f$, che sappiamo essere un isomorfismo.

DEFINIZIONE 2.4.14: Si definisce **endomorfismo** ogni applicazione $f: V \rightarrow V$ lineare.

DEFINIZIONE 2.4.15: Si definisce spazio degli endomorfismi $\text{End}(V) = \{f: V \rightarrow V \mid f \text{ è lineare}\}$.

Osservazione: $\text{End}(V) = \text{Hom}(V, V)$, dunque $\text{End}(V)$ è sottospazio di $\mathcal{F}(V, V)$.

PROPOSIZIONE 2.4.13: $(\text{End}(V), +, \circ)$ è un anello.

Dimostrazione:
Lasciata al lettore.

DEFINIZIONE 2.4.16: Una quaterna $(S, +, \cdot, \circ)$ si dice **algebra** se $(S, +, \cdot)$ è uno spazio vettoriale, $(S, +, \circ)$ è un anello e (S, \cdot, \circ) ha la seguente proprietà:

$$\forall \alpha \in \mathbb{K}, \forall f, g \in S \quad \alpha(f \circ g) = (\alpha f) \circ g = f \circ (\alpha g)$$

PROPOSIZIONE 2.4.14: $(\text{End}(V), +, \cdot, \circ)$ è un'algebra.

Dimostrazione:

L'ultima verifica è lasciata al lettore.

Osservazione: Date $f: \mathbb{K}^n \rightarrow \mathbb{K}^p$, $g: \mathbb{K}^p \rightarrow \mathbb{K}^q$ lineari, sappiamo che:

$$\exists A \in \mathcal{M}(p, n, \mathbb{K}) \mid f(X) = A \cdot X \quad \forall x \in \mathbb{K}^n;$$

$$\exists B \in \mathcal{M}(q, p, \mathbb{K}) \mid g(X) = B \cdot X \quad \forall x \in \mathbb{K}^p;$$

$g \circ f: \mathbb{K}^n \rightarrow \mathbb{K}^q$ è lineare.

Quindi $\exists C \in \mathcal{M}(q, n, \mathbb{K}) \mid (g \circ f)(X) = C \cdot X \quad \forall X \in \mathbb{K}^n$.

Vediamo che $C^i = (g \circ f)(e_i) = g(f(e_i)) = g(A^i) = B \cdot A^i \quad \forall i$, perciò:

$$C = (BA^1 \quad \dots \quad BA^n)$$

DEFINIZIONE 2.4.17: Si definisce prodotto fra due matrici $B \in \mathcal{M}(q, p)$ e $A \in \mathcal{M}(p, n)$ la matrice $C \in \mathcal{M}(q, n)$ tale che:

$$C = B \cdot A = (BA^1 \quad \dots \quad BA^n)$$

ossia $[C]_{ji} = B_j \cdot A^i$.

Questo prodotto viene chiamato **prodotto righe per colonne**.

PROPOSIZIONE 2.4.15: Valgono le seguenti proprietà $\forall A, B, C$ di formato opportuno:

1) $(AB)C = A(BC)$;

2) $(\lambda A)B = \lambda(AB) = A(\lambda B)$;

3) $(A + B)C = AC + BC$;

4) $A(B + C) = AB + AC$;

5) $IA = AI = A$, dove $I = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$ è la **matrice identica**.

Osservazioni: 1) Non ha senso parlare in generale di commutatività del prodotto fra matrici, poiché se A, B non sono quadrate, se posso eseguire $A \cdot B$ non posso eseguire $B \cdot A$ e viceversa.

2) Anche se in $\mathcal{M}(n)$ ha senso parlare di commutatività del prodotto, in generale $AB \neq BA$:

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 1 & 3 \end{pmatrix}; \quad \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 2 \end{pmatrix}.$$

3) $AB = 0 \not\Rightarrow A = 0 \vee B = 0$:

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

Questo significa che $\mathcal{M}(n)$ non è un campo, quindi vuol dire che esistono matrici che non hanno un'inversa:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ 0 & 0 \end{pmatrix} \neq I \quad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

4) $A^n = 0 \not\Rightarrow A = 0$:

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

DEFINIZIONE 2.4.18: $A \in \mathcal{M}(n)$ si dice **nilpotente** se $\exists s \in \mathbb{N} \mid A^s = 0$.

PROPOSIZIONE 2.4.16: 1) ${}^t(AB) = {}^tB {}^tA \quad \forall A, B \in \mathcal{M}(n)$

2) $\forall A \in \mathcal{M}(n), \forall S \in \mathcal{S}(n), {}^tASA \in \mathcal{S}(n)$

3) $\text{tr}(AB) = \text{tr}(BA) \quad \forall A, B \in \mathcal{M}(n)$.

Dimostrazione:

1) $[{}^t(AB)]_{ij} = [AB]_{ji} = A_j B^i = \sum_{k=1}^n [A]_{jk} [B]_{ki}$;
 $[{}^tB {}^tA]_{ij} = ({}^tB)_i ({}^tA)^j = B^i A_j = \sum_{k=1}^n [B]_{ki} [A]_{jk}$.
 Dunque $[{}^t(AB)]_{ij} = [{}^tB {}^tA]_{ij} \quad \forall i, j \Rightarrow {}^t(AB) = {}^tB {}^tA$.

2) Dimostriamo che ${}^t({}^tASA) = {}^tASA$:

$${}^t({}^tASA) = {}^t(SA) \cdot {}^tA = {}^tA {}^tSA = {}^tASA.$$

3) $\text{tr}(AB) = \sum_{k=1}^n A_k B^k = \sum_{k=1}^n \sum_{s=1}^n [A]_{ks} [B]_{sk}$
 $\text{tr}(BA) = \sum_{k=1}^n B_k A^k = \sum_{k=1}^n \sum_{s=1}^n [B]_{ks} [A]_{sk}$,

che sono uguali perché ogni elemento della prima sta nella seconda con s, k scambiati.

DEFINIZIONE 2.4.19: Definiamo $GL(n, \mathbb{K}) = \{A \in \mathcal{M}(n, \mathbb{K}) \mid A \text{ è un isomorfismo di } \mathbb{K}^n\}$.

Osservazione: So che $\forall A, B \in GL(n, \mathbb{K}), A \cdot B \in GL(n, \mathbb{K})$, poiché ho definito il prodotto fra matrici come composizione di A e B .

So anche che la composizione di isomorfismi è un isomorfismo, perciò \cdot è un'operazione in $GL(n, \mathbb{K})$.

PROPOSIZIONE 2.4.17: $(GL(n, \mathbb{K}), \cdot)$ è un gruppo, detto **gruppo lineare generale in \mathbb{K}** .

Dimostrazione:

- Il prodotto fra matrici è associativo in $\mathcal{M}(n, \mathbb{K}) \supset GL(n, \mathbb{K})$, dunque lo è anche in $GL(n, \mathbb{K})$;
- $I_n \in GL(n, \mathbb{K})$;
- $\forall A \in GL(n, \mathbb{K}), \exists A^{-1} \mid A^{-1} \in GL(n, \mathbb{K})$, poiché ho già dimostrato che l'inversa di un isomorfismo esiste ed è un isomorfismo.

PROPOSIZIONE 2.4.18: 1) $\forall A \in GL(n, \mathbb{K})$, allora ${}^tA \in GL(n, \mathbb{K})$ e $({}^tA)^{-1} = {}^t(A^{-1})$;

2) Se $A, B \in GL(n, \mathbb{K})$, allora $(AB)^{-1} = B^{-1}A^{-1}$;

3) Se $A \in GL(n, \mathbb{K}), B \in \mathcal{M}(n, \mathbb{K})$ e $AB = I \Rightarrow BA = I$.

Dimostrazione:

1) ${}^t(A^{-1}) {}^tA = {}^t(AA^{-1}) = {}^tI = I$;
 ${}^tA {}^t(A^{-1}) = {}^t(A^{-1}A) = {}^tI = I$,

dunque tA ha un'inversa destra che è anche un'inversa sinistra, dunque ${}^tA \in GL(n, \mathbb{K})$.

2) $(B^{-1}A^{-1})AB = B^{-1}IB = B^{-1}B = I$;
 $AB(B^{-1}A^{-1}) = A^{-1}IA = A^{-1}A = I$.

Perciò $(AB)^{-1} = B^{-1}A^{-1}$.

3) A è biettiva e B inversa sinistra di A , quindi B è anche inversa destra, cioè $BA = I$.

Osservazione: Siano $A \in \mathcal{M}(p, n, \mathbb{K})$, $B \in \mathcal{M}(n, q, \mathbb{K})$, e siano $p_1, p_2, n_1, n_2, q_1, q_2$ tali che $p = p_1 + p_2, n = n_1 + n_2, q = q_1 + q_2$.

Allora osserviamo che il prodotto fra matrici può essere fatto **a blocchi**:

$$A = \left(\begin{array}{cc} \underbrace{A_1}_{n_1} & \underbrace{A_2}_{n_2} \\ \underbrace{A_3}_{n_1} & \underbrace{A_4}_{n_2} \end{array} \right) \begin{array}{l} \} p_1 \\ \} p_2 \end{array}, \quad B = \left(\begin{array}{cc} \underbrace{B_1}_{q_1} & \underbrace{B_2}_{q_2} \\ \underbrace{B_3}_{q_1} & \underbrace{B_4}_{q_2} \end{array} \right) \begin{array}{l} \} n_1 \\ \} n_2 \end{array} \Rightarrow A \cdot B = \left(\begin{array}{cc} \underbrace{A_1 B_1 + A_2 B_3}_{q_1 \text{ colonne}} & \underbrace{A_1 B_2 + A_2 B_4}_{q_2 \text{ colonne}} \\ \underbrace{A_3 B_1 + A_4 B_3}_{q_1 \text{ colonne}} & \underbrace{A_3 B_2 + A_4 B_4}_{q_2 \text{ colonne}} \end{array} \right) \begin{array}{l} \} p_1 \text{ righe} \\ \} p_2 \text{ righe} \end{array}.$$

Il lettore può verificare per esercizio che il prodotto così definito coincide con il prodotto definito precedentemente.

2.5 SISTEMI LINEARI

DEFINIZIONE 2.5.1: Definiamo **sistema lineare** di p equazioni in n incognite:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \dots + a_{pn}x_n = b_p \end{cases}.$$

Osservazione: Un sistema lineare si può scrivere nella forma $AX = B$, dove:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \dots & a_{pn} \end{pmatrix} \in \mathcal{M}(p, n, \mathbb{K}),$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n; \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \in \mathbb{K}^p.$$

Quindi $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ è soluzione del sistema $\Leftrightarrow AY = B$.

DEFINIZIONE 2.5.2: Se $B = 0$, il sistema si dice **omogeneo**.

Osservazione: I sistemi omogenei ammettono sempre $0 \in \mathbb{K}^n$ come soluzione.

DEFINIZIONE 2.5.3: Risolvere il sistema $AX = B$ significa trovare tutte le soluzioni del sistema.

Osservazione: $AX = B$ è risolubile $\Leftrightarrow B \in \text{Im}(A)$.

Notazione: Denotiamo l'insieme delle soluzioni del sistema $AX = B$ con:

$$\text{Sol}_B = \{X \in \mathbb{K}^n \mid AX = B\}.$$

Osservazione: $\text{Sol}_0 = \{X \in \mathbb{K}^n \mid AX = 0\} = \text{Ker}(A)$, dunque Sol_0 è un sottospazio vettoriale di \mathbb{K}^n (mentre Sol_B non lo è perché non contiene 0).

DEFINIZIONE 2.5.4: Definiamo **sistema omogeneo associato** al sistema $AX = B$ il sistema $AX = 0$.

PROPOSIZIONE 2.5.1: Sia y_B una qualsiasi soluzione di $AX = B$. Allora:

$$Sol_B = y_B + Sol_0 \stackrel{\text{def}}{=} \{y_B + X \mid X \in Sol_0\}.$$

Dimostrazione:

\supseteq) Sia $X \in Sol_0$. Devo verificare che $y_B + X \in Sol_B$:

$$y_B + X \in Sol_B \Leftrightarrow A(y_B + X) = B \text{ e } A(y_B + X) = Ay_B + AX = B + 0 = B$$

\subseteq) Sia $X \in Sol_B$. Poiché $X = y_B + (X - y_B)$, verifico che $X - y_B \in Sol_0$.

$$\text{Infatti } A(X - y_B) = AX - Ay_B = B - B = 0.$$

DEFINIZIONE 2.5.5: Due sistemi lineari si dicono **equivalenti** se hanno esattamente le stesse soluzioni.

DEFINIZIONE 2.5.6: Definiamo **operazioni elementari sul sistema** le seguenti operazioni:

1° tipo: Scambiare due equazioni;

2° tipo: Moltiplicare un'equazione per uno scalare $\neq 0$;

3° tipo: Sostituire un'equazione con quella ottenuta sommando ad essa un multiplo di un'altra equazione.

Osservazione: In notazione matriciale, ciò corrisponde ad eseguire sulla matrice $A' = (A : B)$, detta **matrice completa del sistema**, una delle seguenti **operazioni elementari per riga**:

1° tipo: Scambiare due righe;

2° tipo: Moltiplicare una riga per uno scalare $\neq 0$;

3° tipo: Aggiungere ad una riga un multiplo di un'altra riga.

Tutte queste operazioni non modificano l'insieme delle soluzioni del sistema.

$$\text{Esempio: } \begin{cases} x_1 - x_2 + 2x_3 - x_4 = 1 \\ 2x_1 - 2x_2 + 5x_3 + x_4 = 3 \end{cases}$$

$$A' = \left(\begin{array}{cccc|c} 1 & -1 & 2 & -1 & 1 \\ 2 & -2 & 5 & 1 & 3 \end{array} \right) \xrightarrow{A_2 \rightarrow A_2 - 2A_1} \left(\begin{array}{cccc|c} 1 & -1 & 2 & -1 & 1 \\ 0 & 0 & 1 & 3 & 1 \end{array} \right)$$

Perciò $x_3 = -3x_4 + 1$.

Sostituendo nell'altra equazione:

$$x_1 - x_2 + 2(-3x_4 + 1) - x_4 = 1 \Rightarrow x_1 - x_2 - 7x_4 = -1 \Rightarrow x_1 = x_2 + 7x_4 - 1.$$

Quindi:

$$Sol_B = \left\{ \left(\begin{array}{c} x_2 + 7x_4 - 1 \\ x_2 \\ -3x_4 + 1 \\ x_4 \end{array} \right) \mid x_2, x_4 \in \mathbb{R} \right\} = \left\{ \left(\begin{array}{c} -1 \\ 0 \\ 1 \\ 0 \end{array} \right) + x_2 \left(\begin{array}{c} 1 \\ 1 \\ 0 \\ 0 \end{array} \right) + x_4 \left(\begin{array}{c} 7 \\ 0 \\ -3 \\ 1 \end{array} \right) \mid x_2, x_4 \in \mathbb{R} \right\}$$

Osservazione: Il termine $\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ non è altro che una soluzione y_B del sistema (nel caso $x_2 =$

$x_4 = 0$), mentre il termine $x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 7 \\ 0 \\ -3 \\ 1 \end{pmatrix}$ è la soluzione generale del sistema omogeneo

associato, perciò:

$$Sol_B = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \underbrace{Span \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 0 \\ -3 \\ 1 \end{pmatrix} \right)}_{=Sol_0}$$

Osservazione: Un sistema del tipo:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \dots + a_{1n}x_n = b_1 \\ a_{2j_2}x_{j_2} + a_{2(j_2+1)}x_{(j_2+1)} + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{pj_p}x_{j_p} + \dots + a_{pn}x_{pn} = b_p \end{cases}$$

con $a_{11} \neq 0, a_{2j_2} \neq 0, \dots, a_{pj_p} \neq 0$, cioè se in una riga sono nulli i coefficienti di x_1, \dots, x_k , nella successiva sono nulli almeno quelli di x_1, \dots, x_k, x_{k+1} , è facilmente risolvibile.

Infatti ricavo x_{j_p} nell'ultima equazione (poiché $a_{pj_p} \neq 0$), poi $x_{j_{p-1}}$ dalla penultima e così via fino a x_1 dalla prima equazione, tutti in funzione dei x_i con $i \neq 1, j_1, \dots, j_p$.

DEFINIZIONE 2.5.7: Una matrice A del tipo:

$$A = \begin{pmatrix} 0 \dots 0 & | & \underline{p_1} \dots \dots \\ 0 \dots 0 & 0 \dots 0 & | & \underline{p_2} \dots \\ & & & | & \vdots \\ 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & | & \underline{p_r} \dots \\ 0 \dots & & & & \dots 0 \end{pmatrix}$$

cioè in cui se nella n -esima riga ci sono k zeri iniziali, nella $(n + 1)$ -esima ce ne sono almeno $k + 1$, viene detta **a scalini**.

Il primo termine $\neq 0$ di ogni riga viene detto **pivot**.

Osservazione: Se $A' = (A : B)$ è a scalini, il sistema $AX = B$ è risolvibile \Leftrightarrow la colonna B non contiene nessun pivot.

In tal caso, se i pivots sono contenuti nelle colonne A^{j_1}, \dots, A^{j_r} , ricavo le incognite x_{j_1}, \dots, x_{j_r} in funzione delle altre.

ALGORITMO DI GAUSS:

Data una $M \in \mathcal{M}(p, q)$, l'algoritmo trasforma M in una matrice a scalini attraverso un numero finito di operazioni elementari per riga.

- Sia M^{j_1} la prima colonna da sinistra non nulla.
- A meno di scambi di riga, posso supporre $[M]_{1,j_1} \neq 0$.
- Per $i = 2, \dots, p$ sostituisco la riga M_i con la riga $M_i - [M]_{i,j_1} \cdot ([M]_{1,j_1})^{-1} \cdot M_1$ (cioè rendo $[M]_{i,j_1} = 0$).
- Ottengo:

$$\tilde{M} = \left(0 \left| \begin{array}{c} [M]_{1,j_1} \\ 0 \\ \vdots \\ 0 \end{array} \right| * \right)$$

- Considero in \tilde{M} la sottomatrice ottenuta eliminando la prima riga e le prime j_1 colonne. Itero il procedimento.

- Termino quando ho trattato tutte le righe o quando restano solo righe nulle.

TEOREMA DI GAUSS: Ogni sistema lineare $AX = B$ è equivalente ad un altro sistema lineare $SX = T$, dove $S' = (S : T)$ è a scalini.

Il sistema $AX = B$ è risolubile \Leftrightarrow le matrici S e S' hanno lo stesso numero di pivots (cioè se T non contiene pivots).

Osservazione: La riduzione a scalini di una matrice non è unica.

DEFINIZIONE 2.5.8: Definiamo **forma parametrica** di un sottospazio di \mathbb{K}^n

$W = Span(w_1, \dots, w_p) = \{X \in \mathbb{K}^n | \exists t_1, \dots, t_p \in \mathbb{K} \text{ t. c. } x = t_1 w_1 + \dots + t_p w_p\}$ la scrittura:
 $W = Im(A)$

dove $A: \mathbb{K}^p \rightarrow \mathbb{K}^n$ è la matrice:

$$A = (w_1 | \dots | w_p)$$

con i vettori w_1, \dots, w_p per colonne.

DEFINIZIONE 2.5.9: Definiamo **forma cartesiana** di un sottospazio W di \mathbb{K}^n la scrittura:

$$W = \{X | BX = 0\} = Ker(B)$$

con $\mathcal{M}(q, n) \ni B: \mathbb{K}^n \rightarrow \mathbb{K}^q$.

Le equazioni del sistema $BX = 0$ si dicono **equazioni cartesiane** di W .

Osservazione: • Si passa dalle equazioni cartesiane $BX = 0$ alla forma parametrica risolvendo il sistema $BX = 0$.

- Per passare dalla forma parametrica alla forma cartesiana, si costruisce il sistema $(A : X)$, dove A è tale che $W = Im(A)$ e $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Si porta il sistema $(A : X)$ nella forma a scalini $(S : X')$, dove $X' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$, e, dette S_{k_1}, \dots, S_{k_r}

le righe con pivot di S , poniamo $x_i' = 0 \forall i \neq k_1, \dots, k_r$ (ottenendo quindi un sistema con $n - r$ equazioni).

Esempio: Sia $W \subset \mathbb{R}^3$ il sottospazio vettoriale $\begin{cases} x = t \\ y = t + s \\ z = 3t + 2s \end{cases}$, ossia $W = Span\left(\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right)$.

Perciò $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in W \Leftrightarrow \exists t, s \in \mathbb{R} \begin{cases} t = x \\ t + s = y \\ 3t + 2s = z \end{cases} \Leftrightarrow$ il sistema $\begin{pmatrix} 1 & 0 & : & x \\ 1 & 1 & : & y \\ 3 & 2 & : & z \end{pmatrix}$ ha soluzione \Leftrightarrow
 $\begin{pmatrix} 1 & 0 & : & x \\ 0 & 1 & : & y - x \\ 0 & 2 & : & z - 3x \end{pmatrix}$ ha soluzione $\Leftrightarrow \begin{pmatrix} 1 & 0 & : & x \\ 0 & 1 & : & y - x \\ 0 & 0 & : & z - x - 2y \end{pmatrix}$ ha soluzione $\Leftrightarrow z - x - 2y = 0$.

Quindi la forma cartesiana per W è:

$$W = \{x + 2y - z = 0\} = Ker(B)$$

dove $B = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$.

DEFINIZIONE 2.5.10: Sia V uno spazio vettoriale e $v \in V$. Si definisce **traslazione** di v l'applicazione:

$$\tau_v: V \rightarrow V \mid \tau_v(x) = x + v.$$

PROPOSIZIONE 2.5.2: 1) $\forall v \in V$, τ_v è bigettiva;

$$2) \forall v, w \in V, \tau_{v+w} = \tau_v \circ \tau_w = \tau_w \circ \tau_v;$$

$$3) \forall v \in V, (\tau_v)^{-1} = \tau_{-v}.$$

Dimostrazione:

$$1) \tau_v \text{ è iniettiva, poiché se } x \neq y \in V, \tau_v(x) = x + v \neq y + v = \tau_v(y);$$

$$\tau_v \text{ è surgettiva, poiché } \forall x \in V, \exists f^{-1}(x) \in V \mid f^{-1}(x) + v = x.$$

$$2) \forall x \in V:$$

$$\tau_{v+w}(x) = x + v + w;$$

$$(\tau_v \circ \tau_w)(x) = \tau_v(\tau_w(x)) = \tau_v(x + w) = x + v + w;$$

$$(\tau_w \circ \tau_v)(x) = \tau_w(\tau_v(x)) = \tau_w(x + v) = x + v + w.$$

$$3) (\tau_v \circ \tau_{-v})(x) = \tau_v(\tau_{-v}(x)) = \tau_v(x - v) = x \quad \forall x \in V.$$

Quindi le traslazioni di V formano un gruppo abeliano.

DEFINIZIONE 2.5.11: Sia W un sottospazio vettoriale di V e $v \in V$. Si definisce **sottospazio affine** di V con **giacitura** W l'immagine di τ_v , cioè:

$$H = \tau_v(W) = \{v + w \mid w \in W\}$$

Esempi: 1) Sia dato il sistema $AX = B$, con $A \in \mathcal{M}(p, n, \mathbb{K})$.

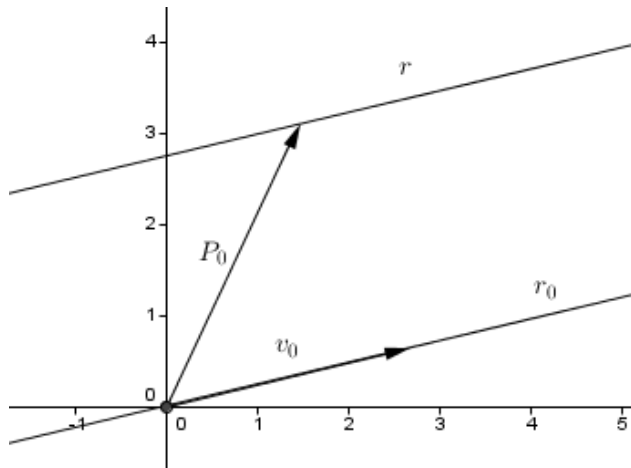
$Sol_B = y_0 + Sol_0$, quindi Sol_B è sottospazio affine di \mathbb{K}^n ;

2) Ogni retta r di \mathbb{R}^3 è un sottospazio affine con giacitura la retta $r_0 // r$ e passante per l'origine.

$$\text{Se } P_0 \in r, r = \tau_{P_0}(r_0);$$

$$\text{se } r_0 = \text{Span}(v_0), r = \{X \in \mathbb{R}^3 \mid X = P_0 + tv_0, t \in \mathbb{R}\}.$$

Graficamente:



Possiamo rappresentare un sottospazio affine di \mathbb{K}^n in forma parametrica e cartesiana:

Sia W sottospazio vettoriale di \mathbb{K}^n e sia $H = P_0 + W$.

Per la forma parametrica, scrivo $W = \text{Im}(A) = \{AY \mid Y \in \mathbb{K}^p\}$.

Allora $H = \{AY + P_0 \mid Y \in \mathbb{K}^p\}$.

Per la forma cartesiana, scrivo $W = \text{Ker}(B) = \{X \in \mathbb{K}^n \mid BX = 0\}$.

Allora:

$$H = \{X \in \mathbb{K}^n \mid X = Y + P_0, Y \in W\} = \{X \in \mathbb{K}^n \mid X - P_0 \in W\} = \{X \in \mathbb{K}^n \mid B(X - P_0) = 0\} = \\ = \{X \in \mathbb{K}^n \mid BX = BP_0\}$$

Si passa da una rappresentazione all'altra in modo analogo al caso vettoriale.

2.6 BASI E DIMENSIONE

DEFINIZIONE 2.6.1: Uno spazio vettoriale V si dice **finitamente generato** se $\exists v_1, \dots, v_n \in V \mid \forall v \in V \exists \alpha_1, \dots, \alpha_n \in \mathbb{K} \mid v = \alpha_1 v_1 + \dots + \alpha_n v_n$, ossia $V = \text{Span}(v_1, \dots, v_n)$.

In tal caso, v_1, \dots, v_n sono detti **generatori** di V .

Esempio: $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ generano \mathbb{K}^n .

Osservazione: $\mathbb{K}[x]$ non è finitamente generato, poiché se per assurdo $\mathbb{K}[x] = \text{Span}(1, x, \dots, x^a)$, con $a \in \mathbb{N}$, non si potrebbero rappresentare i polinomi di grado $> a$.

DEFINIZIONE 2.6.2: $v_1, \dots, v_n \in V$ sono detti **linearmente indipendenti** se $a_1 v_1 + \dots + a_n v_n = 0 \Rightarrow a_1 = \dots = a_n = 0$.

Altrimenti sono detti **linearmente dipendenti**.

Esempio: $e_1, \dots, e_n \in \mathbb{K}^n$ sono linearmente indipendenti, infatti:

$$a_1 e_1 + \dots + a_n e_n = 0 \Rightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0 \Rightarrow a_1 = \dots = a_n = 0.$$

Osservazione: $v \in V$ è linearmente indipendente $\Leftrightarrow v \neq 0$.

Osservazione: Se uno fra i vettori v_1, \dots, v_n è nullo, allora v_1, \dots, v_n sono linearmente dipendenti.

Infatti, se ad esempio $v_1 = 0$:

$$a v_1 + 0 v_2 + \dots + 0 v_n = 0 \nRightarrow a = 0.$$

PROPOSIZIONE 2.6.1: Sia $n \geq 2$. I vettori v_1, \dots, v_n sono linearmente dipendenti \Leftrightarrow almeno uno di essi si può esprimere come combinazione lineare degli altri.

Dimostrazione:

\Rightarrow): Per ipotesi $\exists a_1, \dots, a_n$ non tutti nulli $\mid a_1 v_1 + \dots + a_n v_n = 0$.

Se $a_1 \neq 0$, allora $v_1 = -a_1^{-1}(a_2 v_2 + \dots + a_n v_n)$, tesi.

\Leftarrow): Se $v_1 = a_2 v_2 + \dots + a_n v_n$, allora $v_1 - a_2 v_2 - \dots - a_n v_n = 0$, da cui la tesi.

Osservazione: Se v_1, \dots, v_n sono linearmente indipendenti e $k \leq n$, allora v_1, \dots, v_k sono linearmente indipendenti.

PROPOSIZIONE 2.6.2: Se $v_m \in \text{Span}(v_1, \dots, v_{m-1})$, allora $\text{Span}(v_1, \dots, v_m) = \text{Span}(v_1, \dots, v_{m-1})$.

Dimostrazione:

\subseteq) Se $v \in \text{Span}(v_1, \dots, v_m) \Rightarrow v = a_1 v_1 + \dots + a_m v_m = a_1 v_1 + \dots + a_{m-1} v_{m-1} + a_m (b_1 v_1 + \dots + b_{m-1} v_{m-1}) = (a_1 + a_m b_1) v_1 + \dots + (a_{m-1} + a_m b_{m-1}) v_{m-1} \Rightarrow v \in \text{Span}(v_1, \dots, v_{m-1})$.

\supseteq) Se $v \in \text{Span}(v_1, \dots, v_{m-1}) \Rightarrow v = a_1 v_1 + \dots + a_{m-1} v_{m-1} = a_1 v_1 + \dots + a_{m-1} v_{m-1} + 0 v_m$, quindi $v \in \text{Span}(v_1, \dots, v_m)$.

DEFINIZIONE 2.6.3: Un insieme ordinato $\{v_1, \dots, v_n\}$ di vettori di V è detto **base** di V se v_1, \dots, v_n sono linearmente indipendenti e generano V .

Esempio: $\{e_1, \dots, e_n\}$ è una base di \mathbb{K}^n , detta **base canonica**.

PROPOSIZIONE 2.6.3: Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V , allora ogni $v \in V$ può essere scritto in modo unico come combinazione lineare dei v_1, \dots, v_n .

Dimostrazione:

Poiché i v_1, \dots, v_n sono generatori, allora $v \in V = \text{Span}(v_1, \dots, v_n)$, dunque supponiamo:

$$v = a_1 v_1 + \dots + a_n v_n;$$

$$v = b_1 v_1 + \dots + b_n v_n.$$

Allora $(a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n = 0$, ma i v_i sono linearmente indipendenti, dunque $a_i = b_i \forall i$, da cui segue la tesi.

DEFINIZIONE 2.6.4: I coefficienti dell'unica combinazione lineare dei v_1, \dots, v_n che dà v si chiamano **coordinate** di v rispetto alla base \mathcal{B} e denotati con $[v]_{\mathcal{B}}$.

DEFINIZIONE 2.6.5: Fissando una base \mathcal{B} si determina quindi una corrispondenza biunivoca:

$$[\]_{\mathcal{B}}: V \rightarrow \mathbb{K}^n \mid v \rightarrow [v]_{\mathcal{B}}$$

chiamata "**coordinate rispetto a \mathcal{B}** ".

PROPOSIZIONE 2.6.4: $\forall \mathcal{B}$ base, $[\]_{\mathcal{B}}$ è un isomorfismo.

Dimostrazione:

$[\]_{\mathcal{B}}$ è evidentemente lineare; inoltre è iniettiva, poiché

$\text{Ker}([\]_{\mathcal{B}}) = \{v \in V \mid [v]_{\mathcal{B}} = 0\} = \{v \in V \mid v = 0v_1 + \dots + 0v_n\} = \{0\}$, mentre è surgettiva in quanto ogni $(a_1 \dots a_n) \in \mathbb{K}^n$ è immagine di $v = a_1 v_1 + \dots + a_n v_n \in V$.

COROLLARIO 2.6.5: Se V è un \mathbb{K} -spazio vettoriale che ammette una base formata da n vettori, allora $V \cong \mathbb{K}^n$.

PROPOSIZIONE 2.6.6: Sia $\{v_1, \dots, v_n\}$ base di V e w_1, \dots, w_k dei vettori di V .

Se $k > n$, allora w_1, \dots, w_k sono linearmente dipendenti.

Dimostrazione:

$$\text{Si ha: } w_1 = a_{11} v_1 + \dots + a_{1n} v_n$$

$$w_2 = a_{21} v_1 + \dots + a_{2n} v_n$$

\vdots

$$w_k = a_{k1} v_1 + \dots + a_{kn} v_n.$$

Devo trovare degli α_i non tutti nulli tali che:

$$\alpha_1(a_{11}v_1 + \dots + a_{1n}v_n) + \dots + \alpha_k(a_{k1}v_1 + \dots + a_{kn}v_n) = 0, \text{ ossia:}$$

$$(a_{11}\alpha_1 + \dots + a_{k1}\alpha_k)v_1 + \dots + (a_{1n}\alpha_1 + \dots + a_{kn}\alpha_k)v_n = 0.$$

Ma v_1, \dots, v_n sono linearmente indipendenti, perciò:

$$\begin{cases} a_{11}\alpha_1 + \dots + a_{k1}\alpha_k = 0 \\ \vdots \\ a_{1n}\alpha_1 + \dots + a_{kn}\alpha_k = 0 \end{cases}$$

Poiché è un sistema omogeneo di n equazioni in $k > n$ incognite, ha infinite soluzioni, dunque in particolare ne ha una non nulla, per cui i w_i sono linearmente dipendenti.

COROLLARIO 2.6.7: Se $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_k\}$ sono basi di V , allora $n = k$.

Dimostrazione:

Se $k > n$, i w_i sono linearmente dipendenti per la proposizione precedente;

se $n < k$, i v_i sono linearmente dipendenti,

perciò $k = n$.

DEFINIZIONE 2.6.6: Se V possiede una base finita $\{v_1, \dots, v_n\}$, diciamo che V ha **dimensione** n ($\dim V = n$).

Se $V = \{0\}$, poniamo $\dim V = 0$.

Esempi: 1) $\dim \mathbb{K}^n = n$;

2) $\dim_{\mathbb{C}} \mathbb{C} = 1$, in quanto $\{1\}$ è una base di \mathbb{C} come \mathbb{C} -spazio vettoriale;

3) $\dim_{\mathbb{R}} \mathbb{C} = 2$, in quanto $\{1, i\}$ è una base di \mathbb{C} come \mathbb{R} -spazio vettoriale;

4) $\dim \mathcal{M}(p, n) = p \cdot n$, in quanto $\{E_{ij}\}_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ è una base, dove:

$$[E_{ij}]_{hk} = \delta_{ih} \cdot \delta_{jk} = \begin{cases} 1 & \text{se } (i, j) = (h, k) \\ 0 & \text{se } (i, j) \neq (h, k) \end{cases} \quad (\delta_{ij} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \text{ è detto } \mathbf{\delta} \text{ di Kronecker}).$$

5) $\dim \mathbb{K}_n[x] = n + 1$, in quanto $\{1, x, \dots, x^n\}$ è una base.

DEFINIZIONE 2.6.7: Siano V, W \mathbb{K} -spazi vettoriali. Definiamo una somma e un prodotto per scalari in $V \times W$:

$$(v_1, w_1) + (v_2, w_2) \stackrel{\text{def}}{=} (v_1 + v_2, w_1 + w_2);$$

$$\alpha(v, w) \stackrel{\text{def}}{=} (\alpha v, \alpha w).$$

PROPOSIZIONE 2.6.8: $V \times W$ è spazio vettoriale e $\dim(V \times W) = \dim V + \dim W$.

Dimostrazione:

Lasciamo la verifica che $V \times W$ è spazio vettoriale.

Sia $\{v_1, \dots, v_n\}$ base di V e $\{w_1, \dots, w_k\}$ base di W .

È immediato mostrare che $\{(v_1, 0), \dots, (v_n, 0), (0, w_1), \dots, (0, w_k)\}$ è base di $V \times W$, dunque segue la tesi.

ALGORITMO PER L'ESTRAZIONE DI UNA BASE: Sia $V \neq \{0\}$ uno spazio vettoriale.

Da ogni insieme finito di generatori di V si può estrarre una base.

Dimostrazione:

Siano v_1, \dots, v_k generatori di V . Posso supporre $v_i \neq 0 \ \forall i$, poiché, se ce ne fossero, li potrei togliere e non altererei lo spazio generato.

Allora v_1 è linearmente indipendente.

Guardo $\{v_1, v_2\}$:

- se v_1, v_2 sono linearmente indipendenti, li tengo;
- altrimenti $v_2 \in \text{Span}(v_1)$ e quindi $\text{Span}(v_1, \dots, v_k) = \text{Span}(v_1, v_3, \dots, v_k)$.

Allora elimino v_2 .

Continuo così fino a quando ho considerato tutti i vettori.

COROLLARIO 2.6.9: Sia $\dim V = n$. Se v_1, \dots, v_k sono generatori di V , allora $k \geq n$.

PROPOSIZIONE 2.6.10: Se v_1, \dots, v_k sono linearmente indipendenti e $v \notin \text{Span}(v_1, \dots, v_k)$, allora v, v_1, \dots, v_k sono linearmente indipendenti.

Dimostrazione:

Sia $a_1 v_1 + \dots + a_k v_k + av = 0$.

Deve essere $a = 0$, poiché altrimenti $v = -a^{-1}(a_1 v_1 + \dots + a_k v_k) \Rightarrow v \in \text{Span}(v_1, \dots, v_k)$.

Allora $a_1 v_1 + \dots + a_k v_k = 0$. Poiché i v_1, \dots, v_k sono linearmente indipendenti, segue $a_1 = \dots = a_k = 0$ e quindi la tesi.

TEOREMA DI COMPLETAMENTO A BASE: Sia V uno spazio finitamente generato.

Se $v_1, \dots, v_k \in V$ sono linearmente indipendenti, esistono $v_{k+1}, \dots, v_n \in V$

$\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ è una base di V .

Dimostrazione:

Se $\{v_1, \dots, v_k\}$ generano V , allora $\{v_1, \dots, v_k\}$ è una base di V .

Se non lo generano, allora $\exists v_{k+1} \notin \text{Span}(v_1, \dots, v_k)$.

Per la proposizione precedente, i v_1, \dots, v_k, v_{k+1} sono linearmente indipendenti.

Se generano V , ho trovato una base.

Altrimenti itero il procedimento.

V è finitamente generato, perciò dopo un numero finito di passi il procedimento deve finire.

Osservazione: È un procedimento non algoritmico, poiché non c'è un metodo semplice e diretto per trovare i v_{k+h} , con $h > 0$.

ALGORITMO DI COMPLETAMENTO A BASE: Se v_1, \dots, v_k sono linearmente indipendenti e se conosco una base $\{z_1, \dots, z_n\}$ di V , posso completare $\{v_1, \dots, v_k\}$ a base applicando l'algoritmo di estrazione di una base all'insieme di generatori $\{v_1, \dots, v_k, z_1, \dots, z_n\}$.

PROPOSIZIONE 2.6.11: Ogni sottospazio vettoriale W di uno spazio vettoriale V finitamente generato ha un supplementare.

Dimostrazione:

Sia $\dim V = n$. Allora $\dim W = k \leq n$.

Sia $\{w_1, \dots, w_k\}$ una base di W .

Posso completarla a una base $\{w_1, \dots, w_k, v_{k+1}, \dots, v_n\}$ di V .

Perciò $V = W \oplus \text{Span}(v_{k+1}, \dots, v_n)$ e dunque $\text{Span}(v_{k+1}, \dots, v_n)$ è un supplementare.

PROPOSIZIONE 2.6.12: Sia $V = U \oplus W$, $\{u_1, \dots, u_k\}$ base di U , $\{w_1, \dots, w_m\}$ base di W .

Allora $\{u_1, \dots, u_k, w_1, \dots, w_m\}$ è base di V .

Dimostrazione:

Ogni $v \in V$ si può scrivere come $v = u + w$, con $u \in U$ e $w \in W$, ma

$u = a_1 u_1 + \dots + a_k u_k$ e $w = b_1 w_1 + \dots + b_m w_m$, dunque i $u_1, \dots, u_k, w_1, \dots, w_m$ generano V .

Inoltre sono linearmente indipendenti, poiché:

$$\underbrace{a_1 u_1 + \dots + a_k u_k}_{=u \in U} + \underbrace{b_1 w_1 + \dots + b_m w_m}_{=w \in W} = 0 \Rightarrow u + w = 0 \Rightarrow u = -w, \text{ dunque } W \ni w = -u \in U,$$

perciò $u, w \in U \cap W = \{0\}$, cioè $u = 0$ e $w = 0$.

Poiché $\{u_1, \dots, u_k\}$ e $\{w_1, \dots, w_m\}$ sono linearmente indipendenti, allora

$a_1 = \dots = a_k = b_1 = \dots = b_m = 0$, da cui la tesi.

PROPOSIZIONE 2.6.13: Se $V \neq \{0\}$ non è finitamente generato, allora $\forall n \geq 1$ esistono $v_1, \dots, v_n \in V$ linearmente indipendenti.

Dimostrazione:

Per induzione su n :

Passo base): $n = 1$, basta scegliere $v_1 \neq 0$;

Passo induttivo): Per ipotesi induttiva $\exists v_1, \dots, v_{n-1}$ linearmente indipendenti.

Osservo che $\text{Span}(v_1, \dots, v_{n-1}) \neq V$, poiché V non è finitamente generato.

Dunque $\exists v_n \notin \text{Span}(v_1, \dots, v_{n-1}) \mid v_1, \dots, v_n$ sono linearmente indipendenti.

PROPOSIZIONE 2.6.14: Se V è finitamente generato e W è un sottospazio vettoriale di V , allora:

- 1) W è finitamente generato;
- 2) $\dim W \leq \dim V$;
- 3) se $\dim W = \dim V \Rightarrow W = V$.

Dimostrazione:

- 1) Sia $n = \dim V$. Se W non fosse finitamente generato, per la proposizione precedente esisterebbero $w_1, \dots, w_{n+1} \in W \subset V$ linearmente indipendenti, assurdo.
- 2) Sia $n = \dim V$. Se $\dim W > n$, allora esisterebbero $w_1, \dots, w_{n+1} \in W \subset V$ linearmente indipendenti, assurdo.
- 3) Se $\dim W = n$ e $\{w_1, \dots, w_n\}$ è base di W , allora w_1, \dots, w_n sono linearmente indipendenti anche in V e, poiché $\dim V = n$, devono essere una base di V . Dunque $W = V$.

FORMULA DI GRASSMANN: Siano U, W sottospazi vettoriali di dimensione finita di V . Allora: $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

Dimostrazione:

Sia $\dim U = h$, $\dim W = k$, $\dim(U \cap W) = s$.

Sia $\{z_1, \dots, z_s\}$ una base di $U \cap W$. Allora z_1, \dots, z_s sono linearmente indipendenti in U e W .

Per il teorema di completamento a base $\exists u_1, \dots, u_{h-s} \in U$, $\exists w_1, \dots, w_{k-s} \in W$ tali che:

$\{z_1, \dots, z_s, u_1, \dots, u_{h-s}\}$ è base di U ;

$\{z_1, \dots, z_s, w_1, \dots, w_{k-s}\}$ è base di W .

Se mostro che $\{z_1, \dots, z_s, u_1, \dots, u_{h-s}, w_1, \dots, w_{k-s}\}$ è base di $U + W$ ho la tesi, poiché dimostro che $\dim(U + W) = h + k - s$.

Quei vettori generano, in quanto, preso $v \in U + W$, $\exists u \in U, w \in W \mid v = u + w$.

Inoltre $u = a_1z_1 + \dots + a_s z_s + b_1u_1 + \dots + b_{h-s}u_{h-s}$ e $w = \alpha_1z_1 + \dots + \alpha_s z_s + \beta_1w_1 + \dots + \beta_{k-s}w_{k-s}$, dunque $v = (a_1 + \alpha_1)z_1 + \dots + (a_s + \alpha_s)z_s + b_1u_1 + \dots + b_{h-s}u_{h-s} + \beta_1w_1 + \dots + \beta_{k-s}w_{k-s}$. Mostriamo quindi che sono linearmente indipendenti: sia

$$\underbrace{a_1z_1 + \dots + a_s z_s}_{=z} + \underbrace{b_1u_1 + \dots + b_{h-s}u_{h-s}}_{=u} + \underbrace{c_1w_1 + \dots + c_{k-s}w_{k-s}}_{=w} = 0.$$

Allora $z + u = -w$. Ma $z + u \in U, w \in W$, quindi $z + u = -w \in U \cap W$.

Posso dunque scrivere $w = \alpha_1z_1 + \dots + \alpha_s z_s$, per cui ho:

$$a_1z_1 + \dots + a_s z_s + b_1u_1 + \dots + b_{h-s}u_{h-s} + \alpha_1z_1 + \dots + \alpha_s z_s = 0 \Rightarrow$$

$$(a_1 + \alpha_1)z_1 + \dots + (a_s + \alpha_s)z_s + b_1u_1 + \dots + b_{h-s}u_{h-s} = 0.$$

Questi vettori sono una base di U , quindi $b_1 = \dots = b_{h-s} = 0$.

Allora:

$a_1z_1 + \dots + a_s z_s + c_1w_1 + \dots + c_{k-s}w_{k-s} = 0$, ma questi vettori sono una base di W , quindi sono linearmente indipendenti, per cui $a_1 = \dots = a_s = c_1 = \dots = c_{k-s} = 0$, da cui segue la tesi.

Osservazione: $\dim \mathcal{S}(n) = \frac{n(n+1)}{2}$, $\dim \mathcal{A}(n) = \frac{n(n-1)}{2}$.

Infatti, detta $\{E_{ij}\}_{1 \leq i, j \leq n}$ la base canonica di $\mathcal{M}(n)$, $\{E_{ij} + E_{ji}\}_{1 \leq i < j \leq n} \cup \{E_{ii}\}_{1 \leq i \leq n}$ è una base di $\mathcal{S}(n)$. Lasciamo questa verifica per esercizio. Il numero dei vettori di base è $\frac{n(n-1)}{2} + n$.

Inoltre per Grassmann, $\dim \mathcal{A}(n) = \dim \mathcal{M}(n) - \dim \mathcal{S}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

TEOREMA 2.6.15: Siano V, W \mathbb{K} -spazi vettoriali, con V finitamente generato.

Sia $\{v_1, \dots, v_n\}$ una base di V ; siano w_1, \dots, w_n vettori di W .

Allora $\exists!$ $f: V \rightarrow W$ lineare tale che $f(v_i) = w_i \forall i$.

Dimostrazione:

Esistenza: Sia $v \in V \Rightarrow \exists$ unici $a_1, \dots, a_n \in \mathbb{K} \mid v = a_1v_1 + \dots + a_nv_n$.

Poniamo $f(v) = a_1w_1 + \dots + a_nw_n$.

Si ha evidentemente che $f(v_i) = w_i \forall i$.

Inoltre f è lineare, infatti, se $v = a_1v_1 + \dots + a_nv_n$ e $z = b_1v_1 + \dots + b_nv_n$, allora

$$v + z = (a_1 + b_1)v_1 + \dots + (a_n + b_n)v_n;$$

$$\text{quindi } f(v + z) = (a_1 + b_1)w_1 + \dots + (a_n + b_n)w_n = a_1w_1 + \dots + a_nw_n + b_1w_1 + \dots + b_nw_n = f(v) + f(z).$$

Analogamente per il multiplo.

Unicità: Prendiamo una qualsiasi g lineare tale che $g(v_i) = w_i \forall i$.

Per linearità:

$$g(v) = g\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i g(v_i) = \sum_{i=1}^n a_i w_i = f(v)$$

dunque f è unica.

PROPOSIZIONE 2.6.16: Sia $f: V \rightarrow W$ lineare. Allora:

- 1) Se v_1, \dots, v_n sono linearmente indipendenti e f è iniettiva, allora $f(v_1), \dots, f(v_n)$ sono linearmente indipendenti;
- 2) Se v_1, \dots, v_n generano V , allora $f(v_1), \dots, f(v_n)$ generano $Im(f)$.

Dimostrazione:

- 1) Sia $a_1 f(v_1) + \dots + a_n f(v_n) = 0$. Allora per linearità $f(a_1 v_1 + \dots + a_n v_n) = 0$, dunque $a_1 v_1 + \dots + a_n v_n \in \text{Ker}(f) = \{0\}$, ma v_1, \dots, v_n sono linearmente indipendenti, per cui $a_1 = \dots = a_n = 0$, tesi.
- 2) Sia $y = f(x) \in \text{Im}(f)$. Mostriamo che può essere scritto come combinazione lineare dei $f(v_i)$.

So che $\exists a_1, \dots, a_n \in \mathbb{K} \mid x = a_1 v_1 + \dots + a_n v_n$. Allora:

$$y = f(x) = f\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i f(v_i)$$

da cui la tesi.

Osservazione: Se $A \in \mathcal{M}(p, n, \mathbb{K})$, allora sappiamo che A è una applicazione lineare e la sua immagine è lo spazio generato dalle colonne.

Quindi $\text{Im}(A) = \text{Span}(A^1, \dots, A^n) = \mathcal{C}(A)$.

COROLLARIO 2.6.17: $f: V \rightarrow W$ è un isomorfismo $\Rightarrow f$ trasforma ogni base di V in una base di W .

Dimostrazione:

Prendo una base di V . Poiché f è iniettiva, allora le immagini dei vettori della base sono linearmente indipendenti. Inoltre quegli stessi vettori generano V , quindi le loro immagini generano $\text{Im}(f)$, ma f è surgettiva, quindi $\text{Im}(f) = W$.

FORMULA DELLE DIMENSIONI: Sia V uno spazio vettoriale finitamente generato, e sia $f: V \rightarrow W$ lineare. Allora:

$$\dim V = \dim \text{Ker}(f) + \dim \text{Im}(f)$$

Dimostrazione:

Sia $n = \dim V$ e $k = \dim \text{Ker}(f)$.

Sia $\{v_1, \dots, v_k\}$ una base di $\text{Ker}(f)$; la completo a $\{v_1, \dots, v_k, \dots, v_n\}$ base di V .

Allora $f(v_1), \dots, f(v_n)$ generano $\text{Im}(f)$, ma $f(v_1) = \dots = f(v_k) = 0$, poiché appartengono a $\text{Ker}(f)$.

Perciò posso toglierli e i rimanenti $f(v_{k+1}), \dots, f(v_n)$ generano comunque $\text{Im}(f)$.

Dimostriamo che $f(v_{k+1}), \dots, f(v_n)$ sono linearmente indipendenti:

$$\begin{aligned} a_{k+1} f(v_{k+1}) + \dots + a_n f(v_n) = 0 &\Rightarrow f(a_{k+1} v_{k+1} + \dots + a_n v_n) = 0 \Rightarrow \sum_{i=k+1}^n a_i v_i \in \text{Ker}(f) \\ &\Rightarrow \exists a_1, \dots, a_k \in \mathbb{K} \mid \sum_{i=k+1}^n a_i v_i = a_1 v_1 + \dots + a_k v_k \\ &\Rightarrow a_1 v_1 + \dots + a_k v_k - a_{k+1} v_{k+1} - \dots - a_n v_n = 0 \end{aligned}$$

Ma v_1, \dots, v_n sono linearmente indipendenti, quindi $a_i = 0 \forall i$.

Per cui $\{f(v_{k+1}), \dots, f(v_n)\}$ è una base di $\text{Im}(f)$, cioè $\dim \text{Im}(f) = n - k$, tesi.

COROLLARIO 2.6.18: Sia $f: V \rightarrow W$ lineare.

Se $\dim V = \dim W$, allora f è iniettiva $\Leftrightarrow f$ è surgettiva.

Dimostrazione:

f è iniettiva $\Leftrightarrow \text{Ker}(f) = \{0\} \Leftrightarrow \dim V = \dim \text{Im}(f) \Leftrightarrow \dim W = \dim \text{Im}(f) \Leftrightarrow \text{Im}(f) = W$
 (poiché $\text{Im}(f) \subseteq W$) $\Leftrightarrow f$ è surgettiva.

COROLLARIO 2.6.18: $V \cong W \Leftrightarrow \dim V = \dim W$.

Dimostrazione:

\Leftarrow) Se $\dim V = \dim W$, allora $V \cong \mathbb{K}^n \cong W$.

\Rightarrow) Se $\exists f: V \rightarrow W$ isomorfismo, per la formula delle dimensioni $\dim V = 0 + \dim W$.

Osservazione: Siano V_1, V_2 sottospazi vettoriali di V e sia $f: V_1 \times V_2 \rightarrow V$ data da $f(v_1, v_2) = v_1 + v_2$. f è evidentemente lineare e $\text{Im}(f) = V_1 + V_2$.

Inoltre $\text{Ker}(f)$ è canonicamente isomorfo a $V_1 \cap V_2$, infatti:

$\text{Ker}(f) = \{(v_1, v_2) \in V_1 \times V_2 \mid v_1 + v_2 = 0\} = \{(v_1, v_2) \in V_1 \times V_2 \mid v_1 = -v_2\} =$
 $= \{(v, -v) \mid v \in V_1 \cap V_2\}$, poiché $V_1 \ni v_1 = -v_2 \in V_2$, dunque:

$L: V_1 \cap V_2 \rightarrow V_1 \times V_2 \mid L(v) = (v, -v)$ induce l'isomorfismo cercato (le proprietà sono immediatamente verificabili).

Per cui, per la formula delle dimensioni, $\dim(V_1 + V_2) = \dim \text{Im}(f) = \dim(V_1 \times V_2) - \dim \text{Ker}(f) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$,

che conclude quindi la dimostrazione alternativa della formula di Grassmann.

Osservazione: Nell'insieme quoziente $\{\mathbb{K}\text{-spazi vettoriali finitamente generati}\}/\cong$ esistono tante classi di equivalenza quante \mathbb{N} e in ognuna un rappresentante è \mathbb{K}^n .

La dimensione è un sistema completo di invarianza per la relazione di equivalenza \cong , perciò se $\dim V \neq \dim W$, allora sicuramente V e W non sono isomorfi.

DEFINIZIONE 2.6.8: Due \mathbb{K} -spazi vettoriali V, W si dicono **canonicamente isomorfi** se $\exists f: V \rightarrow W$ isomorfismo che non dipende dalla scelta di una base.

PROPOSIZIONE 2.6.19: Sia $V = U \oplus W$ e $V = U \oplus W'$. $\forall w \in W, \exists! u \in U, w' \in W' \mid w = u + w'$. Allora è ben definita l'applicazione $\varphi: W \rightarrow W' \mid \varphi(w) = w'$.

φ è un isomorfismo canonico.

Dimostrazione:

- φ è lineare, poiché dati $w_1 = u_1 + w_1'$ e $w_2 = u_2 + w_2'$, allora:
 $\varphi(w_1 + w_2) = \varphi((u_1 + u_2) + (w_1' + w_2')) = w_1' + w_2' = \varphi(w_1) + \varphi(w_2);$
 $\varphi(\lambda w) = \varphi(\lambda u + \lambda w') = \lambda w' = \lambda \varphi(w).$
- φ è iniettiva, poiché se $w \in \text{Ker}(\varphi)$, $\varphi(w) = 0 \Rightarrow w = \underset{\in U}{u} + \underset{= \varphi(w)}{0} \Rightarrow w \in U \cap W = \{0\}$.
- Poiché $\dim W' = \dim V - \dim U = \dim W$ e φ è iniettiva, allora φ è surgettiva.

Infine evidentemente φ non dipende dalla scelta di una base, dunque ho la tesi.

Osservazione: Se $\pi_{W'}$ è la proiezione indotta da $V = U \oplus W'$ e $i_W: W \rightarrow V$ è l'inclusione (cioè $i_W(w) = w \ \forall w \in W$), allora $\varphi = \pi_{W'} \circ i_W$.

2.7 RANGO

PROPOSIZIONE 2.7.1: Siano $f: V \rightarrow W$, $g: W \rightarrow Z$ lineari. Allora:

- 1) $\dim \text{Im}(g \circ f) \leq \min(\dim \text{Im}(f), \dim \text{Im}(g))$;
- 2) Se f è un isomorfismo, $\dim \text{Im}(g \circ f) = \dim \text{Im}(g)$;
- 3) Se g è un isomorfismo, $\dim \text{Im}(g \circ f) = \dim \text{Im}(f)$.

Dimostrazione:

- 1) $\text{Im}(g \circ f) = \text{Im}(g|_{\text{Im}(f)})$ e $g|_{\text{Im}(f)}$ è lineare, poiché restrizione di g lineare.
Quindi $\dim \text{Im}(g \circ f) = \dim \text{Im}(f) - \dim \text{Ker}(g|_{\text{Im}(f)}) \leq \dim \text{Im}(f)$.
Inoltre $\text{Im}(g \circ f) \subseteq \text{Im}(g)$, quindi $\dim \text{Im}(g \circ f) \leq \dim \text{Im}(g)$, tesi.
- 2) Se f è un isomorfismo, allora $f(V) = W$, perciò $\text{Im}(g \circ f) = \text{Im}(g|_{\text{Im}(f)}) = \text{Im}(g)$, tesi.
- 3) Se g è un isomorfismo, allora $\text{Ker}(g) = \{0\}$, dunque $\dim \text{Im}(g \circ f) = \dim \text{Im}(f) - 0$, tesi.

DEFINIZIONE 2.7.1: Sia $f: V \rightarrow W$ lineare. Definiamo **rango** di f $rk(f) = \dim \text{Im}(f)$.

In particolare, se $A \in \mathcal{M}(p, n, \mathbb{K})$ e dunque $A: \mathbb{K}^n \rightarrow \mathbb{K}^p | X \rightarrow AX$, allora $rk(A) = \dim \text{Im}(A) = \dim \mathcal{C}(A)$.

DEFINIZIONE 2.7.2: Sia $A \in \mathcal{M}(p, n, \mathbb{K})$. Definiamo **spazio delle righe** $\mathcal{R}(A) = \text{Span}(A_1, \dots, A_p)$

Definiamo inoltre **rango per righe** il numero $\dim \mathcal{R}(A)$.

PROPOSIZIONE 2.7.2: Sia S una ridotta a scalini di A . Allora:

- 1) $\mathcal{R}(A) = \mathcal{R}(S)$;
- 2) $\dim \mathcal{C}(A) = \dim \mathcal{C}(S)$, cioè $rk(A) = rk(S)$.

Dimostrazione:

- 1) Facendo operazioni elementari di riga, non altero lo spazio delle righe, dunque $\mathcal{R}(A) = \mathcal{R}(S)$.
- 2) I sistemi $AX = 0$ e $SX = 0$ sono equivalenti, perciò $\text{Ker}(A) = \text{Ker}(S)$.
Per la formula delle dimensioni:
 $n = \dim \text{Ker}(A) + rk(A)$ e $n = \dim \text{Ker}(S) + rk(S)$,
dunque $rk(A) = rk(S)$.

PROPOSIZIONE 2.7.3: Sia S una matrice a scalini con r pivots nelle colonne S^{j_1}, \dots, S^{j_r} . Allora:

- 1) $\{S_1, \dots, S_r\}$ è una base di $\mathcal{R}(S)$, dunque $\dim \mathcal{R}(S) = r$;
- 2) $\{S^{j_1}, \dots, S^{j_r}\}$ è una base di $\mathcal{C}(S)$, dunque $rk(S) = r$.

Dimostrazione:

- 1) • Poiché le altre righe sono nulle, sicuramente S_1, \dots, S_r generano $\mathcal{R}(S)$.
• S_1, \dots, S_r sono linearmente indipendenti, poiché se:

$$a_1 S_1 + \dots + a_r S_r = 0$$

allora $a_1 = 0$ in quanto S_1 è l'unica riga ad avere un elemento $\neq 0$ nella colonna S^{j_1} ,
 $a_2 = 0$, in quanto $a_1 = 0$ e quindi S_2 è l'unica riga ad avere un elemento $\neq 0$ nella
colonna S^{j_2} , e così via.

- 2) S^{j_1}, \dots, S^{j_r} sono linearmente indipendenti, la dimostrazione è analoga alla precedente nel caso delle righe.

Mostriamo che S^{j_1}, \dots, S^{j_r} generano $\mathcal{C}(S)$, cioè che $S^i \in \text{Span}(S^{j_1}, \dots, S^{j_r}) \forall i \neq j_1, \dots, j_r$, cioè che $\exists a_1, \dots, a_r \in \mathbb{K} \mid S^i = a_1 S^{j_1} + \dots + a_r S^{j_r}$, cioè che il sistema

$$(S^{j_1} \mid \dots \mid S^{j_r}) \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} = (S^i)$$

ha soluzione.

Ma la matrice è a scalini, perciò ha r pivots.

Se aggiungo la colonna dei termini noti non posso dunque aggiungere pivots.

Dunque $\#pivots(S) = \#pivots(S')$, quindi il sistema è risolubile, tesi.

Osservazione: Se in qualche modo riesco a calcolare $rk(A)$, allora:

- so calcolare $\dim \text{Sol}(AX = 0) = \dim \text{Sol}_0$, poiché $\dim \text{Sol}_0 = \dim \text{Ker}(A) = n - rk(A)$;
- se voglio calcolare $\dim \text{Span}(v_1, \dots, v_k)$, pongo $A = (v_1 \mid \dots \mid v_k)$ e dunque $\dim \text{Span}(v_1, \dots, v_k) = \dim \mathcal{C}(A) = \dim \text{Im}(A) = rk(A)$.

Osservazione: Un modo per calcolare $rk(A)$ è via l'algoritmo di Gauss; infatti abbiamo visto che $rk(A) = \#pivots(S)$, dove S è una ridotta a scalini di A . Per formalizzare meglio questo procedimento, ci serviremo del seguente risultato:

PROPOSIZIONE 2.7.4: Sia $A \in \mathcal{M}(p, q)$. Allora:

- 1) il numero di pivots di una sua ridotta a scalini non dipende dalla riduzione a scalini;
- 2) $\dim \mathcal{R}(A) = \dim \mathcal{C}(A)$;

Dimostrazione:

- 1) Se S è una ridotta a scalini di A , con r pivots, allora $r = \dim \mathcal{R}(S) = \dim \mathcal{R}(A)$. Quindi r dipende solamente da A .
- 2) $\dim \mathcal{R}(A) = \dim \mathcal{R}(S) = r = \dim \mathcal{C}(S) = \dim \mathcal{C}(A) = rk(A)$.

COROLLARIO 2.7.5: $\forall A \in \mathcal{M}(p, q), rk({}^t A) = rk(A)$.

Dimostrazione:

$$rk({}^t A) = \dim \mathcal{C}({}^t A) = \dim \mathcal{R}(A) = rk(A).$$

TEOREMA DI ROUCHÉ – CAPELLI: $AX = B$ è risolubile $\Leftrightarrow rk(A) = rk(A')$,

dove $A' = (A \mid B)$.

Dimostrazione:

Se S è una ridotta a scalini di A , sapevamo che $AX = B$ è risolubile $\Leftrightarrow rk(S) = rk(S')$, ma $rk(S) = rk(A)$, dunque segue la tesi.

DEFINIZIONE 2.7.3: Una matrice $A \in \mathcal{M}(n, \mathbb{K})$ si dice **invertibile** se

$$\exists B \in \mathcal{M}(n, \mathbb{K}) \mid A \cdot B = B \cdot A = I.$$

DEFINIZIONE 2.7.4: Una matrice $A \in \mathcal{M}(n, \mathbb{K})$ si dice **singolare** se $rk(A) < n$.

PROPOSIZIONE 2.7.6: Sia $A \in \mathcal{M}(n, \mathbb{K})$. Allora sono fatti equivalenti:

- 1) A è invertibile;
- 2) $A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è un isomorfismo;

3) $rk(A) = n$.

Dimostrazione:

1) \Leftrightarrow 2): ovvia.

2) \Rightarrow 3): ovvia.

3) \Rightarrow 2): So che A è lineare e che A è surgettiva, in quanto $rk(A) = \dim Im(A) = n$, perciò A è iniettiva, dunque ho la tesi.

Osservazione: $A \in \mathcal{M}(n)$ è singolare \Leftrightarrow non è invertibile.

DEFINIZIONE 2.7.5: Definiamo **matrice elementare** di $\mathcal{M}(n, \mathbb{K})$ ogni matrice ottenuta da I_n eseguendo una sola operazione elementare per riga:

1° tipo: Denotiamo con E_{ij} la matrice ottenuta da I_n scambiando l' i -esima riga con la j -esima riga;

2° tipo: Denotiamo con $E_i(\lambda)$ la matrice ottenuta da I_n moltiplicando l' i -esima riga per la costante $\lambda \neq 0$;

3° tipo: $E_{ij}(\lambda)$ è la matrice ottenuta da I_n sommando alla riga i -esima λ volte la riga j -esima.

Osservazione: Se B è ottenuta da A con un'operazione elementare per riga, allora $B = EA$, dove E è la matrice elementare corrispondente all'operazione effettuata.

Esempio: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{A_2+3A_1} B = \begin{pmatrix} a & b \\ c+3a & d+3b \end{pmatrix}$;

$E \cdot A = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c+3a & d+3b \end{pmatrix}$.

Osservazione: Le matrici elementari sono tutte invertibili:

1) $E_{ij} \cdot E_{ij} = I$, poiché scambio le righe i e j e poi le riscambio; perciò $(E_{ij})^{-1} = E_{ij}$.

2) $E_i(\lambda^{-1}) \cdot E_i(\lambda) = I$, poiché moltiplico la riga i -esima prima per λ^{-1} e poi per λ ; perciò $(E_i(\lambda))^{-1} = E_i(\lambda^{-1})$.

3) $E_{ij}(\lambda) \cdot E_{ij}(-\lambda) = I$, dunque $(E_{ij}(\lambda))^{-1} = E_{ij}(-\lambda)$.

Quindi se prendo una matrice $A \in \mathcal{M}(p, q, \mathbb{K})$ e gli applico n operazioni di riga, ottengo la ridotta a scalini S :

$$A \rightarrow M_1 A \rightarrow M_2 M_1 A \rightarrow \dots \rightarrow M_n \dots M_1 A = S$$

Se A è invertibile, allora è un isomorfismo; gli M_i sono tutti isomorfismi, e la composizione di isomorfismi è un isomorfismo, perciò S è invertibile.

Dunque, detta $M = M_n \dots M_1 \in GL(p)$:

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{A} & \mathbb{K}^p \xrightarrow{M} \mathbb{K}^p \\ & & \mathbb{K}^n \xrightarrow{S=M \cdot A} \mathbb{K}^p \end{array}$$

Sappiamo che $S^j = M \cdot A^j$, ma M è un isomorfismo, quindi trasforma basi in basi.

Inoltre $\{S^{j_1}, \dots, S^{j_r}\}$ è una base di $Im(S)$, dunque $\{A^{j_1}, \dots, A^{j_r}\}$ è una base di $Im(A)$.

ALGORITMO PER L'ESTRAZIONE DI UNA BASE DA UN GRUPPO DI GENERATORI: Dati $v_1, \dots, v_k \in \mathbb{K}^n$, sia $A = (v_1 \mid \dots \mid v_k)$. Detta S una ridotta a scalini di A , se S^{j_1}, \dots, S^{j_r} sono le colonne contenenti i pivots di S , allora $\{v_{j_1}, \dots, v_{j_r}\}$ è una base $Span(v_1, \dots, v_k)$.

Osservazione: Per estendere $v_1, \dots, v_m \in \mathbb{K}^n$ linearmente indipendenti a base di \mathbb{K}^n costruiamo $A = (v_1 \mid \dots \mid v_m \mid e_1 \mid \dots \mid e_n)$. Poiché i vettori in colonna generano \mathbb{K}^n , applicando l'algoritmo precedente estendo v_1, \dots, v_m a base di \mathbb{K}^n .

Osservazione: Sia $A \in \mathcal{M}(p, n, \mathbb{K})$. So che $\exists M \in GL(p) \mid MA = S$ a scalini.

Per trovare una tale M riduco $(A \mid I_p)$ a scalini fino a ottenere $(S \mid B)$.

Sicuramente $\exists M \in GL(p) \mid M(A \mid I_p) = (S \mid B)$; allora:

$$\begin{cases} MA = S \\ MI = B \end{cases} \quad \begin{cases} M = B \\ BA = S \end{cases}$$

dunque la matrice cercata è B .

CALCOLO DELL'INVERSA: Sia $A \in \mathcal{M}(n, \mathbb{K})$. Riduco per righe $(A \mid I_n)$ fino a $(S \mid *)$, con S a scalini. Poiché $rk(A) = rk(S)$, A è invertibile $\Leftrightarrow rk(S) = n$.

Se $rk(S) < n$, l'algoritmo si ferma, poiché A non è invertibile.

Se $rk(S) = n$, proseguo con la riduzione fino a ottenere $(I \mid B)$.

Per l'osservazione precedente, $BA = I$, cioè B è un'inversa sinistra di A . Ma essendo A invertibile, allora B è anche inversa destra: $B = A^{-1}$.

Notazione: Sia $A \in \mathcal{M}(p, n)$. Denotiamo con $(A_{i_1}, \dots, A_{i_m} \mid A^{j_1}, \dots, A^{j_q})$ la sottomatrice di A ottenuta in modo che contenga gli elementi nelle intersezioni fra le righe e le colonne considerate.

DEFINIZIONE 2.7.6: Una sottomatrice quadrata si dice **minore**.

PROPOSIZIONE 2.7.7: Sia $A \in \mathcal{M}(p, n)$. Sia B un minore invertibile di A .

Allora le righe (o le colonne) che concorrono a formare B sono linearmente indipendenti.

Dimostrazione:

Sia $B = (A_{i_1}, \dots, A_{i_q} \mid A^{j_1}, \dots, A^{j_q})$.

Se $\alpha_1 A_{i_1} + \dots + \alpha_q A_{i_q} = 0$, a maggior ragione $\alpha_1 B_1 + \dots + \alpha_q B_q = 0$.

Ma B_1, \dots, B_q sono linearmente indipendenti perché $rk(B) = q$, quindi $\alpha_1 = \dots = \alpha_q = 0$.

TEOREMA 2.7.8: Il rango di una matrice coincide con il massimo degli ordini dei suoi minori invertibili.

Dimostrazione:

Sia $A \in \mathcal{M}(p, n)$; sia $r = rk(A)$; sia ρ il massimo degli ordini dei minori di A invertibili.

- $\rho \leq r$: Sia B un minore $\rho \times \rho$ di A invertibile. Allora, per la proposizione precedente, esistono in A ρ righe indipendenti, quindi $r \geq \rho$.

- $\rho \geq r$: Siano A_{i_1}, \dots, A_{i_r} r righe indipendenti in A . Allora ho la sottomatrice $B = (A_{i_1}, \dots, A_{i_r} \mid A^1, \dots, A^n)$ di rango r .

Allora il rango per colonne è r , dunque esistono in B r colonne indipendenti B^{j_1}, \dots, B^{j_r} .

Allora la sottomatrice M di B , $M = (B_1, \dots, B_r \mid B^{j_1}, \dots, B^{j_r})$ ha rango r , cioè è un minore $r \times r$ invertibile di A .

Dunque $\rho \geq r$, da cui la tesi.

DEFINIZIONE 2.7.7: Sia $B = (A_{i_1}, \dots, A_{i_q} | A^{j_1}, \dots, A^{j_q})$ un minore di A . Definiamo minore orlato di B un qualunque minore $B' = (A_{i_1}, \dots, A_{i_q}, A_h | A^{j_1}, \dots, A^{j_q}, A^k)$, con $h \neq i_1, \dots, i_q$ e $k \neq j_1, \dots, j_q$.

TEOREMA DEGLI ORLATI: Sia $A \in \mathcal{M}(m, n, \mathbb{K})$. Allora $rk(A) = k \Leftrightarrow \exists$ un minore $k \times k$ invertibile i cui orlati sono tutti non invertibili.

Dimostrazione:

\Rightarrow) Sappiamo che se $rk(A) = k$ allora esiste un minore $k \times k$ invertibile e tutti i minori $h \times h$, con $h > k$, sono non invertibili. Gli orlati appartengono a questo tipo di minori, dunque segue la tesi.

\Leftarrow) Se esiste un minore Q $k \times k$ invertibile, so che $rk(A) \geq k$.

Dunque devo mostrare che se tutti gli orlati di Q sono non invertibili, effettivamente non può essere $rk(A) > k$.

Ovvero se $rk(A) > k$, trovo un orlato di Q invertibile.

Sia Q dato dalle righe R_{i_1}, \dots, R_{i_k} e dalle colonne C_{j_1}, \dots, C_{j_k} .

L'invertibilità di Q implica che la matrice $(C_{j_1} | \dots | C_{j_k})$ ha rango k .

Dunque C_{j_1}, \dots, C_{j_k} sono elementi di \mathbb{K}^n linearmente indipendenti.

Se $rk(A) > k$, allora $\dim \text{Span}(C_1, \dots, C_m) > k$, per cui $\exists C_{j_{k+1}} | C_{j_1}, \dots, C_{j_k}, C_{j_{k+1}}$ sono linearmente indipendenti.

Quindi $rk(C_{j_1} | \dots | C_{j_k} | C_{j_{k+1}}) = k + 1$.

Inoltre le righe i_1, \dots, i_k di questa matrice sono linearmente indipendenti, perché identificano una sottomatrice che contiene Q .

Dunque \exists riga $R_{i_{k+1}} |$ le righe $R_{i_1}, \dots, R_{i_k}, R_{i_{k+1}}$ di questa matrice $(k + 1) \times (k + 1)$ sono linearmente indipendenti.

Questo è un orlato invertibile di Q , dunque ho la tesi.

2.8 SD-EQUIVALENZA

DEFINIZIONE 2.8.1: Sia $f: V \rightarrow W$ lineare e siano $\mathcal{S} = \{v_1, \dots, v_n\}$ una base di V e $\mathcal{T} = \{w_1, \dots, w_m\}$ una base di W . Definiamo **matrice associata a f rispetto a \mathcal{S} e \mathcal{T}** :

$$\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f) = ([f(v_1)]_{\mathcal{T}} | \dots | [f(v_n)]_{\mathcal{T}})$$

dove $[f(v_i)]_{\mathcal{T}} = \begin{pmatrix} \alpha_{i,1} \\ \vdots \\ \alpha_{i,m} \end{pmatrix}$ sono le coordinate di $f(v_i)$ rispetto a \mathcal{T} , cioè tali che

$$f(v_i) = \alpha_{i,1}w_1 + \dots + \alpha_{i,m}w_m.$$

Osservazione: Sia $v \in V$; allora $v = x_1v_1 + \dots + x_nv_n$.

$$\begin{aligned} f(v) &= x_1f(v_1) + \dots + x_nf(v_n) = x_1(\alpha_{1,1}w_1 + \dots + \alpha_{1,m}w_m) + \dots + x_n(\alpha_{n,1}w_1 + \dots + \alpha_{n,m}w_m) = \\ &= (\alpha_{1,1}x_1 + \dots + \alpha_{n,1}x_n)w_1 + \dots + (\alpha_{1,m}x_1 + \dots + \alpha_{n,m}x_n)w_m, \text{ cioè:} \end{aligned}$$

$$[f(v)]_{\mathcal{T}} = \begin{pmatrix} \alpha_{1,1}x_1 + \dots + \alpha_{n,1}x_n \\ \vdots \\ \alpha_{1,m}x_1 + \dots + \alpha_{n,m}x_n \end{pmatrix} = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f) \cdot [v]_{\mathcal{S}}$$

TEOREMA 2.8.1: Siano V, W spazi vettoriali tali che $\dim V = n, \dim W = m$.

Sia $\mathcal{S} = \{v_1, \dots, v_n\}$ base di V e $\mathcal{T} = \{w_1, \dots, w_m\}$ base di W .

Allora l'applicazione:

$$\mathfrak{M}_{\mathcal{S}, \mathcal{T}}: \text{Hom}(V, W) \rightarrow \mathcal{M}(m, n, \mathbb{K}) \mid f \rightarrow \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$$

è un isomorfismo.

Dimostrazione:

- $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}$ è evidentemente lineare;
- È iniettiva, poiché se $f \in \text{Ker}(\mathfrak{M}_{\mathcal{S}, \mathcal{T}})$, $[\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)]^1 = \dots = [\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)]^n = 0$, cioè $f(v_1) = \dots = f(v_n) = 0$, e per il teorema che dice che $\exists!$ applicazione lineare che manda una base in vettori preassegnati, allora f è l'applicazione nulla.
- $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}$ è surgettiva, poiché $\forall A \in \mathcal{M}(m, n) \exists!$ $f: V \rightarrow W$ lineare tale che $[f(v_1)]_{\mathcal{T}} = A^1, \dots, [f(v_n)]_{\mathcal{T}} = A^n$, dunque $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f) = A$.

COROLLARIO 2.8.2: Siano V, W spazi vettoriali tali che $\dim V = n, \dim W = m$. Allora $\dim \text{Hom}(V, W) = m \cdot n$.

Dimostrazione:

Segue dal fatto che \forall basi di V e W , l'applicazione $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}: \text{Hom}(V, W) \rightarrow \mathcal{M}(m, n, \mathbb{K})$ è un isomorfismo e $\dim \mathcal{M}(m, n, \mathbb{K}) = m \cdot n$.

Notazione: Se V è uno spazio vettoriale e \mathcal{B} è base di V , denotiamo con $V_{\mathcal{B}}$ lo spazio V rispetto alla base \mathcal{B} .

PROPOSIZIONE 2.8.3: Siano $f: U_{\mathcal{S}} \rightarrow V_{\mathcal{T}}$ e $g: V_{\mathcal{T}} \rightarrow W_{\mathcal{R}}$ lineari e siano $A = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$ e $B = \mathfrak{M}_{\mathcal{T}, \mathcal{R}}(g)$. Allora $\mathfrak{M}_{\mathcal{S}, \mathcal{R}}(g \circ f) = B \cdot A$.

Dimostrazione:

$\forall u \in U, [(g \circ f)(u)]_{\mathcal{R}} = [g(f(u))]_{\mathcal{R}} = B \cdot [f(u)]_{\mathcal{T}} = B \cdot A \cdot [u]_{\mathcal{S}}$, da cui la tesi.

Osservazione: Se \mathcal{S} è base di V, \mathcal{T} è base di W e $A = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$, allora il diagramma

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow [\]_{\mathcal{S}} & & \downarrow [\]_{\mathcal{T}} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^m \end{array}$$

è commutativo, cioè per "andare" da uno spazio all'altro si può seguire un qualsiasi percorso (dunque ad esempio $A \circ [\]_{\mathcal{S}} = [\]_{\mathcal{T}} \circ f$).

Inoltre il diagramma:

$$\begin{array}{ccccc} & & \overbrace{f \quad g}^{g \circ f} & & \\ V & \xrightarrow{f} & W & \xrightarrow{g} & Z \\ \downarrow [\]_{\mathcal{S}} & & \downarrow [\]_{\mathcal{T}} & & \downarrow [\]_{\mathcal{R}} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^m & \xrightarrow[B]{} & \mathbb{K}^p \\ & & \underbrace{A \quad B}_{B \cdot A} & & \end{array}$$

è commutativo.

PROPOSIZIONE 2.8.4: Sia $f: V \rightarrow W$ lineare, $A = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$. Allora $\text{rk}(f) = \text{rk}(A)$.

Dimostrazione:

Sia $\mathcal{S} = \{v_1, \dots, v_n\}$; allora $\text{Im}(f) = \text{Span}(f(v_1), \dots, f(v_n))$.

Inoltre $[f(v_i)]_{\mathcal{T}} = A^i \forall i$.

Se $\varphi = []_{\mathcal{T}}: W \rightarrow \mathbb{K}^m$ è l'isomorfismo indotto dalla base \mathcal{T} , allora $\varphi(Im(f)) = \mathcal{C}(A) = Im(A)$.

Per cui $rk(f) = \dim Im(f) = \dim Im(A) = rk(A)$.

Osservazione: Per la proposizione precedente, se $\{A^{j_1}, \dots, A^{j_r}\}$ è una base di $Im(A)$, allora $\{f(v_{j_1}), \dots, f(v_{j_r})\}$ è una base di $Im(f)$.

COROLLARIO 2.8.5: Sia $f: V \rightarrow W$ lineare, $\dim V = \dim W = n$.

Allora f è invertibile $\Leftrightarrow A = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$ è invertibile.

Dimostrazione:

f è invertibile $\Leftrightarrow rk(f) = n \Leftrightarrow rk(A) = n \Leftrightarrow A$ è invertibile.

DEFINIZIONE 2.8.2: Siano \mathcal{S}, \mathcal{T} basi di V . Definiamo **matrice del cambiamento di base** da \mathcal{S} a \mathcal{T} la matrice $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(id)$.

Osservazioni: 1) Se $N = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(id)$ e $v \in V$, allora $[v]_{\mathcal{T}} = N \cdot [v]_{\mathcal{S}}$. Dunque N trasforma le coordinate di v rispetto a \mathcal{S} nelle coordinate di v rispetto a \mathcal{T} .

2) Evidentemente $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(id) \cdot \mathfrak{M}_{\mathcal{T}, \mathcal{S}}(id) = I$, dunque $\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(id)$ è invertibile e

$$\left(\mathfrak{M}_{\mathcal{S}, \mathcal{T}}(id)\right)^{-1} = \mathfrak{M}_{\mathcal{T}, \mathcal{S}}(id).$$

3) Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è base di V , allora $[]_{\mathcal{B}}(v_i) = e_i$, cioè $[]_{\mathcal{B}}$ trasforma \mathcal{B} nella base canonica di \mathbb{K}^n .

4) Se $g: V \rightarrow \mathbb{K}^n$ è un isomorfismo, allora $\exists!$ base \mathcal{B} di V tale che $g = []_{\mathcal{B}}$.

Infatti, per l'osservazione precedente, $\mathcal{B} = \{g^{-1}(e_1), \dots, g^{-1}(e_n)\}$, dunque è unica.

PROPOSIZIONE 2.8.6: Sia V uno spazio vettoriale, $\dim V = n$, \mathcal{B} base di V , $A \in GL(n)$. Allora:

1) $\exists!$ base \mathcal{S} di $V \mid A = \mathfrak{M}_{\mathcal{S}, \mathcal{B}}(id)$;

2) $\exists!$ base \mathcal{T} di $V \mid A = \mathfrak{M}_{\mathcal{B}, \mathcal{T}}(id)$.

Dimostrazione:

1) Le ipotesi creano una situazione del genere:

$$\begin{array}{ccc} V & \xrightarrow{id} & V \\ & & \downarrow []_{\mathcal{B}} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^n \end{array}$$

A è un isomorfismo, dunque anche A^{-1} è un isomorfismo.

Allora $\exists!$ isomorfismo $g: V \rightarrow \mathbb{K}^n \mid$ il diagramma:

$$\begin{array}{ccc} V & \xrightarrow{id} & V \\ g \downarrow & & \downarrow []_{\mathcal{B}} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^n \end{array}$$

commuti. In particolare $g = A^{-1} \circ []_{\mathcal{B}}$.

Per l'osservazione 4) $\exists!$ base \mathcal{S} di V tale che $g = []_{\mathcal{S}}$ (che dunque sarà

$$\mathcal{S} = \{g^{-1}(e_1), \dots, g^{-1}(e_n)\} = \{[]_{\mathcal{B}}^{-1}(A(e_1)), \dots, []_{\mathcal{B}}^{-1}(A(e_n))\} = \{[A^1]_{\mathcal{B}}^{-1}, \dots, [A^n]_{\mathcal{B}}^{-1}\}.$$

2) $\exists!$ $g = A \circ []_{\mathcal{B}}$ isomorfismo che rende commutativo il diagramma:

$$\begin{array}{ccc} V & \xrightarrow{id} & V \\ []_B \downarrow & & \downarrow g \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^n \end{array}$$

Per l'osservazione 4) $\exists!$ base \mathcal{T} di V tale che $g = []_{\mathcal{T}}$.

Osservazione: Sia $f: V \rightarrow W$ lineare, $\mathcal{S}, \mathcal{S}'$ basi di V , $\mathcal{T}, \mathcal{T}'$ basi di W . Siano $A = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f)$ e $A' = \mathfrak{M}_{\mathcal{S}', \mathcal{T}'}(f)$. Siano inoltre $N = \mathfrak{M}_{\mathcal{S}', \mathcal{S}}(id)$ e $M = \mathfrak{M}_{\mathcal{T}, \mathcal{T}'}(id)$.

La situazione dei dati è dunque:

$$\begin{array}{ccc} V_{\mathcal{S}} & \xrightarrow{A} & W_{\mathcal{T}} \\ N \uparrow & & \downarrow M \\ V_{\mathcal{S}'} & \xrightarrow{A'} & W_{\mathcal{T}'} \end{array}$$

Il diagramma è commutativo e dunque $A' = MAN$, ossia:

$$\mathfrak{M}_{\mathcal{S}', \mathcal{T}'}(f) = \mathfrak{M}_{\mathcal{T}, \mathcal{T}'}(id) \cdot \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(f) \cdot \mathfrak{M}_{\mathcal{S}', \mathcal{S}}(id).$$

DEFINIZIONE 2.8.3: $f, g \in Hom(V, W)$. f e g si dicono **SD-equivalenti** ($f \equiv_{SD} g$) $\Leftrightarrow \exists h \in GL(W), \exists k \in GL(V) | g = h \circ f \circ k$.

Osservazioni: 1) \equiv_{SD} è una relazione di equivalenza (la verifica è lasciata al lettore);
2) Se $f \equiv_{SD} g$, allora $rk(f) = rk(g)$, poiché componendo isomorfismi il rango non cambia (il rango è dunque un invariante per \equiv_{SD});

DEFINIZIONE 2.8.4: $A, B \in \mathcal{M}(p, n, \mathbb{K})$. $A \equiv_{SD} B \Leftrightarrow \exists M \in GL(p), \exists N \in GL(n) | B = MAN$.

Osservazione: Dalle definizioni segue immediatamente che, se $A = \mathfrak{M}_{B, \mathcal{S}}(f), B = \mathfrak{M}_{B, \mathcal{S}}(g)$, allora $f \equiv_{SD} g \Leftrightarrow A \equiv_{SD} B$.

Osservazione: Se $B = MAN$, con M, N invertibili, posso vedere A come un'applicazione lineare: $A = \mathfrak{M}_{\mathcal{C}, \mathcal{C}}(A)$, dove \mathcal{C} è la base canonica.

Inoltre, se interpreto $N = \mathfrak{M}_{\mathcal{S}, \mathcal{C}}(id_{\mathbb{K}^n}), M = \mathfrak{M}_{\mathcal{C}, \mathcal{T}}(id_{\mathbb{K}^p})$:

$$\mathbb{K}_{\mathcal{S}}^n \xrightarrow{N} \mathbb{K}_{\mathcal{C}}^n \xrightarrow[A]{} \mathbb{K}_{\mathcal{C}}^p \xrightarrow[M]{} \mathbb{K}_{\mathcal{T}}^p$$

Allora $B = \mathfrak{M}_{\mathcal{S}, \mathcal{T}}(A)$.

Per cui $A \equiv_{SD} B \Leftrightarrow$ rappresentano la stessa applicazione lineare in basi diverse.

Estendiamo questa osservazione al caso delle applicazioni lineari con la seguente proposizione:

PROPOSIZIONE 2.8.7: $f, g \in Hom(V, W)$. Allora:

$$f \equiv_{SD} g \Leftrightarrow \exists \mathcal{B}, \mathcal{B}' \text{ basi di } V, \exists \mathcal{S}, \mathcal{S}' \text{ basi di } W \text{ tali che } \mathfrak{M}_{\mathcal{B}', \mathcal{S}'}(f) = \mathfrak{M}_{\mathcal{B}, \mathcal{S}}(g).$$

Dimostrazione:

\Rightarrow) Fisso \mathcal{B} base di V e \mathcal{S} base di W .

Per ipotesi $\exists h \in GL(W), \exists k \in GL(V) | g = h \circ f \circ k$:

$$\begin{array}{ccccccc}
 & & \overbrace{\hspace{10em}}^{g=h \circ f \circ k} & & & & \\
 & & \xrightarrow{k} & & \xrightarrow{f} & & \xrightarrow{h} \\
 V & \rightarrow & V & \rightarrow & W & \rightarrow & W \\
 []_{\mathcal{B}} \downarrow & & []_{\mathcal{B}} \downarrow & & []_{\mathcal{S}} \downarrow & & \downarrow []_{\mathcal{S}} \\
 \mathbb{K}^n & \xrightarrow{N} & \mathbb{K}^n & \xrightarrow{A} & \mathbb{K}^p & \xrightarrow{M} & \mathbb{K}^p
 \end{array}$$

Allora $\mathfrak{M}_{\mathcal{B},\mathcal{S}}(g) = MAN$, con $A = \mathfrak{M}_{\mathcal{B},\mathcal{S}}(f)$.

Interpreto N e M come matrici del cambiamento di base:

$\exists! \mathcal{B}'$ base di $V \mid \mathfrak{M}_{\mathcal{B}',\mathcal{B}}(id_V) = N$,

$\exists! \mathcal{S}'$ base di $W \mid \mathfrak{M}_{\mathcal{S},\mathcal{S}'}(id_W) = M$.

Allora $\mathfrak{M}_{\mathcal{B}',\mathcal{S}'}(f) = \mathfrak{M}_{\mathcal{S},\mathcal{S}'}(id_W) \cdot \mathfrak{M}_{\mathcal{B},\mathcal{S}}(f) \cdot \mathfrak{M}_{\mathcal{B}',\mathcal{B}}(id_V) = MAN = \mathfrak{M}_{\mathcal{B},\mathcal{S}}(g)$.

\Leftrightarrow Analogo.

PROPOSIZIONE 2.8.8: $f: V \rightarrow W$ lineare, $\dim V = n$, $\dim W = p$, $rk(f) = r$.

Allora $\exists \mathcal{B}$ base di V , $\exists \mathcal{S}$ base di W :

$$\mathfrak{M}_{\mathcal{B},\mathcal{S}}(f) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix} \in \mathcal{M}(p, n, \mathbb{K})$$

Dimostrazione:

$\dim Ker(f) = n - r$.

Sia $\{v_{r+1}, \dots, v_n\}$ una base di $Ker(f)$.

Sia $\mathcal{B} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ base di V .

Allora $\{f(v_1), \dots, f(v_r)\}$ è una base di $Im(f)$ (in quanto $f(v_i) = 0 \ \forall r+1 \leq i \leq n$).

La completo a $\mathcal{S} = \{f(v_1), \dots, f(v_r), w_{r+1}, \dots, w_p\}$ base di W .

Allora \mathcal{B} e \mathcal{S} verificano la tesi.

TEOREMA 2.8.9: $f \equiv_{SD} g \Leftrightarrow rk(f) = rk(g)$.

Dimostrazione:

\Rightarrow) Già vista.

\Leftarrow) Se $rk(f) = rk(g) = r$, allora per la proposizione precedente $\exists \mathcal{B}, \mathcal{S}$ basi $\mid \mathfrak{M}_{\mathcal{B},\mathcal{S}}(g) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}$,

$$\exists \mathcal{B}', \mathcal{S}' \text{ basi} \mid \mathfrak{M}_{\mathcal{B}',\mathcal{S}'}(f) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}.$$

Poiché \equiv_{SD} è una relazione di equivalenza, ho la tesi.

Osservazione: Il rango è dunque un invariante completo per \equiv_{SD} , in altre parole l'insieme quoziente $Hom(V, W) / \equiv_{SD}$ ha $r = \min(\dim V, \dim W) + 1$ classi di equivalenza, in quanto il rango di una matrice $m \cdot n$ può oscillare fra 0 e $\min(\dim V, \dim W)$.

Esprimiamo le due precedenti proposizioni anche a livello matriciale:

PROPOSIZIONE 2.8.10: $A \in \mathcal{M}(p, n, \mathbb{K})$. Se $rk(A) = r \Rightarrow A \equiv_{SD} \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}$.

TEOREMA 2.8.11: $A, B \in \mathcal{M}(p, n, \mathbb{K})$. Allora $A \equiv_{SD} B \Leftrightarrow rk(A) = rk(B)$.

Si ritrova dunque il seguente risultato:

COROLLARIO 2.8.12: $rk(A) = rk({}^tA)$.

Dimostrazione:

Denotiamo $J_r(p, n) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix} \in \mathcal{M}(p, n)$.

Se $rk(A) = r$, $\exists M, N$ invertibili $| A = M \cdot J_r(p, n) \cdot N$.

Dunque ${}^tA = {}^tN \cdot {}^tJ_r(p, n) \cdot {}^tM = {}^tN \cdot J_r(n, p) \cdot {}^tM$.

Ma tM e tN sono invertibili, dunque ${}^tA \equiv_{SD} J_r(n, p)$.

Poiché $rk(J_r(n, p)) = r$, segue che $rk({}^tA) = r$.

2.9 SPAZIO DUALE

DEFINIZIONE 2.9.1: Sia V un \mathbb{K} -spazio vettoriale. Si definisce **spazio duale** $V^* = Hom(V, \mathbb{K})$.

Gli elementi v_i^* di V^* sono detti **funzionali lineari**.

PROPOSIZIONE 2.9.1: Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ base di V . $\forall i$, sia $v_i^*: V \rightarrow \mathbb{K}$ il funzionale definito da $v_i^*(v_j) = \delta_{ij}$, dove δ_{ij} è il delta di Kronecker.

Allora $\mathcal{B}^* = \{v_1^*, \dots, v_n^*\}$ è base di V^* , detta **base duale** di \mathcal{B} .

Dimostrazione:

- I v_i^* sono linearmente indipendenti, infatti se $a_1v_1^* + \dots + a_nv_n^* = 0$, allora $(a_1v_1^* + \dots + a_nv_n^*)(v_j) = 0 \ \forall j$. Poiché $0 = (a_1v_1^* + \dots + a_nv_n^*)(v_j) = (a_jv_j^*)(v_j) = a_j \ \forall j$, concludiamo che $a_1 = \dots = a_n = 0$.
- Dimostriamo che generano:
sia $f \in V^*$; cerco $a_1, \dots, a_n \in \mathbb{K} \mid f = a_1v_1^* + \dots + a_nv_n^*$.
Poiché $\forall j, f(v_j) = (a_1v_1^* + \dots + a_nv_n^*)(v_j) = a_j$, basta scegliere $a_i = f(v_i)$ e ottengo la tesi.

Osservazione: Per la dimostrazione precedente, $[f]_{\mathcal{B}^*} = \begin{pmatrix} f(v_1) \\ \vdots \\ f(v_n) \end{pmatrix}$.

DEFINIZIONE 2.9.2: Si definisce spazio **biduale** $V^{**} = (V^*)^* = Hom(V^*, \mathbb{K})$.

Osservazione: $\dim V = \dim V^* = \dim V^{**}$.

Notazione: Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V . Poniamo $\varphi_{\mathcal{B}}: V \rightarrow V^* \mid v_i \rightarrow \varphi_{\mathcal{B}}(v_i) = v_i^* \ \forall i$.

Quindi:

$$V \xrightarrow{\varphi_{\mathcal{B}}} V^* \xrightarrow{\varphi_{\mathcal{B}^*}} V^{**}$$

TEOREMA 2.9.2: L'applicazione $\psi_V: V \rightarrow V^{**} \mid v \rightarrow \psi_V(v)$, dove

$\psi_V(v): V^* \rightarrow \mathbb{K} \mid g \rightarrow \psi_V(v)(g) = g(v)$:

- 1) è un isomorfismo canonico;
- 2) \forall base \mathcal{B} di V , $\varphi_{\mathcal{B}^*} \circ \varphi_{\mathcal{B}} = \psi_V$.

Dimostrazione:

- 1) Dimostriamo innanzitutto che effettivamente $\psi_V(v) \in V^{**}$, cioè che $\psi_V(v)$ è lineare $\forall v$:
 $\forall v \in V, \forall \lambda, \mu \in \mathbb{K}, \forall f, g \in V^*, \psi_V(v)(\lambda f + \mu g) = (\lambda f + \mu g)(v) = \lambda f(v) + \mu g(v) = \lambda \psi_V(v)(f) + \mu \psi_V(v)(g)$, che implica che $\psi_V(v)$ è lineare perché conserva le combinazioni lineari; dunque $\psi_V: V \rightarrow V^{**}$ è ben definita.

Lasciamo la verifica che ψ_V è lineare, cioè che $\forall v_1, v_2 \in V, \psi_V(a_1 v_1 + a_2 v_2) = a_1 \psi_V(v_1) + a_2 \psi_V(v_2)$, cioè che $\psi_V(a_1 v_1 + a_2 v_2)(g) = (a_1 \psi_V(v_1) + a_2 \psi_V(v_2))(g) \quad \forall g \in V^*$.

Poiché $\dim V^{**} = \dim V$, dimostriamo solo l'iniettività di ψ_V :

sia $v \in \text{Ker}(\psi_V) \Rightarrow \psi_V(v) = 0 \Rightarrow \psi_V(v)(g) = g(v) = 0 \quad \forall g \in V^* \Rightarrow v = 0$, poiché se $v \neq 0, \exists g: V \rightarrow \mathbb{K}$ lineare | $g(v) \neq 0$, assurdo.

Dunque ψ_V è un isomorfismo ed evidentemente non dipende da nessuna base.

- 2) Devo mostrare che, fissata $B = \{v_1, \dots, v_n\}$, $(\varphi_{B^*} \circ \varphi_B)(v_i) = \psi_V(v_i) \quad \forall i$, poiché se due applicazioni lineari coincidono su una base, evidentemente coincidono su qualunque elemento dello spazio e dunque sono uguali.

Ma poiché $(\varphi_{B^*} \circ \varphi_B)(v_i)$ e $\psi_V(v_i)$ sono due funzionali di V^{**} , per mostrare che sono uguali bisogna far vedere che coincidono su una base di V^* , cioè che

$$(\varphi_{B^*} \circ \varphi_B)(v_i)(v_j^*) = \psi_V(v_i)(v_j^*) \quad \forall j.$$

Ora:

$$(\varphi_{B^*} \circ \varphi_B)(v_i)(v_j^*) = (\varphi_{B^*}(\varphi_B(v_i)))(v_j^*) = (\varphi_{B^*}(v_i^*))(v_j^*) = (v_i^{**})(v_j^*) = \delta_{ij};$$

$$\psi_V(v_i)(v_j^*) = (v_j^*)(v_i) = \delta_{ij}, \text{ dunque ho la tesi.}$$

DEFINIZIONE 2.9.3: Sia $S \subset V$. Si definisce **annullatore** di S $\text{Ann}(S) = \{f \in V^* | f|_S \equiv 0\}$.

PROPOSIZIONE 2.9.3: 1) $\forall S \subset V, \text{Ann}(S)$ è sottospazio vettoriale di V^*

2) $S \subseteq T \Rightarrow \text{Ann}(T) \subseteq \text{Ann}(S)$

3) Se U è sottospazio vettoriale di V e $\dim U = k \Rightarrow \dim \text{Ann}(U) = n - k$

4) $\forall f \in V^*, \text{Ann}(f) = \psi_V(\text{Ker}(f))$

5) $\forall U$ sottospazio vettoriale di $V, \text{Ann}(\text{Ann}(U)) = \psi_V(U)$.

Dimostrazione:

1) È una semplice verifica.

2) Se $f \in \text{Ann}(T) \Rightarrow f(v) = 0 \quad \forall v \in T \supseteq S \Rightarrow f \in \text{Ann}(S)$.

3) Sia $\{u_1, \dots, u_k\}$ base di U . La completo a $\{u_1, \dots, u_k, v_{k+1}, \dots, v_n\}$ base di V . Provo che $\{v_{k+1}^*, \dots, v_n^*\}$ è base di $\text{Ann}(U)$:

- Sicuramente $v_i^* \in \text{Ann}(U) \quad \forall i \geq k + 1$, poiché $v_i^*(u_j) = 0 \quad \forall j \leq k$;

- v_{k+1}^*, \dots, v_n^* sono linearmente indipendenti perché elementi della base duale;

- Mostriamo ora che v_{k+1}^*, \dots, v_n^* generano:

Sia $f \in \text{Ann}(U) \subset V^* \Rightarrow \exists a_1, \dots, a_n \in \mathbb{K} | f = a_1 u_1^* + \dots + a_k u_k^* + a_{k+1} v_{k+1}^* + \dots + a_n v_n^*$.

Poiché $f \in \text{Ann}(U)$, allora $f(u_i) = 0 \quad \forall i \leq k$, quindi $a_1 = \dots = a_k = 0$.

Dunque $f = a_{k+1} v_{k+1}^* + \dots + a_n v_n^*$, da cui la tesi.

4) $\text{Ann}(f) = \{h \in V^* | h(f) = 0\} = \{\psi_V(x) \in V^{**} | \psi_V(x)(f) = f(x) = 0\} = \psi_V(\{x \in V | f(x) = 0\}) = \psi_V(\text{Ker}(f))$.

5) Poiché $\dim \psi_V(U) = \dim U = n - \dim \text{Ann}(U) = n - n + \dim \text{Ann}(\text{Ann}(U)) = \dim \text{Ann}(\text{Ann}(U))$, dimostro solo che $\psi_V(U) \subseteq \text{Ann}(\text{Ann}(U))$:

$\forall x \in U, \psi_V(x)|_{\text{Ann}(U)} = 0$, perché $\forall f \in \text{Ann}(U), \psi_V(x)(f) = f(x) = 0$.

Notazione: Al posto di $\psi_V(U)$ scriveremo semplicemente U .

DEFINIZIONE 2.9.4: Sia $f: V \rightarrow W$ lineare. Definiamo **trasposta** di f :

$${}^t f: W^* \rightarrow V^* \mid {}^t f(g) = g \circ f.$$

Osservazione: È una buona definizione, poiché se $g: W \rightarrow \mathbb{K}$, allora $g \circ f: V \rightarrow W \rightarrow \mathbb{K}$, cioè $g \circ f \in V^*$.

PROPOSIZIONE 2.9.4: 1) ${}^t f: W^* \rightarrow V^*$ è lineare

2) ${}^t({}^t f) = f$ (grazie all'identificazione degli isomorfismi canonici ψ_V e ψ_W), cioè è commutativo il diagramma:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \psi_V \downarrow & & \downarrow \psi_W \\ V^{**} & \xrightarrow{{}^t({}^t f)} & W^{**} \end{array}$$

3) Se $h: W \rightarrow Z$ è lineare, allora ${}^t(h \circ f) = {}^t f \circ {}^t h$

4) $\text{Ker}({}^t f) = \text{Ann}(\text{Im}(f))$

5) $\text{Im}({}^t f) = \text{Ann}(\text{Ker}(f))$

6) Se \mathcal{B} è base di V e \mathcal{S} è base di W , $\mathfrak{M}_{\mathcal{S}^*, \mathcal{B}^*}({}^t f) = {}^t(\mathfrak{M}_{\mathcal{B}, \mathcal{S}}(f))$.

Dimostrazione:

1) $\forall a_1, a_2 \in \mathbb{K}, \forall g_1, g_2 \in W^*$:

$${}^t f(a_1 g_1 + a_2 g_2) = (a_1 g_1 + a_2 g_2) \circ f = a_1(g_1 \circ f) + a_2(g_2 \circ f) = a_1 {}^t f(g_1) + a_2 {}^t f(g_2).$$

2) Devo mostrare che $\psi_W \circ f = {}^t({}^t f) \circ \psi_V$, ossia che

$$\forall v \in V, (\psi_W \circ f)(v) = ({}^t({}^t f) \circ \psi_V)(v), \text{ ossia che}$$

$$\forall g \in W^*, (\psi_W \circ f)(v)(g) = ({}^t({}^t f) \circ \psi_V)(v)(g).$$

$$(\psi_W \circ f)(v)(g) = \psi_W(f(v))(g) = g(f(v)) = (g \circ f)(v);$$

$$\begin{aligned} ({}^t({}^t f) \circ \psi_V)(v)(g) &= {}^t({}^t f)(\psi_V(v))(g) = (\psi_V(v) \circ {}^t f)(g) = \psi_V(v)({}^t f(g)) = \\ &= \psi_V(v)(g \circ f) = (g \circ f)(v). \end{aligned}$$

3) $\forall g \in Z^*, {}^t(h \circ f)(g) = g \circ h \circ f = {}^t h(g) \circ f = {}^t f({}^t h(g)) = ({}^t f \circ {}^t h)(g)$.

4) \Leftrightarrow Sia $g \in \text{Ker}({}^t f)$, cioè ${}^t f(g) = g \circ f = 0 \Rightarrow \forall f(x) \in \text{Im}(f), g(f(x)) = 0$, quindi $g \in \text{Ann}(\text{Im}(f))$;

\Rightarrow Sia $g \in \text{Ann}(\text{Im}(f)) \Rightarrow \forall x \in V, g(f(x)) = 0$, cioè $({}^t f(g))(x) = 0 \forall x \in V$, quindi ${}^t f(g) = 0$, cioè $g \in \text{Ker}({}^t f)$.

5) Per la 4) so che $\text{Ker}({}^t({}^t f)) = \text{Ann}(\text{Im}({}^t f))$, cioè $\text{Ker}(f) = \text{Ann}(\text{Im}({}^t f))$, quindi, applicando l'annullatore, $\text{Ann}(\text{Ker}(f)) = \text{Ann}(\text{Ann}(\text{Im}({}^t f))) = \text{Im}({}^t f)$.

6) $\mathcal{B} = \{v_1, \dots, v_n\}, \mathcal{S} = \{w_1, \dots, w_p\}$. Sia $A = \mathfrak{M}_{\mathcal{B}, \mathcal{S}}(f), N = \mathfrak{M}_{\mathcal{S}^*, \mathcal{B}^*}({}^t f)$.

$$\text{Allora } N^j = [{}^t f(w_j^*)]_{\mathcal{B}^*} = \begin{pmatrix} (w_j^* \circ f)(v_1) \\ \vdots \\ (w_j^* \circ f)(v_n) \end{pmatrix}, \text{ dunque } [N]_{ij} = (w_j^* \circ f)(v_i) = w_j^*(f(v_i)).$$

Ora $[f(v_i)]_{\mathcal{S}} = A^i$, cioè $f(v_i) = [A]_{1i}w_1 + \dots + [A]_{pi}w_p$, da cui:

$$w_j^*(f(v_i)) = [A]_{ji}, \text{ ossia } [N]_{ij} = [A]_{ji}, \text{ da cui } N = {}^t A.$$

Osservazione: Ancora: $\text{rk}({}^t A) = \dim \text{Im}({}^t A) = \dim \text{Ann}(\text{Ker}(f)) = n - \dim \text{Ker}(f) = \text{rk}(A)$

3 ENDOMORFISMI

3.0 ALCUNE NOZIONI SULLE PERMUTAZIONI

Notazione: Denoteremo $J_n = \{1, \dots, n\}$ e $S_n = S(J_n)$ le permutazioni di J_n .

Notazione: Denoteremo con $\begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$ la permutazione σ tale che $i \rightarrow \sigma(i) \forall i$.

DEFINIZIONE 3.0.1: Definiamo **orbita** di i secondo σ la successione:

$$i \rightarrow \sigma(i) \rightarrow \dots \rightarrow \sigma^k(i) = i.$$

L'orbita si dice **banale** quando consiste di un solo elemento, cioè $\sigma(i) = i$.

DEFINIZIONE 3.0.2: $c \in S_n$ si dice **ciclo** se contiene una sola orbita non banale.

Due cicli si dicono **disgiunti** se non hanno elementi in comune.

Notazione: Denoteremo con $(n_1 \ \dots \ n_k)$ il ciclo tale che $n_i \rightarrow n_{i+1} \ \forall 1 \leq i < k$ e $n_k \rightarrow n_1$.

PROPOSIZIONE 3.0.1: Cicli disgiunti commutano.

Esempio: Se $\sigma = (1 \ 2 \ 4) \circ (3 \ 5)$ e $\tau = (3 \ 5) \circ (1 \ 2 \ 4)$, $\sigma, \tau \in S_5$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}, \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}.$$

PROPOSIZIONE 3.0.2: Ogni $\sigma \in S_n$ si scrive come composizione di cicli disgiunti, in modo unico a meno dell'ordine.

Esempio: $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 5 & 1 & 2 & 6 \end{pmatrix}$. La decompongo in cicli:

$1 \rightarrow 4 \rightarrow 1$, cioè ho il ciclo $(1 \ 4)$,

$2 \rightarrow 3 \rightarrow 5 \rightarrow 2$, cioè ho il ciclo $(2 \ 3 \ 5)$,

$6 \rightarrow 6$, cioè ho il ciclo banale.

Quindi $\sigma = (1 \ 4) \circ (2 \ 3 \ 5)$.

DEFINIZIONE 3.0.3: Se $c = (n_1 \ \dots \ n_k)$ è un ciclo, definiamo **lunghezza** di c $l(c) = k$.

Per convenzione $l(id) = 1$.

DEFINIZIONE 3.0.4: Definiamo **trasposizione** un ciclo di lunghezza 2: $\tau = (n_1 \ n_2)$.

DEFINIZIONE 3.0.5: Se $\sigma = c_1 \circ \dots \circ c_p$, con c_1, \dots, c_p cicli disgiunti:

1) poniamo $N(\sigma) = (l(c_1) - 1) + \dots + (l(c_p) - 1)$

2) diciamo che σ è **pari (dispari)** se $N(\sigma)$ è pari (dispari)

3) definiamo **segno** di σ $sgn(\sigma) = (-1)^{N(\sigma)}$.

Osservazioni: Tutte le trasposizioni sono permutazioni dispari.

PROPOSIZIONE 3.0.3: Ogni ciclo c di lunghezza k si può scrivere come composizione di $N(c) = k - 1$ trasposizioni (non disgiunte).

Dimostrazione:

Se $c = (n_1 \dots n_k)$, non è difficile verificare che $c = (n_1 \ n_k) \circ \dots \circ (n_1 \ n_2)$.

Osservazione: La decomposizione di un ciclo nel prodotto di trasposizioni non è unica, ad esempio $(1 \ 2 \ 3) = (1 \ 3)(1 \ 2) = (1 \ 2)(1 \ 3)(2 \ 3)(1 \ 2)$.

Si può però dimostrare il seguente fatto:

PROPOSIZIONE 3.0.4: La parità del numero di trasposizioni che compongono un ciclo è costante.

PROPOSIZIONE 3.0.5: Sia $\sigma \in S_n$. Allora $N(\sigma) = N(\sigma^{-1})$.

Dimostrazione:

Sia $\sigma = c_1 \circ \dots \circ c_p$ la decomposizione di σ in cicli disgiunti.

Allora $\sigma^{-1} = c_p^{-1} \circ \dots \circ c_1^{-1}$.

Inoltre $l(c_i) = l(c_i^{-1})$, poiché $(n_1 \dots n_k)^{-1} = (n_k \dots n_1)$.

Perciò $\forall i, N(c_i) = N(c_i^{-1}) \Rightarrow N(\sigma) = N(\sigma^{-1})$.

Osservazione: Poiché $\sigma \circ \sigma^{-1} = id$, componendo σ con $N(\sigma)$ trasposizioni si ottiene l'identità.

3.1 DETERMINANTE

DEFINIZIONE 3.1.1: Sia $A \in \mathcal{M}(n, \mathbb{K})$. Definiamo **determinante** una funzione $D: \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$ tale che $D(A) = 0 \Leftrightarrow$ le righe di A sono linearmente dipendenti ($\Leftrightarrow rk(A) < n$).

Osservazione: Cerchiamo una tale D_2 nello spazio $\mathcal{M}(2, \mathbb{K})$.

Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

- Se $a = 0 \wedge c = 0$ le righe sono dipendenti;
- Se $a = 0 \wedge c \neq 0$, allora le righe sono dipendenti $\Leftrightarrow b = 0$.
- Se $a \neq 0$, riduco a scalini:

$$A' = \begin{pmatrix} a & b \\ 0 & d - bca^{-1} \end{pmatrix},$$

dunque le righe di A sono dipendenti $\Leftrightarrow d - bca^{-1} = 0 \Leftrightarrow ad - bc = 0$.

Riassumendo, se pongo $D_2(A) = ad - bc$, ho l'applicazione voluta, tale che $D(A) = 0 \Leftrightarrow$ le righe di A sono linearmente dipendenti.

DEFINIZIONE 3.1.2: Siano V, W \mathbb{K} -spazi vettoriali. Sia $f: \underbrace{V \times \dots \times V}_{n \text{ volte}} \rightarrow W$.

$\forall i = 1, \dots, n$ fisso $w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_n \in V$ e sia $f_i = f(w_1, \dots, w_{i-1}, v, w_{i+1}, \dots, w_n): V \rightarrow W$.

L'applicazione f si dice **multilineare** se f_i è lineare $\forall i$.

Osservazione: L'applicazione determinante che stiamo cercando deve essere multilineare, poiché deve essere lineare in ogni riga.

PROPOSIZIONE 3.1.1: La funzione $D_2: \mathcal{M}(2, \mathbb{K}) \rightarrow \mathbb{K} \mid D_2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ verifica le seguenti proprietà:

- 1) D_2 è lineare in ogni riga;
- 2) Se A ha due righe uguali, $D_2(A) = 0$;
- 3) $D_2(I) = 1$.

Dimostrazione:

- 1) Verifichiamolo solo per la prima riga; poniamo $B = (a_1 \ b_1)$, $C = (a_2 \ b_2)$. Allora:

$$D_2 \begin{pmatrix} \lambda B + \mu C \\ A_2 \end{pmatrix} = D_2 \begin{pmatrix} \lambda a_1 + \mu a_2 & \lambda b_1 + \mu b_2 \\ c & d \end{pmatrix} = (\lambda a_1 + \mu a_2)d - (\lambda b_1 + \mu b_2)c = \\ \lambda(a_1d - b_1c) + \mu(a_2d - b_2c) = \lambda D_2 \begin{pmatrix} B \\ A_2 \end{pmatrix} + \mu D_2 \begin{pmatrix} C \\ A_2 \end{pmatrix}, \text{ da cui la tesi.}$$

- 2) Ovvio.
- 3) $1 \cdot 1 - 0 \cdot 0 = 1$.

PROPOSIZIONE 3.1.2: Se D verifica le proprietà 1), 2), 3), allora verifica anche le seguenti:

- a) Se A ha una riga nulla, $D(A) = 0$;
- b) $D(\dots, A_i, \dots, A_j, \dots) = -D(\dots, A_j, \dots, A_i, \dots)$;
- c) Se B è ottenuta da A sommando ad una riga una combinazione lineare delle altre righe (operazione elementare di 3° tipo), allora $D(B) = D(A)$;
- d) Se le righe di A sono linearmente dipendenti, $D(A) = 0$;
- e) Se $A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$, allora $D(A) = a_1 \cdot \dots \cdot a_n$.

Dimostrazione:

- a) Se $A_i = 0$, allora $A = (A_1 \mid \dots \mid 0 \cdot B \mid \dots \mid A_n)$.
Dunque $D(A) = 0 \cdot D(A_1 \mid \dots \mid B \mid \dots \mid A_n) = 0$.
- b) Considero la matrice $(\dots, A_i + A_j, \dots, A_i + A_j, \dots)$.
Per la proprietà 2), $D(\dots, A_i + A_j, \dots, A_i + A_j, \dots) = 0$.
Per multilinearità:
 $0 = D(\dots, A_i + A_j, \dots, A_i + A_j, \dots) = D(\dots, A_i, \dots, A_i, \dots) + D(\dots, A_i, \dots, A_j, \dots) + \\ D(\dots, A_j, \dots, A_i, \dots) + D(\dots, A_j, \dots, A_j, \dots) = 0 + D(\dots, A_i, \dots, A_j, \dots) + D(\dots, A_j, \dots, A_i, \dots) + 0,$
da cui la tesi.
- c) Supponiamo $B = A_1 + \sum_{i=2}^n \alpha_i A_i$. Allora:
 $D(B) = D(A_1 + \sum_{i=2}^n \alpha_i A_i, A_2, \dots, A_n) = D(A) + \sum_{i=2}^n \alpha_i D(A_i, A_2, \dots, A_n)$.
Ma la sommatoria è nulla, in quanto per la proprietà 2) tutti i termini sono nulli, dunque $D(B) = D(A)$.
- d) Supponiamo $A_1 = \sum_{i=2}^n \alpha_i A_i$. Allora:
 $D(A) = D(\sum_{i=2}^n \alpha_i A_i, A_2, \dots, A_n) = \sum_{i=2}^n \alpha_i D(A_i, A_2, \dots, A_n) = 0$.
- e) Se A è diagonale, allora $A_i = a_i I_i \ \forall i$. Perciò:
 $D(A) = a_1 D(I_1, a_2 I_2, \dots, a_n I_n) = \dots = a_1 \cdot \dots \cdot a_n \cdot D(I) = a_1 \cdot \dots \cdot a_n$.

Nella seguente esposizione riguardo al determinante dimostreremo prima che, se la funzione determinante esiste, allora è unica, e solo dopo ne mostreremo l'esistenza.

PROPOSIZIONE 3.1.3: Se D verifica 1), 2) e 3), allora è unico.

Dimostrazione:

Sia S a scalini ottenuta da A con m operazioni di 1° tipo e k di 3° tipo.

Allora $D(A) = (-1)^m D(S)$.

Se S ha una riga nulla $\Rightarrow D(S) = 0 \Rightarrow D(A) = 0$.

Altrimenti con solo operazioni di 3° tipo portiamo S nella forma $S' = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$.

Dunque $D(S) = D(S') = a_1 \cdot \dots \cdot a_n \Rightarrow D(A) = (-1)^m a_1 \cdot \dots \cdot a_n$, perciò in ogni caso è unico.

COROLLARIO 3.1.4: Se D è una funzione che verifica 1), 2) e 3), allora:

$D(A) = 0 \Leftrightarrow$ le righe di A sono linearmente dipendenti.

Dimostrazione:

\Leftarrow Già fatta.

\Rightarrow Se le righe di A fossero indipendenti, allora $D(A) = (-1)^m a_1 \cdot \dots \cdot a_n \neq 0$, assurdo.

PROPOSIZIONE 3.1.5: Se $D: \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$ verifica 1), 2) e 3), allora:

$$D(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)}$$

dove $a_{i,j} = [A]_{ij}$.

Dimostrazione:

$$D(A) = D \left(\begin{array}{c} \sum_{i_1=1}^n a_{1,i_1} \cdot I_{i_1} \\ \hline A_2 \\ \hline \vdots \\ \hline A_n \end{array} \right) = \sum_{i_1=1}^n a_{1,i_1} D \left(\begin{array}{c} I_{i_1} \\ \hline \sum_{i_2=1}^n a_{2,i_2} \cdot I_{i_2} \\ \hline A_3 \\ \hline \vdots \\ \hline A_n \end{array} \right) = \dots = \sum_{\substack{i_1 \in J_n \\ \vdots \\ i_n \in J_n}} a_{1,i_1} \cdot \dots \cdot a_{n,i_n} D \left(\begin{array}{c} I_{i_1} \\ \hline \vdots \\ \hline I_{i_n} \end{array} \right)$$

Ma se fra gli i_j ce ne sono due uguali, allora $D \left(\begin{array}{c} I_{i_1} \\ \hline \vdots \\ \hline I_{i_n} \end{array} \right) = 0$, poiché ha due righe uguali.

Perciò:

$$D(A) = \sum_{\{i_1, \dots, i_n\} \in J_n^n} a_{1,i_1} \cdot \dots \cdot a_{n,i_n} D \left(\begin{array}{c} I_{i_1} \\ \hline \vdots \\ \hline I_{i_n} \end{array} \right) = \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} D \left(\begin{array}{c} I_{\sigma(1)} \\ \hline \vdots \\ \hline I_{\sigma(n)} \end{array} \right)$$

Si riporta la matrice $\begin{pmatrix} I_{\sigma(1)} \\ \hline \vdots \\ \hline I_{\sigma(n)} \end{pmatrix}$ a I con $N(\sigma)$ scambi di righe, per cui:

$$D(A) = \sum_{\sigma \in S_n} \underbrace{(-1)^{N(\sigma)}}_{\text{sgn}(\sigma)} \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)} \underbrace{D(I)}_{=1} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{n,\sigma(n)}$$

Osservazione: La precedente proposizione è una dimostrazione alternativa dell'unicità di D .

Esempi: Applichiamo la formula trovata ai casi più semplici (è molto laborioso applicarla alle matrici di ordine > 3):

- $n = 2$: $S_2 = \left\{ id, \underbrace{(1 \ 2)}_{\sigma} \right\}$ e $sgn(id) = 1$, $sgn(\sigma) = -1$.

$D \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = 1 \cdot a_{1,id(1)} \cdot a_{2,id(2)} + (-1) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} = a_{11}a_{22} - a_{12}a_{21}$, che coincide con la formula che già avevamo.

- $n = 3$: $S_3 = \left\{ \underbrace{id}_{sgn=1}, \underbrace{(1 \ 2), (1 \ 3), (2 \ 3)}_{sgn=-1}, \underbrace{(1 \ 2 \ 3), (1 \ 3 \ 2)}_{sgn=1} \right\}$;

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

Osservazione: La precedente espressione per il determinante nel caso $n = 3$ è detta **regola (o formula) di Sarrus**.

DEFINIZIONE 3.1.3: Definiamo $D_n: \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$ la funzione definita da:

se $n = 1$, $D_1(a) = a$;

se $n = 2$, $D_2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$;

se $n > 2$, $D_n(A) = \sum_{i=1}^n (-1)^{i+1} \cdot [A]_{i1} \cdot D_{n-1}(A_{i1})$, dove A_{ij} è la sottomatrice di A di ordine $n - 1$ ottenuta da A cancellando la riga A_i e la colonna A^j .

Questa sottomatrice prende il nome di **complemento algebrico** dell'elemento a_{ij} .

L'applicazione D_n appena definita si chiama **sviluppo di Laplace secondo la prima colonna**.

Osservazione: L'applicazione D_n è stata definita ricorsivamente.

PROPOSIZIONE 3.1.6: L'applicazione D_n verifica le proprietà 1), 2) e 3).

Dimostrazione:

Procediamo in ogni caso con l'induzione su n ; in tutti e tre i casi abbiamo già provato il passo base, dunque qua mostriamo solo il passo induttivo.

$$1) \text{ Siano } A = \begin{pmatrix} \overline{A_1} \\ \vdots \\ \overline{\lambda B + \mu C} \\ \vdots \\ \overline{A_n} \end{pmatrix} \leftarrow A_j, \quad A' = \begin{pmatrix} \overline{A_1} \\ \vdots \\ \overline{B} \\ \vdots \\ \overline{A_n} \end{pmatrix} \leftarrow A'_j, \quad A'' = \begin{pmatrix} \overline{A_1} \\ \vdots \\ \overline{C} \\ \vdots \\ \overline{A_n} \end{pmatrix} \leftarrow A''_j.$$

Dobbiamo dimostrare che $D_n(A) = \lambda D_n(A') + \mu D_n(A'')$.

Osserviamo che:

$$\begin{cases} [A]_{i1} = [A']_{i1} = [A'']_{i1} \quad \forall i \neq j \\ [A]_{j1} = \lambda [A']_{j1} + \mu [A'']_{j1} \end{cases}, \text{ inoltre } A_{j1} = A'_{j1} = A''_{j1},$$

mentre se $i \neq j$ il minore A_{i1} ha una riga che è combinazione lineare di due righe dei minori A'_{i1} e A''_{i1} .

Dunque per ipotesi induttiva:

$$\forall i \neq j, D_{n-1}(A_{i1}) = \lambda D_{n-1}(A'_{i1}) + \mu D_{n-1}(A''_{i1}).$$

Allora:

$$\begin{aligned} D_n(A) &= \sum_{i=1}^n (-1)^{i+1} \cdot [A]_{i1} \cdot D_{n-1}(A_{i1}) = \\ &= \sum_{i \neq j} \left((-1)^{i+1} \cdot [A]_{i1} \cdot D_{n-1}(A_{i1}) \right) + (-1)^{j+1} \cdot [A]_{j1} \cdot D_{n-1}(A_{j1}) = \\ &= \sum_{i \neq j} \left((-1)^{i+1} \cdot [A]_{i1} \cdot (\lambda D_{n-1}(A'_{i1}) + \mu D_{n-1}(A''_{i1})) \right) + (-1)^{j+1} \cdot \\ &\quad \cdot (\lambda [A']_{j1} + \mu [A'']_{j1}) \cdot D_{n-1}(A_{j1}) = \\ &= \lambda \sum_{i=1}^n \left((-1)^{i+1} \cdot [A']_{i1} \cdot D_{n-1}(A'_{i1}) \right) + \mu \sum_{i=1}^n \left((-1)^{i+1} \cdot [A'']_{i1} \cdot D_{n-1}(A''_{i1}) \right) \\ &= \lambda D_n(A') + \mu D_n(A''). \end{aligned}$$

2) Supponiamo che A abbia due righe uguali, ad esempio $A_j = A_h, j < h$.

Se $i \neq j$ e $i \neq h$, anche il minore A_{i1} ha due righe uguali e quindi, per ipotesi induttiva, $D_{n-1}(A_{i1}) = 0$. Dunque:

$$D_n(A) = (-1)^{j+1} [A]_{j1} D_{n-1}(A_{j1}) + (-1)^{h+1} [A]_{h1} D_{n-1}(A_{h1})$$

Poiché $A_j = A_h$, si ha che $[A]_{j1} = [A]_{h1}$.

Inoltre i minori A_{j1} e A_{h1} contengono le stesse righe ma in posizioni diverse.

Più precisamente, se A'_m denota la riga A_m privata del primo elemento, si ha:

$$A_{j1} = \begin{pmatrix} \vdots \\ \hline A'_{j-1} \\ \hline A'_{j+1} \\ \hline \vdots \\ \hline A'_h = A'_j \\ \hline A'_{h+1} \\ \hline \vdots \end{pmatrix}; \quad A_{h1} = \begin{pmatrix} \vdots \\ \hline A'_{j-1} \\ \hline A'_j = A'_h \\ \hline \vdots \\ \hline A'_{h-1} \\ \hline A'_{h+1} \\ \hline \vdots \end{pmatrix}.$$

Allora A_{j1} può essere trasformato in A_{h1} attraverso $h - 1 - j$ scambi di righe, per cui $D_{n-1}(A_{h1}) = (-1)^{h-1-j} D_{n-1}(A_{j1})$.

Dunque:

$$\begin{aligned} D_n(A) &= (-1)^{j+1} [A]_{j1} D_{n-1}(A_{j1}) + (-1)^{h+1} [A]_{h1} (-1)^{h-1-j} D_{n-1}(A_{j1}) = \\ &= [A]_{j1} D_{n-1}(A_{j1}) \cdot \left((-1)^{j+1} + (-1)^{2h-j} \right) \end{aligned}$$

Ma $j + 1 + 2h - j = 2h + 1$ dispari, allora $\left((-1)^{j+1} + (-1)^{2h-j} \right) = 0$, da cui la tesi.

3) L'unico contributo allo sviluppo di Laplace è dato da $[A]_{11} = 1$, il cui complemento algebrico è I_{n-1} .

Per ipotesi induttiva $D_{n-1}(I_{n-1}) = 1$ e dunque $D_n(I_n) = 1$.

Osservazione: Con la precedente dimostrazione abbiamo dimostrato l'effettiva esistenza della funzione determinante.

DEFINIZIONE 3.1.4: Chiamiamo determinante l'unica funzione $\det : \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$ che verifica 1), 2) e 3).

Osservazione: Abbiamo visto che $\det(A) = 0 \Leftrightarrow$ le righe di A sono dipendenti $\Leftrightarrow A$ è singolare. Inoltre A è invertibile $\Leftrightarrow \det(A) \neq 0$.

Osservazione: Con la stessa dimostrazione si prova che lo sviluppo di Laplace secondo una colonna A^j :

$$D_n(A) = \sum_{i=1}^n (-1)^{i+j} [A]_{ij} D_{n-1}(A_{ij})$$

verifica 1), 2) e 3) e dunque coincide con $\det(A)$ (per l'unicità).

Osservazione: Dalla definizione, $\det(\lambda A) = \lambda^n \det(A)$.

TEOREMA DI BINET: $\forall A, B \in \mathcal{M}(n, \mathbb{K}), \det(AB) = \det(A) \cdot \det(B)$.

Dimostrazione:

Se $\det(B) = 0 \Rightarrow \text{rk}(B) < n$ e dunque $\text{rk}(AB) < n \Rightarrow \det(AB) = 0$.

Se $\det(B) \neq 0$, si consideri $f: \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$ definita da $f(A) = \frac{\det(AB)}{\det(B)}$;

f verifica le proprietà 1), 2) e 3), infatti:

1) Se una riga di A è combinazione lineare di due righe, allora lo stesso vale per AB .

Ne segue che f è lineare nelle righe (poiché il denominatore $\det(B)$ è una costante).

2) Se A ha due righe uguali, allora ce le ha anche AB (poiché ancora B è isomorfismo); dunque $\det(AB) = 0 \Rightarrow f(A) = 0$.

3) $f(I) = \frac{\det(IB)}{\det(B)} = 1$.

Per l'unicità della funzione \det , $f(A) = \det(A)$ e dunque $\det(A) = \frac{\det(AB)}{\det(B)}$, tesi.

COROLLARIO 3.1.7: Se A è invertibile, allora $\det(A^{-1}) = (\det(A))^{-1}$.

Dimostrazione:

$1 = \det(I) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1}) \Rightarrow \det(A^{-1}) = (\det(A))^{-1}$.

PROPOSIZIONE 3.1.8: $\forall A \in \mathcal{M}(n, \mathbb{K}), \det({}^t A) = \det(A)$.

Dimostrazione:

$$\det({}^t A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot [{}^t A]_{1,\sigma(1)} \cdot \dots \cdot [{}^t A]_{n,\sigma(n)} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot [A]_{\sigma(1),1} \cdot \dots \cdot [A]_{\sigma(n),n}$$

Sia $\tau = \sigma^{-1}$. Se $\sigma(i) = j \Rightarrow \tau(j) = i$, dunque $[A]_{\sigma(i),i} = [A]_{j,\tau(j)}$.

Riordinando il prodotto e poiché $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$:

$$\det({}^t A) = \sum_{\tau \in S_n} \text{sgn}(\tau) \cdot [A]_{1,\tau(1)} \cdot \dots \cdot [A]_{n,\tau(n)} = \det(A)$$

COROLLARIO 3.1.9: $\det(A)$ può essere calcolato mediante sviluppo di Laplace secondo una qualsiasi riga.

Dimostrazione:

$$\begin{aligned} \sum_{i=1}^n (-1)^{i+j} [A]_{ji} \det(A_{ji}) &= \sum_{i=1}^n (-1)^{i+j} [{}^t A]_{ij} \det({}^t A_{ji}) = \sum_{i=1}^n (-1)^{i+j} [{}^t A]_{ij} \det(({}^t A)_{ij}) \equiv \\ &= \det({}^t A) = \det(A), \end{aligned}$$

dove il passaggio \equiv deriva dal fatto che quella precedente è esattamente lo sviluppo di Laplace della j -esima colonna di ${}^t A$.

Osservazione: Poiché $\det({}^t A) = \det(A)$, la funzione determinante verifica le proprietà 1), 2) e 3) anche per le colonne.

REGOLA DI CRAMER: Sia $AX = B$ un sistema lineare quadrato con n equazioni e n incognite, $\det(A) \neq 0$. Allora la sua unica soluzione è $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, dove:

$$y_i = \frac{\det(B(i))}{\det(A)}, \text{ dove } B(i) = \left(A^1 \mid \dots \mid \underbrace{B}_{i\text{-esima colonna}} \mid \dots \mid A^n \right).$$

Dimostrazione:

Poiché Y è soluzione di $AX = B$, allora $AY = y_1 A^1 + \dots + y_n A^n = B$. Allora:

$$\det(B(i)) = \det \left(A^1 \mid \dots \mid \sum_{j=1}^n y_j A^j \mid \dots \mid A^n \right) = \sum_{j=1}^n y_j \det(A^1 \mid \dots \mid A^j \mid \dots \mid A^n) = y_i \det(A)$$

poiché se $j \neq i$, allora $(A^1 \mid \dots \mid A^j \mid \dots \mid A^n)$ ha due colonne uguali e dunque $\det(A^1 \mid \dots \mid A^j \mid \dots \mid A^n) = 0$.

Quindi, visto che $\det(A) \neq 0$, si ha la tesi.

CALCOLO DELL'INVERSA: Se A è invertibile, allora la matrice B definita da

$$[B]_{ij} = (-1)^{i+j} \frac{\det(A_{ji})}{\det(A)}$$

è l'inversa di A .

Dimostrazione:

Poiché A è invertibile, mi basta mostrare che $AB = I$, poiché se B è inversa destra allora è anche inversa sinistra.

$$[AB]_{hk} = \sum_{i=1}^n [A]_{hi} [B]_{ik} = \sum_{i=1}^n (-1)^{i+k} [A]_{hi} \frac{\det(A_{ki})}{\det(A)}$$

$$\text{Se } h = k \Rightarrow [AB]_{hh} = \sum_{i=1}^n (-1)^{i+h} [A]_{hi} \frac{\det(A_{hi})}{\det(A)} = \frac{\det(A)}{\det(A)} = 1;$$

se $h \neq k \Rightarrow [AB]_{hk} = 0$, poiché il numeratore è lo sviluppo secondo la riga k -esima di una matrice ottenuta da A sostituendo ad A_k la riga A_h e che quindi ha due righe uguali \Rightarrow tesi.

Osservazione: La formula dell'inversa implica la regola di Cramer.

Infatti sia dato il sistema lineare $n \times n$ $AX = B$.

L'unica soluzione è $Y = A^{-1}B$.

$$\text{Se } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow y_i = [A^{-1}B]_{i1} = \sum_{j=1}^n [A^{-1}]_{ij} [B]_{j1} = \sum_{j=1}^n (-1)^{i+j} \frac{\det(A_{ji})}{\det(A)} [B]_{j1} = \frac{\det(B(i))}{\det(A)}$$

L'ultimo passaggio deriva dal fatto che quello sulla sinistra è lo sviluppo secondo la prima colonna di una matrice ottenuta da A sostituendo la colonna B alla colonna A^i .

Osservazione: Siano A, C matrici quadrate. Allora, sfruttando il prodotto a blocchi, possiamo vedere che:

$$\begin{pmatrix} A & | & B \\ 0 & | & C \end{pmatrix} = \begin{pmatrix} I & | & 0 \\ 0 & | & C \end{pmatrix} \cdot \begin{pmatrix} A & | & B \\ 0 & | & I \end{pmatrix}$$

Dunque:

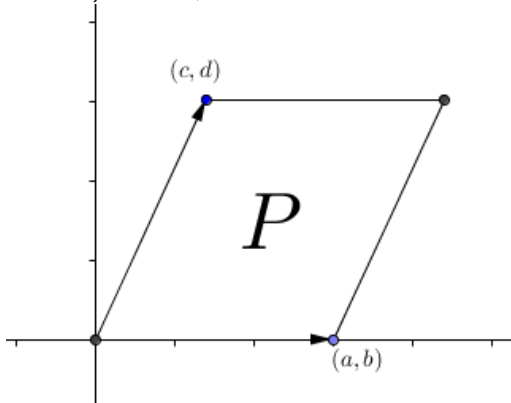
$$\det \begin{pmatrix} A & | & B \\ 0 & | & C \end{pmatrix} = \det \begin{pmatrix} I & | & 0 \\ 0 & | & C \end{pmatrix} \cdot \det \begin{pmatrix} A & | & B \\ 0 & | & I \end{pmatrix} = \det(C) \cdot \det(A)$$

poiché eseguendo lo sviluppo di Laplace secondo la prima riga nella matrice $\begin{pmatrix} I & | & 0 \\ 0 & | & C \end{pmatrix}$, si ottiene $1 \cdot \dots \cdot 1 \cdot \det(C) = \det(C)$. Analogamente per l'altra matrice.

Osservazione: Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$. Sia P il parallelogramma con lati (a, b) e (c, d) .

Posso supporre $a > 0$.

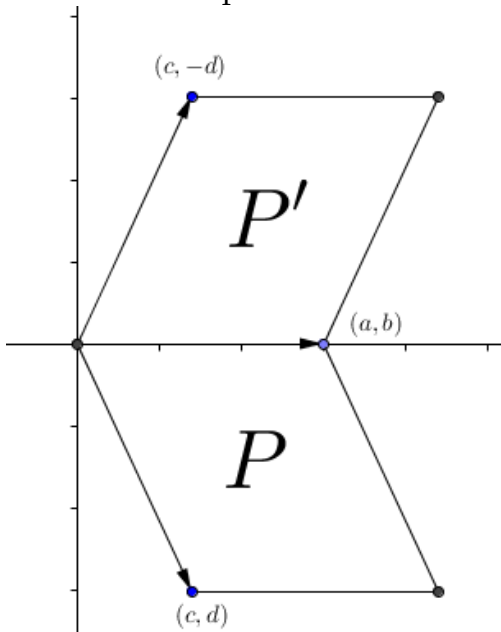
- Caso 1): $b = 0, d > 0$.



$$\text{Area}(P) = ad = \det \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Caso 2): $b = 0, d < 0$.

Ribaltiamo P rispetto all'asse x ottenendo P' :

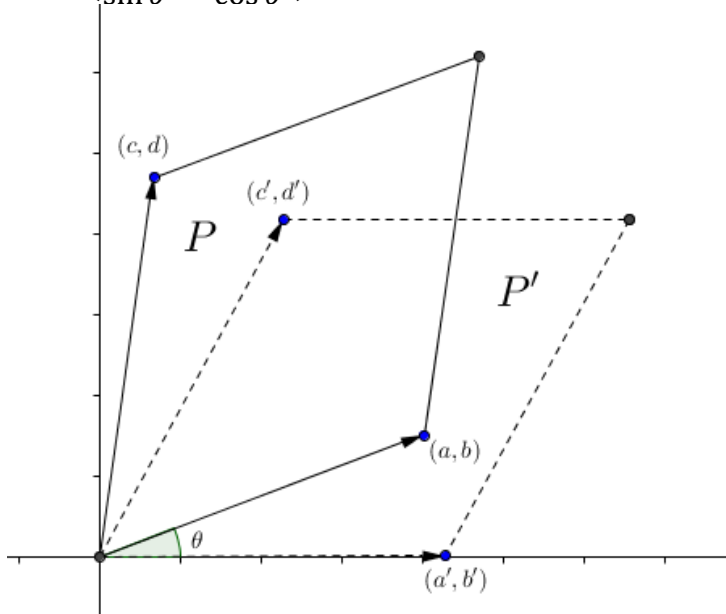


$$\text{Area}(P) = \text{Area}(P') = \det \begin{pmatrix} a & 0 \\ c & -d \end{pmatrix} = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dunque se $b = 0 \Rightarrow Area(P) = \left| \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right|$.

- Caso generale): Sia R_θ la rotazione antioraria degli assi di angolo θ .

$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$; sia P' il parallelogramma P nei nuovi assi.



Sicuramente $Area(P) = Area(P')$.

P' ha lati $(a', 0), (c', d')$, perciò:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a' \\ b' \end{pmatrix} \Rightarrow \begin{cases} a = a' \cos \theta \\ b = a' \sin \theta \end{cases}$$

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} c' \\ d' \end{pmatrix} \Rightarrow \begin{cases} c = c' \cos \theta - d' \sin \theta \\ d = c' \sin \theta + d' \cos \theta \end{cases}$$

Vediamo che:

$$\begin{aligned} \left| \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| &= |ad - bc| = |a' \cos \theta (c' \sin \theta + d' \cos \theta) - a' \sin \theta (c' \cos \theta - d' \sin \theta)| \\ &= |a' c' \cos \theta \sin \theta + a' d' (\cos \theta)^2 - a' c' \sin \theta \cos \theta + a' d' (\sin \theta)^2| = |a' d'| \\ &= \left| \det \begin{pmatrix} a' & 0 \\ c' & d' \end{pmatrix} \right| = Area(P') = Area(P) \end{aligned}$$

Dunque in generale $|\det(A)| = Area(P)$, dove $A \in \mathcal{M}(2, \mathbb{R})$ e P è il parallelogramma con lati i vettori riga di A .

DETERMINANTE DI VANDERMONDE: Siano $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ e sia:

$$A(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{pmatrix}$$

Allora $\det(A(\lambda_1, \dots, \lambda_n)) = \prod_{i < j} (\lambda_j - \lambda_i)$ (in particolare è $\neq 0 \Leftrightarrow \lambda_i \neq \lambda_j \forall i \neq j$).

Dimostrazione 1:

Per induzione su n :

Passo base): $n = 2$: $\det \begin{pmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{pmatrix} = \lambda_2 - \lambda_1$, verificato.

Passo induttivo): Consideriamo $A(\lambda_1, \dots, \lambda_{n+1})$ e $\forall 2 \leq i \leq n+1$ tolgo $\lambda_1 c_{i-1}$ a c_i .

Ottengo:

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \lambda_2 - \lambda_1 & \lambda_2^2 - \lambda_1 \lambda_2 & \dots & \lambda_2^n - \lambda_1 \lambda_2^{n-1} \\ 1 & \lambda_3 - \lambda_1 & \lambda_3^2 - \lambda_1 \lambda_3 & \dots & \lambda_3^n - \lambda_1 \lambda_3^{n-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_{n+1} - \lambda_1 & \lambda_{n+1}^2 - \lambda_1 \lambda_{n+1} & \dots & \lambda_{n+1}^n - \lambda_1 \lambda_{n+1}^{n-1} \end{pmatrix}$$

Dunque:

$$\begin{aligned} A(\lambda_1, \dots, \lambda_{n+1}) &= \det \begin{pmatrix} \lambda_2 - \lambda_1 & \lambda_2(\lambda_2 - \lambda_1) & \dots & \lambda_2^{n-1}(\lambda_2 - \lambda_1) \\ \vdots & \vdots & \dots & \vdots \\ \lambda_{n+1} - \lambda_1 & \lambda_{n+1}(\lambda_{n+1} - \lambda_1) & \dots & \lambda_{n+1}^{n-1}(\lambda_{n+1} - \lambda_1) \end{pmatrix} \\ &= (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) \dots (\lambda_{n+1} - \lambda_1) \det A(\lambda_2, \dots, \lambda_{n+1}) = \prod_{i < j} (\lambda_j - \lambda_i) \end{aligned}$$

Dimostrazione 2:

$$\text{Sia } p(x) = \det \begin{pmatrix} 1 & x & x^2 & \dots & x^n \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_{n+1} & \lambda_{n+1}^2 & \dots & \lambda_{n+1}^n \end{pmatrix}.$$

$p(x)$ è un polinomio in x di grado al più n (per sviluppo lungo la prima riga);

$\forall 2 \leq i \leq n, p(\lambda_i) = 0$ (poiché la matrice avrebbe due righe uguali).

Per Ruffini $(\lambda_i - x) | p(x) \forall i \geq 2$.

Supponiamo i λ_i tutti diversi (altrimenti la tesi è banale); allora:

$$(\lambda_2 - x) \cdot \dots \cdot (\lambda_{n+1} - x) | p(x).$$

Confrontando i due gradi, deduco che $p(x) = k \cdot (\lambda_2 - x) \cdot \dots \cdot (\lambda_{n+1} - x)$, $k \in \mathbb{K}$.

Ma:

$$p(0) = \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_{n+1} & \lambda_{n+1}^2 & \dots & \lambda_{n+1}^n \end{pmatrix} = \lambda_2 \cdot \dots \cdot \lambda_{n+1} \cdot \prod_{\substack{i < j \\ i \geq 2}} (\lambda_j - \lambda_i)$$

Perciò $k = \prod_{\substack{i < j \\ i \geq 2}} (\lambda_j - \lambda_i) \Rightarrow p(\lambda_1) = k \cdot (\lambda_2 - \lambda_1) \cdot \dots \cdot (\lambda_{n+1} - \lambda_1) \Rightarrow$ tesi.

3.2 ENDOMORFISMI SIMILI

Notazione: Indicheremo $\mathfrak{M}_{B,B}(f)$ come $\mathfrak{M}_B(f)$.

Riprendiamo per un attimo il concetto di SD-equivalenza. Sappiamo che, dato $f \in \text{End}(V)$,

$$\mathfrak{M}_{B,B'}(f) \equiv_{SD} \mathfrak{M}_{S,S'}(f).$$

Supponiamo che $B = B'$ e $S = S'$; allora:

$$\mathfrak{M}_B(f) \equiv_{SD} \mathfrak{M}_S(f) \Leftrightarrow \exists M, N \in GL(V) \mid \mathfrak{M}_S(f) = N \cdot \mathfrak{M}_B(f) \cdot M.$$

In particolare, $M = \mathfrak{M}_{S,B}(id)$ e $N = \mathfrak{M}_{B,S}(id)$.

Dunque $N = M^{-1}$. Questo può essere riassunto nel seguente schema:

$$\begin{array}{ccccccc} V_S & \xrightarrow{id} & V_B & \xrightarrow{f} & V_B & \xrightarrow{id} & V_S \\ []_S \downarrow & & []_B \downarrow & & []_B \downarrow & & \downarrow []_S \\ \mathbb{K}^n & \xrightarrow{M} & \mathbb{K}^n & \xrightarrow{A} & \mathbb{K}^n & \xrightarrow{M^{-1}} & \mathbb{K}^n \end{array}$$

dove $A = \mathfrak{M}_{\mathcal{B}}(f)$.

Dunque $\mathfrak{M}_{\mathcal{S}}(f) = M^{-1} \cdot A \cdot M$.

DEFINIZIONE 3.2.1: $f, g \in \text{End}(V)$ si dicono **coniugati** ($f \sim g$) se $\exists h \in GL(V) \mid g = h^{-1} \circ f \circ h$

DEFINIZIONE 3.2.2: $A, B \in \mathcal{M}(n, \mathbb{K})$ si dicono **simili** ($A \sim B$) se $\exists M \in GL(n, \mathbb{K}) \mid B = M^{-1}AM$.

Osservazioni: 1) Coniugio e similitudine sono relazioni di equivalenza (le verifiche sono lasciate al lettore);

2) $f, g \in \text{End}(V)$. Sono fatti equivalenti:

a) $f \sim g$;

b) $\forall \mathcal{B}$ base di V , $\mathfrak{M}_{\mathcal{B}}(f) \sim \mathfrak{M}_{\mathcal{B}}(g)$;

c) $\exists \mathcal{B}, \mathcal{S}$ basi di V , $\mathfrak{M}_{\mathcal{B}}(f) = \mathfrak{M}_{\mathcal{S}}(g)$.

Questo fatto può essere dimostrato ricalcando l'analoga dimostrazione per endomorfismi SD-equivalenti.

3) $f \sim g \Rightarrow f \equiv_{SD} g$.

Dunque il rango è un invariante di coniugio, ma non è un sistema completo di invarianti.

Infatti $rk(I) = rk \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = 2$, ma $\nexists M \in GL(2) \mid M^{-1}IM = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, poiché $\forall M \in GL(2)$, $M^{-1}IM = I$.

PROPOSIZIONE 3.2.1: $A, B \in \mathcal{M}(n, \mathbb{K})$. $A \sim B \Rightarrow \det(A) = \det(B)$.

Dimostrazione:

$B = M^{-1}AM \Rightarrow \det(B) = \det(M^{-1}AM) = \det(M^{-1}) \cdot \det(A) \cdot \det(M) = \det(A)$, poiché per Binet $\det(M^{-1}) = (\det(M))^{-1}$.

Osservazione: Dunque il determinante è un invariante di similitudine. Per la proposizione precedente sicuramente $\#(\mathcal{M}(n, \mathbb{K})/\sim) \geq \#(\mathbb{K})$, dunque se \mathbb{K} è infinito anche $\mathcal{M}(n, \mathbb{K})/\sim$ ha infinite classi di similitudine.

DEFINIZIONE 3.2.3: $f \in \text{End}(V)$. Definiamo $\det(f) = \det(\mathfrak{M}_{\mathcal{B}}(f))$, dove \mathcal{B} è una base di V .

Osservazione: È una buona definizione, cioè non dipende dalla scelta della base. Infatti, se \mathcal{S} è un'altra base di V , allora le matrici $\mathfrak{M}_{\mathcal{B}}(f)$ e $\mathfrak{M}_{\mathcal{S}}(f)$ sono simili e dunque hanno lo stesso determinante.

PROPOSIZIONE 3.2.2: $f \sim g \Rightarrow \det(f) = \det(g)$.

Dimostrazione:

Sia \mathcal{B} una base di V . Allora:

$f \sim g \Rightarrow \mathfrak{M}_{\mathcal{B}}(f) \sim \mathfrak{M}_{\mathcal{B}}(g) \Rightarrow \det(\mathfrak{M}_{\mathcal{B}}(f)) = \det(\mathfrak{M}_{\mathcal{B}}(g)) \Rightarrow \det(f) = \det(g)$.

Osservazione: Quindi il determinante è un invariante di coniugio; con lo stesso controesempio di prima vediamo che $\{\text{rango}, \text{determinante}\}$ non è un sistema completo di invarianti per coniugio.

DEFINIZIONE 3.2.4: $f \in \text{End}(V)$. $\lambda \in \mathbb{K}$ si dice **autovalore** per f se $\exists v \in V, v \neq 0 \mid f(v) = \lambda v$.

In tal caso v è detto **autovettore** relativo a λ .

DEFINIZIONE 3.2.5: Definiamo **spettro** di f $Sp(f) = \{\lambda \in \mathbb{K} \mid \lambda \text{ è autovalore per } f\}$.

DEFINIZIONE 3.2.6: Diciamo che W sottospazio di V è un **sottospazio f -invariante** se $f(W) \subseteq W$.

Osservazioni: 1) Se v è autovettore $\Rightarrow f(\text{Span}(v)) \subseteq \text{Span}(v)$, quindi $\text{Span}(v)$ è f -invariante.

In particolare, se $\lambda \neq 0 \Rightarrow f(\text{Span}(v)) = \text{Span}(v)$, se $\lambda = 0 \Rightarrow f(\text{Span}(v)) = \{0\}$.

2) L'autovalore relativo ad un autovettore è univocamente determinato; infatti:

$f(v) = \lambda v = \mu v \Rightarrow (\lambda - \mu)v = 0 \Rightarrow (\text{poiché } v \neq 0) \Rightarrow \lambda = \mu$.

3) v è autovettore relativo a $0 \Leftrightarrow v \in \text{Ker}(f) \Leftrightarrow f$ non è iniettiva.

DEFINIZIONE 3.2.7: Definiamo **autospatio** relativo all'autovalore λ

$V_\lambda(f) = \{v \in V \mid f(v) = \lambda v\}$ (potremo scrivere V_λ al posto di $V_\lambda(f)$).

Osservazioni: 1) V_λ è un sottospazio vettoriale di V , poiché $V_\lambda = \text{Ker}(f - \lambda \text{id})$.

2) λ è autovalore per $f \Leftrightarrow \dim(V_\lambda) \geq 1$. In tal caso $V_\lambda = \{0\} \cup \{\text{autovettori relativi a } \lambda\}$.

3) $f(V_\lambda) \subseteq V_\lambda$, cioè V_λ è f -invariante.

4) $f|_{V_\lambda} = \lambda \text{id}|_{V_\lambda}$.

DEFINIZIONE 3.2.8: Se λ è autovalore per $f \in \text{End}(V)$, poniamo $\mu_g(\lambda) = \dim(V_\lambda)$, dove $\mu_g(\lambda)$ è detta **molteplicità geometrica** di λ .

Osservazione: \forall autovalore λ , $1 \leq \mu_g(\lambda) \leq \dim(V)$.

PROPOSIZIONE 3.2.3: Sia $f \sim g$. Allora:

1) $Sp(f) = Sp(g)$;

2) $\forall \lambda \in Sp(f)$, $\dim(V_\lambda(f)) = \dim(V_\lambda(g))$;

(ossia lo spettro e la molteplicità geometrica degli autovalori sono invarianti di coniugio).

Dimostrazione:

1) Per ipotesi $\exists h \in GL(V) \mid g = h^{-1} \circ f \circ h$.

\subseteq) Sia $\lambda \in Sp(f)$. Allora $\exists v \neq 0 \mid f(v) = \lambda v$.

Sia $w = h^{-1}(v)$. Allora $w \neq 0$.

$g(w) = (h^{-1} \circ f \circ h)(h^{-1}(v)) = h^{-1}(f(v)) = \lambda h^{-1}(v) = \lambda w$.

Dunque $\lambda \in Sp(g)$.

\supseteq) Analogo.

2) Abbiamo appena provato che $h^{-1}(V_\lambda(f)) \subseteq V_\lambda(g)$ e dunque $V_\lambda(f) \subseteq h(V_\lambda(g))$.

Allo stesso modo si prova che $V_\lambda(g) \subseteq h(V_\lambda(f))$, allora $V_\lambda(f) = h(V_\lambda(g))$.

Essendo h un isomorfismo, $\dim(V_\lambda(f)) = \dim(V_\lambda(g))$, tesi.

PROPOSIZIONE 3.2.4: $f \in \text{End}(V)$, \mathcal{B} base di V , $A = \mathfrak{M}_{\mathcal{B}}(f)$.

Sia $[\]_{\mathcal{B}}$ l'isomorfismo indotto da \mathcal{B} ($[\]_{\mathcal{B}}: V \rightarrow \mathbb{K}^n$). Allora:

- 1) λ è autovalore per $f \Leftrightarrow \lambda$ è autovalore per A ;
- 2) $V_{\lambda}(A) = [V_{\lambda}(f)]_{\mathcal{B}}$.

Dimostrazione:

- 1) λ è autovalore per $f \Leftrightarrow \exists v \neq 0 \mid f(v) = \lambda v \Leftrightarrow \exists v \neq 0 \mid \underbrace{[f(v)]_{\mathcal{B}}}_{AX} = \lambda \underbrace{[v]_{\mathcal{B}}}_X \Leftrightarrow$
 $\Leftrightarrow \exists X \in \mathbb{K}^n \setminus \{0\} \mid AX = \lambda X \Leftrightarrow \lambda$ è autovalore per A .
- 2) \subseteq) Sia $X \in V_{\lambda}(A)$ e sia $v \in V \mid [v]_{\mathcal{B}} = X$.
 $AX = \lambda X \Rightarrow [f(v)]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}} \Rightarrow f(v) = \lambda v \Rightarrow v \in V_{\lambda}(f)$.
 \supseteq) Sia $v \in V_{\lambda}(f) \Rightarrow f(v) = \lambda v$ e sia $X = [v]_{\mathcal{B}}$.
 Allora $[f(v)]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}} \Rightarrow AX = \lambda X \Rightarrow X = [v]_{\mathcal{B}} \in V_{\lambda}(A)$.

Osservazione: Dunque possiamo calcolare gli autovalori e gli autospazi di f usando una qualsiasi matrice associata.

CALCOLO DI AUTOVALORI E AUTOVETTORI PER $A \in \mathcal{M}(n, \mathbb{K})$:

λ è autovalore per $A \Leftrightarrow \exists X \neq 0 \mid AX = \lambda X \Leftrightarrow \exists X \neq 0 \mid X \in \text{Ker}(A - \lambda I) \Leftrightarrow \det(A - \lambda I) = 0$
 (poiché se $\det(A - \lambda I) \neq 0$, allora il sistema $(A - \lambda I)X = 0$ avrebbe una soluzione, $X = 0$, assurdo).

DEFINIZIONE 3.2.9: Il polinomio $p_A(t) = \det(A - tI)$ è detto **polinomio caratteristico** di A .

Osservazione: Gli autovalori di A sono le radici del polinomio caratteristico p_A .

Osservazioni: 1) Se $T \in \mathcal{T}(n, \mathbb{K})$ è triangolare superiore, cioè è del tipo:

$$T = \begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

allora:

$$T - tI = \begin{pmatrix} a_1 - t & & * \\ & \ddots & \\ 0 & & a_n - t \end{pmatrix}$$

Perciò $p_T(t) = (a_1 - t) \cdot \dots \cdot (a_n - t)$, dunque gli autovalori sono gli elementi sulla diagonale.

- 2) Se $A = \begin{pmatrix} M & | & N \\ \hline 0 & | & P \end{pmatrix}$, con $M, P \in \mathcal{M}(n, \mathbb{K})$, $\det(A) = \det(M) \cdot \det(P)$.

$$A - tI = \begin{pmatrix} M - tI & | & N \\ \hline 0 & | & P - tI \end{pmatrix}$$

Quindi $p_A(t) = \det(A - tI) = \det(M - tI) \cdot \det(P - tI) = p_M(t) \cdot p_P(t)$.

PROPOSIZIONE 3.2.5: $\forall A \in \mathcal{M}(n, \mathbb{K})$, $p_A(t) = (-1)^n t^n + (-1)^{n-1} \text{tr}(A) t^{n-1} + \dots + \det(A)$.

Dimostrazione:

Sicuramente il coefficiente di t^n è $(-1)^n$; inoltre $p_A(0) = \det(A - 0I) = \det(A)$.

Mostriamo per induzione su n che il coefficiente di t^{n-1} è $(-1)^{n-1} \text{tr}(A)$.

Passo base): $n = 2$, $A - tI = \begin{pmatrix} a-t & b \\ c & d-t \end{pmatrix} \Rightarrow \det(A - tI) = t^2 - (a+d)t + ad - bc$, verificato.

Passo induttivo): Sia $A \in \mathcal{M}(n, \mathbb{K})$ e siano $a_{ij} = [A]_{ij}$.

$$\begin{aligned} \det(A - tI) &= \sum_{i=1}^n (-1)^{i+1} [A - tI]_{1i} \det((A - tI)_{1i}) \\ &= (a_{11} - t) \det((A - tI)_{11}) + \sum_{i=2}^n (-1)^{i+1} a_{1i} \det((A - tI)_{1i}) \end{aligned}$$

Notiamo che l'addendo $\sum_{i=2}^n (-1)^{i+1} a_{1i} \det((A - tI)_{1i})$ contiene al massimo termini di grado $n - 2$, poiché "cancellando" la prima riga e l' i -esima colonna, con $i \neq 1$, vengono "cancellati" due elementi sulla diagonale che contengono la variabile t , dunque con il prodotto degli altri $n - 2$ termini sulla diagonale si raggiunge al massimo l'esponente $n - 2$.

Per ipotesi induttiva $\det((A - tI)_{11}) = (-1)^{n-1} t^{n-1} + (-1)^{n-2} (\sum_{j=2}^n a_{jj}) t^{n-2} + r(t)$, con $\deg(r(t)) \leq n - 2$. Dunque:

$$\begin{aligned} (a_{11} - t) \det((A - tI)_{11}) &= (a_{11} - t) \left((-1)^{n-1} t^{n-1} + (-1)^{n-2} \left(\sum_{j=2}^n a_{jj} \right) t^{n-2} + r(t) \right) \\ &= (-1)^n t^n + (-1)^{n-1} a_{11} t^{n-1} + (-1)^{n-1} \left(\sum_{j=2}^n a_{jj} \right) t^{n-1} + r'(t) \end{aligned}$$

con $\deg(r'(t)) \leq n - 2$.

Quindi il coefficiente del termine di grado $n - 1$ di $\det(A - tI)$ è:

$$(-1)^{n-1} a_{11} + (-1)^{n-1} \left(\sum_{j=2}^n a_{jj} \right) = (-1)^{n-1} \sum_{j=1}^n a_{jj} = (-1)^{n-1} \text{tr}(A),$$

tesi.

Osservazione: Si può dimostrare in generale che il coefficiente del termine di grado $n - i$ di $p_A(t)$ è:

$$(-1)^{n-i} \sum_{j=1}^{n-i+1} \det(M_A(j, i))$$

dove $M_A(j, i)$ è la sottomatrice quadrata di A che ha come diagonale gli elementi dal j -esimo al $(i + j - 1)$ -esimo della diagonale di A .

Osservazioni: 1) Abbiamo dimostrato che $\text{tr}(A)$ è la somma degli autovalori di A (contati con molteplicità), poiché se a_1, \dots, a_n sono gli autovalori di A (eventualmente $\notin \mathbb{K}$), $p_A(t) = (a_1 - t) \dots (a_n - t)$, quindi il coefficiente del termine di grado $n - 1$ di $p_A(t)$ è la somma degli autovalori.

2) Se $\mathbb{K} = \mathbb{R}$ e $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}(n, \mathbb{R})$, cioè A è una rotazione di 90° , $p_A(t) = t^2 + 1$, che è irriducibile su \mathbb{R} , dunque A non ha autovalori reali.

3) In ogni caso $A \in \mathcal{M}(n, \mathbb{K})$ ha al massimo n autovalori distinti.

PROPOSIZIONE 3.2.6: Se $A \sim B$, allora $p_A(t) = p_B(t)$.

Dimostrazione:

$B = M^{-1}AM$, dunque:

$$p_B(t) = \det(B - tI) = \det(M^{-1}AM - tI) = \det(M^{-1}(A - tI)M) \stackrel{\square}{=} \det(A - tI) = p_A(t),$$

dove il passaggio contrassegnato con \square deriva dalla formula di Binet.

Osservazione: Dunque il polinomio caratteristico è un invariante di similitudine; con esso lo sono anche tutti i suoi coefficienti, in particolare $\det(A)$ e $tr(A)$.

DEFINIZIONE 3.2.10: $f \in \text{End}(V)$. Poniamo $p_f(t) = p_A(t)$, dove A è la matrice associata a f in una qualsiasi base di V .

Osservazione: È una buona definizione, poiché se A' è la matrice associata ad f in un'altra base di V , $A \sim A'$ ed abbiamo dimostrato che se $A \sim A' \Rightarrow p_A(t) = p_{A'}(t)$.

COROLLARIO 3.2.7: Se $f \sim g$, allora $p_f(t) = p_g(t)$.

Dimostrazione:

Sia \mathcal{B} base di V .

$$f \sim g \Rightarrow \mathfrak{M}_{\mathcal{B}}(f) \sim \mathfrak{M}_{\mathcal{B}}(g) \Rightarrow p_f(t) = p_g(t).$$

DEFINIZIONE 3.2.11: $\forall \lambda \in Sp(f)$ denotiamo con $\mu_a(\lambda)$ la molteplicità algebrica di λ come radice di $p_f(t)$.

PROPOSIZIONE 3.2.8: $\forall n \in \mathbb{N}$, $A \sim B \Rightarrow A^n \sim B^n$.

Dimostrazione:

$$A \sim B \Rightarrow \exists M \in GL(n) \mid B = M^{-1}AM, \text{ dunque } B^n = \underbrace{M^{-1}AM \cdot \dots \cdot M^{-1}AM}_{n \text{ volte}} = M^{-1}A^nM \Rightarrow$$

$$\Rightarrow B^n \sim A^n.$$

Osservazione: La lista di invarianti per coniugio/similitudine trovati fin qui è:

- $rk(f)$;
- $\det(f)$;
- $Sp(f)$;
- p_f ;
- $\forall \lambda \in Sp(f)$, $\mu_g(\lambda)$ e $\mu_a(\lambda)$.

Questa lista è ridondante, infatti:

- Se $p_f = p_g \Rightarrow \begin{cases} \det(f) = \det(g) \\ Sp(f) = Sp(g) \\ \forall \lambda \in Sp(f), \mu_a(\lambda, f) = \mu_a(\lambda, g) \end{cases}$, infatti il determinante è il termine noto del polinomio caratteristico e gli autovalori sono le radici.
- Se $\forall \lambda \in Sp(f) = Sp(g)$, $\mu_g(\lambda, f) = \mu_g(\lambda, g) \Rightarrow rk(f) = rk(g)$.

Infatti:

$$rk(f) = n - \dim(Ker(f)) = n - \dim(V_0(f)) = n - \mu_g(0, f);$$

$$rk(g) = n - \dim(Ker(g)) = n - \dim(V_0(g)) = n - \mu_g(0, g).$$

Dunque gli invarianti significativi trovati finora sono:

- 1) il polinomio caratteristico;
- 2) la dimensione degli autospazi.

Questo non è un sistema completo di invarianti, infatti:

$$\text{sia } A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$p_A(t) = p_B(t) = t^4;$$

$$\dim(V_0(A)) = \dim(\text{Ker}(A)) = 2;$$

$$\dim(V_0(B)) = \dim(\text{Ker}(B)) = 2.$$

Ma poiché $A^2 \neq 0$ e $B^2 = 0$, allora sicuramente $A^2 \not\sim B^2 \Rightarrow A \not\sim B$.

PROPOSIZIONE 3.2.9: $f \in \text{End}(V)$, $\lambda \in \text{Sp}(f)$. Allora $1 \leq \mu_g(\lambda) \leq \mu_a(\lambda) \leq \dim(V)$.

Dimostrazione:

Sia $d = \mu_g(\lambda) = \dim(V_\lambda)$.

Sia $\{v_1, \dots, v_d\}$ una base di V_λ .

La completo a una base \mathcal{B} di V .

Allora $A = \mathfrak{M}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda I_d & | & M \\ \hline 0 & | & N \end{pmatrix}$, da cui $p_f(t) = (\lambda - t)^d \cdot p_N(t)$ e quindi $\mu_a(\lambda) \geq d$.

PROPOSIZIONE 3.2.10: Sia V uno spazio vettoriale e siano W_1, \dots, W_k sottospazi di V .

Sono fatti equivalenti:

- 1) $\dim(W_1 + \dots + W_k) = \dim(W_1) + \dots + \dim(W_k)$
- 2) Se \mathcal{B}_i è base di W_i , allora $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ è base di $W_1 + \dots + W_k$
- 3) Se $w_i \in W_i \ \forall i$ e $w_1 + \dots + w_k = 0 \Rightarrow w_1 = \dots = w_k = 0$
- 4) $\forall v \in W_1 + \dots + W_k$, v si scrive in modo unico come $v = w_1 + \dots + w_k$, con $w_i \in W_i \ \forall i$.

Dimostrazione:

2) \Rightarrow 1): Poiché \mathcal{B}_i è base di $W_i \Rightarrow \#\mathcal{B} = \#\mathcal{B}_1 + \dots + \#\mathcal{B}_k \Rightarrow$
 $\Rightarrow \dim(W_1 + \dots + W_k) = \dim(W_1) + \dots + \dim(W_k)$.

1) \Rightarrow 2): $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ genera $W_1 + \dots + W_k$ per 1). Inoltre $\#\mathcal{B} = \dim(W_1 + \dots + W_k)$, quindi \mathcal{B} è una base di $W_1 + \dots + W_k$.

2) \Rightarrow 3): Scrivo ogni w_i come combinazione lineare di \mathcal{B}_i :

$$w_1 + \dots + w_k = cl(\mathcal{B}_1) + \dots + cl(\mathcal{B}_k)$$

Ma $\mathcal{B}_1, \dots, \mathcal{B}_k$ fanno parte di \mathcal{B} e dunque sono indipendenti, dunque tutti gli addendi sono nulli $\Rightarrow w_i = 0 \ \forall i$.

3) \Rightarrow 4): Sia $v = w_1 + \dots + w_k = w'_1 + \dots + w'_k$. Allora:

$$\underbrace{(w_1 - w'_1)}_{\in W_1} + \dots + \underbrace{(w_k - w'_k)}_{\in W_k} = 0.$$

Quindi per 3), $w_i - w'_i = 0 \ \forall i$, da cui la tesi.

4) \Rightarrow 2): $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ genera $W_1 + \dots + W_k$.

Basta mostrare che \mathcal{B} è un insieme indipendente.

$$\underbrace{cl(\mathcal{B}_1)}_{=w_1} + \dots + \underbrace{cl(\mathcal{B}_k)}_{=w_k} = 0.$$

Allora $w_1 + \dots + w_k = 0 + \dots + 0 = 0$.

Ma poiché la scrittura è unica, $w_i = 0 \ \forall i$.

Poiché \mathcal{B}_i è base di W_i , tutti i coefficienti della combinazione lineare sono nulli \Rightarrow tesi.

DEFINIZIONE 3.2.12: Se vale una qualsiasi delle precedenti condizioni equivalenti, diciamo che W_1, \dots, W_k sono in **somma diretta (multipla)** e $W_1 \oplus \dots \oplus W_k = W_1 + \dots + W_k$.

Si ha che $\dim(W_1 \oplus \dots \oplus W_k) = \sum_{i=1}^k \dim(W_i)$.

PROPOSIZIONE 3.2.11: Sia $f \in \text{End}(V)$, $Sp(f) = \{\lambda_1, \dots, \lambda_k\}$.

Allora gli autospazi $V_{\lambda_1}, \dots, V_{\lambda_k}$ sono in somma diretta.

Dimostrazione:

Basta provare che se $v_1 \in V_{\lambda_1}, \dots, v_k \in V_{\lambda_k}$ e $v_1 + \dots + v_k = 0$, allora $v_i = 0 \forall i$.

Per induzione su k :

Passo base): $k = 1$: $v_1 = 0$ implica ovviamente $v_1 = 0$.

Passo induttivo): Dall'ipotesi $v_1 + \dots + v_k = 0$ ottengo, applicando f :

$\lambda_1 v_1 + \dots + \lambda_k v_k = 0$ e, moltiplicando invece per λ_k :

$\lambda_k v_1 + \dots + \lambda_k v_k = 0$.

Sottraendo ottengo:

$$\underbrace{(\lambda_1 - \lambda_k)v_1}_{\in V_{\lambda_1}} + \dots + \underbrace{(\lambda_{k-1} - \lambda_k)v_{k-1}}_{\in V_{\lambda_{k-1}}} = 0.$$

Per ipotesi induttiva:

$(\lambda_1 - \lambda_k)v_1 = 0, \dots, (\lambda_{k-1} - \lambda_k)v_{k-1} = 0$.

Ma $\lambda_i - \lambda_k \neq 0 \forall i$, poiché gli autovalori sono distinti. Dunque $v_1 = \dots = v_{k-1} = 0$ e quindi anche $v_k = 0$.

Osservazione: In generale $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} \subsetneq V$.

PROPOSIZIONE 3.2.12: Siano $A, B \in \mathcal{M}(n, \mathbb{R})$. Allora A e B sono simili su $\mathbb{R} \Leftrightarrow$ lo sono su \mathbb{C} .

Dimostrazione:

Precisiamo che:

$A \sim_{\mathbb{R}} B \Leftrightarrow \exists M \in GL(n, \mathbb{R}) \mid M^{-1}AM = B$;

$A \sim_{\mathbb{C}} B \Leftrightarrow \exists M \in GL(n, \mathbb{C}) \mid M^{-1}AM = B$.

\Rightarrow) ovvia.

\Leftarrow) Per ipotesi $\exists M \in GL(n, \mathbb{C}) \mid M^{-1}AM = B$, cioè $AM = MB$.

Sia $M = X + iY$, con $X, Y \in \mathcal{M}(n, \mathbb{R})$ (N.B.: non è detto che X e Y siano invertibili!).

$A(X + iY) = (X + iY)B \Rightarrow AX + iAY = XB + iYB \Rightarrow AX = XB, AY = YB$, separando parte reale e parte immaginaria.

Notiamo che:

$\forall t \in \mathbb{R}, A(X + tY) = AX + tAY = XB + tYB = (X + tY)B$,

dunque per arrivare alla tesi mi basta trovare $t \in \mathbb{R} \mid X + tY$ sia invertibile, poiché in quel caso A e B sarebbero simili grazie a $X + tY \in \mathcal{M}(n, \mathbb{R})$, cioè simili su \mathbb{R} .

$\det(X + tY)$ è un polinomio $p(t) \in \mathbb{R}[t]$, ma $p(i) = \det(X + iY) = \det(M) \neq 0$, dunque $p(t)$ non è il polinomio nullo $\Rightarrow \exists t_0 \in \mathbb{R} \mid p(t_0) \neq 0$, cioè $X + t_0Y \in GL(n, \mathbb{R})$, tesi.

3.3 DIAGONALIZZABILITÀ

DEFINIZIONE 3.3.1: $f \in \text{End}(V)$ si dice **diagonalizzabile** se esiste una base \mathcal{B} | $\mathfrak{M}_{\mathcal{B}}(f)$ è diagonale.

Osservazione: $\mathfrak{M}_{\mathcal{B}}(f)$ è diagonale $\Leftrightarrow \mathcal{B}$ è costituita da autovettori per f .

PROPOSIZIONE 3.3.1: $f \in \text{End}(V)$, $\text{Sp}(f) = \{\lambda_1, \dots, \lambda_k\}$.

Allora f è diagonalizzabile $\Leftrightarrow V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$.

Dimostrazione:

\Rightarrow Per ipotesi $\exists \mathcal{B}$ base di autovettori.

La suddivido in $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$, con $\mathcal{B}_i \subset V_{\lambda_i}$.

$V = \text{Span}(\mathcal{B}_1) \oplus \dots \oplus \text{Span}(\mathcal{B}_k) \subseteq V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$.

Ma $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} \subseteq V \Rightarrow$ tesi.

\Leftarrow Se \mathcal{B}_i è base di $V_{\lambda_i} \Rightarrow \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ è base di autovettori di V , dunque segue la tesi.

Osservazioni: 1) La proprietà di essere diagonalizzabile è un invariante di coniugio.

Dimostrazione:

$f \sim g \Rightarrow g = h^{-1} \circ f \circ h$.

Se $\{v_1, \dots, v_n\}$ è base di autovettori per f , $\{h^{-1}(v_1), \dots, h^{-1}(v_n)\}$ è base di autovettori per g .

2) $A \in \mathcal{M}(n, \mathbb{K})$ è diagonalizzabile $\Leftrightarrow A \sim D$ diagonale (poiché D ha una base di autovettori quindi ce l'ha anche A).

3) Sia \mathcal{S} base di V . f diagonalizzabile $\Leftrightarrow \mathfrak{M}_{\mathcal{S}}(f)$ è diagonalizzabile.

4) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$ non è diagonalizzabile (infatti se lo fosse sarebbe simile a una matrice diagonale, ma A ha come autovalore solo 1 \Rightarrow la matrice simile sarebbe $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, ma $A \not\sim I$).

5) $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$ non è diagonalizzabile perché non ha autovalori reali.

6) $f \in \text{End}(V) | f^2 = f$ (si dice che f è un **proiettore**) $\Rightarrow f$ è diagonalizzabile.

Dimostrazione:

Sicuramente $\forall v \in V, v = (v - f(v)) + f(v)$. Notiamo che:

$f(v - f(v)) = f(v) - f(f(v)) = f(v) - f(v) = 0 \Rightarrow v - f(v) \in \text{Ker}(f) = V_0(f)$;

$f(f(v)) = f(v) \Rightarrow f(v) \in \text{Im}(f) = V_1(f)$.

Poiché sappiamo che $V = \text{Ker}(f) \oplus \text{Im}(f) \Rightarrow V = V_0(f) \oplus V_1(f)$, da cui la tesi.

7) $f \in \text{End}(V) | f^2 = id$ (si dice che f è un **involuzione**) $\Rightarrow f$ è diagonalizzabile.

Dimostrazione:

Sicuramente $\forall v \in V, v = \frac{v+f(v)}{2} + \frac{v-f(v)}{2}$. Notiamo che:

$f\left(\frac{v+f(v)}{2}\right) = \frac{f(v)+v}{2} \Rightarrow \frac{v+f(v)}{2} \in V_1(f)$;

$f\left(\frac{v-f(v)}{2}\right) = \frac{f(v)-v}{2} \Rightarrow \frac{v-f(v)}{2} \in V_{-1}(f)$.

Abbiamo mostrato che $V_1(f)$ e $V_{-1}(f)$ generano V , dunque $V_1(f) + V_{-1}(f) \supseteq V$.

Poiché ovviamente $V_1(f) + V_{-1}(f) \subseteq V$, allora $V_1(f) + V_{-1}(f) = V$.

Inoltre $V_1(f) \cap V_{-1}(f) = \{0\}$, dunque $V_1(f) \oplus V_{-1}(f) = V$, tesi.

TEOREMA DI DIAGONALIZZABILITÀ: $f \in \text{End}(V)$, $\dim(V) = n$, $\text{Sp}(f) = \{\lambda_1, \dots, \lambda_k\}$.

Allora f è diagonalizzabile $\Leftrightarrow \begin{cases} \mu_a(\lambda_1) + \dots + \mu_a(\lambda_k) = n \\ \mu_a(\lambda_i) = \mu_g(\lambda_i) \quad \forall i \end{cases}$.

Dimostrazione:

$\Rightarrow \exists \mathcal{B}$ base di autovettori di f .

$$\mathfrak{M}_{\mathcal{B}}(f) = \begin{pmatrix} \boxed{\lambda_1 I_{d_1}} & & \\ & \ddots & \\ & & \boxed{\lambda_k I_{d_k}} \end{pmatrix}$$

con $d_1 + \dots + d_k = n$.

Allora $p_f(t) = (\lambda_1 - t)^{d_1} \cdot \dots \cdot (\lambda_k - t)^{d_k}$.

Da questo segue che $\mu_a(\lambda_i) = d_i$, quindi è provata la prima condizione.

Poiché \mathcal{B} contiene d_i autovettori relativi a λ_i , si ha che $\dim(V_{\lambda_i}) \geq d_i = \mu_a(\lambda_i)$.

Ma $\mu_g(\lambda_i) \leq \mu_a(\lambda_i) \Rightarrow \mu_g(\lambda_i) = \mu_a(\lambda_i)$.

\Leftarrow So che $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} \subseteq V$ e che f è diagonalizzabile \Leftrightarrow vale l'uguaglianza.

Ma $\dim(V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}) = \dim(V_{\lambda_1}) + \dots + \dim(V_{\lambda_k}) = \sum_i \mu_g(\lambda_i) = \sum_i \mu_a(\lambda_i) = n \Rightarrow$ tesi.

Osservazione: La condizione $\mu_a(\lambda_1) + \dots + \mu_a(\lambda_k) = n$ equivale a dire che $p_f(t) = \prod_{i=1}^k (\lambda_i - t)^{m_i}$, cioè che $p_f(t)$ è completamente fattorizzabile.

COROLLARIO 3.3.2: Se $f \in \text{End}(V)$, con $\dim(V) = n$, ha n autovalori distinti, allora f è diagonalizzabile.

Dimostrazione:

Poiché $\mu_a(\lambda_i) > 0$, allora sicuramente $\mu_a(\lambda_i) = 1 \quad \forall i$ e dunque $\sum_i \mu_a(\lambda_i) = n$.

Inoltre $1 \leq \mu_g(\lambda_i) \leq \mu_a(\lambda_i) = 1 \Rightarrow \mu_g(\lambda_i) = 1 = \mu_a(\lambda_i) \quad \forall i \Rightarrow$ tesi grazie al teorema di diagonalizzabilità.

Osservazione: Se $\mathbb{K} = \mathbb{C}$ (o in generale se \mathbb{K} è algebricamente chiuso), la condizione $\sum_i \mu_a(\lambda_i) = n$ è sempre verificata.

Osservazioni: 1) Se A è diagonalizzabile, anche A^n lo è $\forall n \in \mathbb{N}$.

Dimostrazione:

A diagonalizzabile $\Rightarrow A \sim D$ diagonale $\Rightarrow A^n \sim D^n$ e sicuramente D^n è diagonale $\forall n \in \mathbb{N}$

(la dimostrazione formale può essere fatta in questo modo:

a) dimostrare per induzione su m che, prese $B, C \in \mathcal{D}(m, \mathbb{K})$, $BC \in \mathcal{D}(m, \mathbb{K})$;

b) sfruttare il fatto a) per dimostrare per induzione che $D^{n-1} \in \mathcal{D}(m, \mathbb{K}) \Rightarrow D^n \in \mathcal{D}(m, \mathbb{K})$.)

2) Se λ è autovalore per A , allora λ^m è autovalore per $A^m \quad \forall m \in \mathbb{N}$.

Dimostrazione:

Se $AX = \lambda X \Rightarrow A^2X = A(AX) = A(\lambda X) = \lambda(AX) = \lambda^2X$.

In generale, se $A^{n-1}X = \lambda^{n-1}X$, procedendo come sopra si mostra che $A^nX = \lambda^nX$.

PROPOSIZIONE 3.3.3: Siano $f, g \in \text{End}(V) \mid f \circ g = g \circ f$ e sia V_λ l'autospazio relativo a λ per f . Allora $g(V_\lambda) \subseteq V_\lambda$.

Dimostrazione:

Sia $v \in V_\lambda$. Allora $f(g(v)) = g(f(v)) = g(\lambda v) = \lambda g(v) \Rightarrow g(v) \in V_\lambda \Rightarrow$ tesi.

Osservazione: Dunque se due endomorfismi commutano, gli autospazi dell'uno sono invarianti per l'altro.

PROPOSIZIONE 3.3.4: $f \in \text{End}(V)$, $V = A \oplus B$, A, B sottospazi di V f -invarianti.

Allora f è diagonalizzabile $\Leftrightarrow f|_A$ e $f|_B$ sono diagonalizzabili.

Dimostrazione:

\Leftarrow) $\exists \mathcal{B}_A$ base di A di autovettori per f ;

$\exists \mathcal{B}_B$ base di B di autovettori per f ;

Perciò $\mathcal{B} = \mathcal{B}_A \cup \mathcal{B}_B$ è base di V di autovettori per f .

\Rightarrow) $\exists \mathcal{B} = \{v_1, \dots, v_n\}$ base di V di autovettori per f .

$f(v_i) = \lambda_i v_i \forall i$; inoltre $\forall i \exists! a_i \in A, b_i \in B \mid v_i = a_i + b_i$.

$f(v_i) = \lambda_i v_i = \lambda_i(a_i + b_i) = \underbrace{\lambda_i a_i}_{\in A} + \underbrace{\lambda_i b_i}_{\in B}$.

Ma $f(v_i) = f(a_i + b_i) = \underbrace{f(a_i)}_{\in A} + \underbrace{f(b_i)}_{\in B}$, poiché A, B sono f -invarianti.

Ma lo spezzamento è unico, quindi $f(a_i) = \lambda_i a_i$; $f(b_i) = \lambda_i b_i \forall i$.

Dunque ho trovato $a_1, \dots, a_n \in A, b_1, \dots, b_n \in B \mid f(a_i) = \lambda_i a_i$ e $f(b_i) = \lambda_i b_i \forall i$.

Ma $a_1, \dots, a_n \in A$ provengono, tramite la proiezione $\pi_A: V \rightarrow A$, da una base di V , dunque gli a_i sono generatori di $\text{Im}(\pi_A)$. Ma π_A è surgettiva, dunque gli a_i generano A .

Dunque posso estrarre una base per A che è di autovettori perché sono $\neq 0$ (se ci fossero l'algoritmo di estrazione li eliminerebbe) e verificano $f(a_i) = \lambda_i a_i$.

Il discorso per B è analogo \Rightarrow tesi.

PROPOSIZIONE 3.3.5: $f \in \text{End}(V)$ diagonalizzabile; W sottospazio di V f -invariante.

Allora $f|_W$ è diagonalizzabile.

Dimostrazione 1:

Per la proposizione precedente mi basta trovare U sottospazio di $V \mid f(U) \subseteq U$ e $V = W \oplus U$.

f è diagonalizzabile $\Rightarrow \exists$ base di autovettori $\{v_1, \dots, v_n\}$ di V per f .

Se $\{w_1, \dots, w_p\}$ è una base di W , $\{w_1, \dots, w_p, v_1, \dots, v_n\}$ generano V e se applico l'algoritmo di estrazione a base ottengo una base $\{w_1, \dots, w_p, v_{j_1}, \dots, v_{j_{n-p}}\}$ di V .

Sia $U = \text{Span}(v_{j_1}, \dots, v_{j_{n-p}})$. Per costruzione $V = U \oplus W$ e inoltre $f(U) \subseteq U$ poiché questi sono elementi di una base di autovettori.

Dimostrazione 2:

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

Proviamo che $W = (W \cap V_{\lambda_1}) \oplus \dots \oplus (W \cap V_{\lambda_k})$.

$\forall w \in W, \exists v_1, \dots, v_k \mid w = v_1 + \dots + v_k$ e $v_i \in V_{\lambda_i}$.

Se provo che $v_i \in W \forall i$ ho la tesi, poiché avrei che W è composto solo da autovettori.

$w = v_1 + \dots + v_k$;

$f(w) = \lambda_1 v_1 + \dots + \lambda_k v_k$;

$f^2(w) = \lambda_1^2 v_1 + \dots + \lambda_k^2 v_k$;

\vdots

$f^{k-1}(w) = \lambda_1^{k-1} v_1 + \dots + \lambda_k^{k-1} v_k$.

Dunque:

$$\begin{pmatrix} w \\ f(w) \\ \vdots \\ f^{k-1}(w) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_k \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \dots & \lambda_k^{k-1} \end{pmatrix}}_{=V} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix}$$

V è invertibile perché è una matrice di Vandermonde con $\lambda_i \neq \lambda_j \forall i, j$, poiché i corrispondenti autospazi sono in somma diretta, perciò:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} = V^{-1} \begin{pmatrix} w \\ f(w) \\ \vdots \\ f^{k-1}(w) \end{pmatrix}$$

Quindi i v_i si ottengono come combinazione lineare degli $f^j(w) \Rightarrow$

$\Rightarrow v_i \in \text{Span}(w, f(w), \dots, f^{k-1}(w)) \forall i$.

Ma i $f^j(w)$ sono tutti contenuti in W , poiché W è un sottospazio f -invariante, dunque $v_i \in W \forall i$, da cui la tesi.

DEFINIZIONE 3.3.2: $f, g \in \text{End}(V)$ si dicono **simultaneamente diagonalizzabili** se ammettono una base comune di autovettori.

TEOREMA DI DIAGONALIZZAZIONE SIMULTANEA: Siano $f, g \in \text{End}(V) \mid f \circ g = g \circ f$. Allora:

- 1) Se f è diagonalizzabile con $n = \dim(V)$ autovalori distinti, allora g è diagonalizzabile e f e g sono simultaneamente diagonalizzabili
- 2) Se f e g sono diagonalizzabili lo sono simultaneamente.

Dimostrazione:

- 1) $\exists \mathcal{B}$ base di V , $\mathcal{B} = \{v_1, \dots, v_n\}$ ciascun $\text{Span}(v_i)$ coincide con un autospazio (in quanto tutti gli autospazi di f hanno dimensione 1 per ipotesi).

Dunque sappiamo che $g(\text{Span}(v_i)) \subseteq \text{Span}(v_i)$, cioè $g(v_i) = \lambda_i v_i$ per qualche $\lambda_i \in \mathbb{K} \forall i$ (infatti in generale so che $g(\text{Span}(v_i)) \subseteq V_{\lambda_i}$, ma $V_{\lambda_i} = \text{Span}(v_i)$)

Dunque segue la tesi.

- 2) f diagonalizzabile $\Rightarrow V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$.

$f \circ g = g \circ f \Rightarrow g(V_{\lambda_i}) \subseteq V_{\lambda_i} \forall i$.

Sappiamo che g diagonalizzabile e $g(V_{\lambda_i}) \subseteq V_{\lambda_i}$ implica che $g|_{V_{\lambda_i}}$ è diagonalizzabile.

Se \mathcal{B}_{λ_i} è una base di autovettori per $g|_{V_{\lambda_i}}$, allora $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \dots \cup \mathcal{B}_{\lambda_k}$ è una base di V (poiché $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$) di autovettori per f e g (questo per costruzione, perché prendendo un $x \in \mathcal{B}_{\lambda_i}$ ho sicuramente un autovettore per f , ma questi sono autovettori per $g|_{V_{\lambda_i}}$, dunque sono autovettori anche per g).

Dunque segue la tesi.

3.4 TRIANGOLABILITÀ

DEFINIZIONE 3.4.1: $f \in \text{End}(V)$ si dice triangolabile se $\exists B$ base di $V \mid \mathfrak{M}_B(f)$ è triangolare.

Osservazioni: 1) $A \in \mathcal{M}(n, \mathbb{K})$ è triangolabile $\Leftrightarrow A \sim T$ triangolare.

2) f diagonalizzabile $\Rightarrow f$ triangolabile.

Il viceversa è falso, infatti:

$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ è triangolabile ma non è diagonalizzabile.

3) $f \in \text{End}(V)$, \mathcal{S} base qualsiasi di V .

Allora f è triangolabile $\Leftrightarrow A = \mathfrak{M}_{\mathcal{S}}(f)$ è triangolare.

DEFINIZIONE 3.4.2: $\dim(V) = n$. Si chiama **bandiera** per V ogni famiglia $\{V_i\}_{i \in J_n}$ di sottospazi vettoriali di V tali che:

1) $V_1 \subset V_2 \subset \dots \subset V_n$

2) $\forall i \dim(V_i) = i$.

Osservazioni: 1) Ogni base $\{v_1, \dots, v_n\}$ di V induce una bandiera per V : $V_i = \text{Span}(v_1, \dots, v_i)$.

2) \forall bandiera per V , $\exists B$ base che la induce (basta scegliere come i -esimo vettore di base un vettore $\in V_i$ e $\notin V_{i-1}$).

DEFINIZIONE 3.4.3: $f \in \text{End}(V)$, B base di V . B si dice **base a bandiera** per f se i sottospazi della bandiera indotta da B sono f -invarianti, cioè $\forall i f(\text{Span}(v_1, \dots, v_i)) \subseteq \text{Span}(v_1, \dots, v_i)$.

PROPOSIZIONE 3.4.1: f triangolabile $\Leftrightarrow \exists$ base a bandiera per f .

Dimostrazione:

\Rightarrow f triangolabile $\Rightarrow A = \mathfrak{M}_B(f)$ triangolare $\Rightarrow A \sim T$ triangolare \Rightarrow se \mathcal{S} è la base tale che $\mathfrak{M}_{\mathcal{S}}(f) = T$, allora \mathcal{S} è evidentemente a bandiera.

\Leftarrow Se $f(v_i) \in \text{Span}(v_1, \dots, v_i) \forall i$, allora $\forall i$ le coordinate di $f(v_i)$ dalla $(i+1)$ -esima alla n -esima sono nulle, da cui la tesi.

Osservazione: Non tutti gli endomorfismi sono triangolabili (poiché il primo vettore di una base a bandiera deve essere autovettore per f). Ad esempio:

$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, non avendo autovalori reali, non è triangolabile.

TEOREMA DI TRIANGOLABILITÀ: f è triangolabile $\Leftrightarrow p_f(t)$ è completamente fattorizzabile (cioè se $\sum_x \mu_a(x) = \dim(V)$).

Dimostrazione:

\Rightarrow Per ipotesi $\exists B$ base di V tale che:

$$A = \mathfrak{M}_B(f) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Allora $p_f(t) = p_A(t) = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t) = \prod_{i=1}^k (\lambda_i - t)^{m_i}$, tesi.

⇔) Per induzione su $n = \dim(V)$:

Passo base): $n = 1$: ovvio, poiché ogni matrice 1×1 è triangolare.

Passo induttivo): Per ipotesi $\exists \lambda_1$ autovalore per f (poiché $p_f(t)$ è completamente fattorizzabile e dunque ha almeno una radice). Sia v_1 autovettore relativo a λ_1 .

Completo v_1 a base $\mathcal{S} = \{v_1, \dots, v_n\}$ di V .

Sia $V_1 = \text{Span}(v_1)$ e $W = \text{Span}(v_2, \dots, v_n)$. Allora $V = V_1 \oplus W$.

$$\mathfrak{M}_{\mathcal{S}}(f) = \left(\begin{array}{c|ccc} \lambda_1 & & & \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & C \end{array} \right)$$

Siano $\pi_{V_1}: V \rightarrow V_1$, $\pi_W: V \rightarrow W$ le proiezioni indotte dalla somma diretta.

Osservo che $\mathcal{T} = \{v_2, \dots, v_n\}$ è base di W e $\mathfrak{M}_{\mathcal{T}}(\pi_W \circ f|_W) = C$.

Ora:

$$p_f(t) = (\lambda_1 - t) \cdot p_C(t).$$

Ma $p_f(t)$ è completamente fattorizzabile, quindi anche $p_C(t)$ lo è. Inoltre $\pi_W \circ f|_W \in \text{End}(W)$, quindi per ipotesi induttiva $\exists \{w_2, \dots, w_n\}$ base di W a bandiera per $\pi_W \circ f|_W$.

È ovvio che $\{v_1, w_2, \dots, w_n\}$ è base di V , ma è anche a bandiera, infatti:

$f(v_1) = \lambda_1 \cdot v_1$, poiché v_1 è autovettore;

$$f(w_i) = \left(\underbrace{(\pi_{V_1} + \pi_W)}_{=id} \circ f \right) (w_i) = \underbrace{(\pi_{V_1} \circ f)}_{\in V_1} (w_i) + \underbrace{(\pi_W \circ f)}_{\in \text{Span}(w_2, \dots, w_i)} (w_i),$$

e $(\pi_W \circ f)(w_i) \in \text{Span}(w_2, \dots, w_i)$ poiché per ipotesi induttiva $\{w_2, \dots, w_n\}$ è una base di W a bandiera per $\pi_W \circ f|_W$.

Dunque $f(w_i) \in \text{Span}(v_1, w_2, \dots, w_n)$, da cui la tesi.

COROLLARIO 3.4.2: Se \mathbb{K} è algebricamente chiuso (cioè ogni polinomio in \mathbb{K} è completamente fattorizzabile), allora tutti gli endomorfismi di V sono triangolabili.

COROLLARIO 3.4.3: La proprietà di essere triangolabile è una proprietà invariante per coniugio/similitudine (poiché dipende solo dal polinomio caratteristico).

Osservazione: Le matrici $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ sono simili poiché entrambe rappresentano l'endomorfismo $f| f(e_1) = v_1$ e $f(e_2) = v_1 + v_2$, la prima nella base $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, la seconda nella base $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. D'altra parte, entrambe le matrici sono triangolabili (in quanto triangolari), ma la forma triangolabile è sostanzialmente diversa. Dunque, poiché nella stessa classe di similitudine possono coesistere elementi molto diversi, si dice che le matrici triangolari non sono "forme canoniche".

DEFINIZIONE 3.4.4: $A \in \mathcal{M}(n, \mathbb{K})$ si dice nilpotente se $\exists n \in \mathbb{N} | A^n = 0$.

PROPOSIZIONE 3.4.4: $A \in \mathcal{M}(n, \mathbb{K})$. A è nilpotente $\Leftrightarrow 0$ è autovalore con molteplicità n ($\Leftrightarrow p_A(t) = t^n$).

Dimostrazione:

\Rightarrow) Per ipotesi $\exists p \in \mathbb{N} \mid A^p = 0$. Sia $\lambda \in \overline{\mathbb{K}}$ un autovalore per A ($\overline{\mathbb{K}}$ è la chiusura algebrica di \mathbb{K} , cioè un campo dove tutti i polinomi di $\mathbb{K}[t]$ sono completamente fattorizzabili).

Allora λ^p è autovalore per $A^p = 0$, ma l'unico autovalore per 0 è $0 \Rightarrow \lambda^p = 0 \Rightarrow \lambda = 0$.

\Leftarrow) Per ipotesi A è triangolabile, cioè $A \sim T$ triangolare strettamente (poiché matrici simili hanno gli stessi autovalori, quindi T ha solo 0 come autovalore e dunque ha la diagonale nulla).

Dimostriamo ora per induzione su n che se $T \in \mathcal{M}(n, \mathbb{K})$ triangolare strettamente $\Rightarrow T^n = 0$:

Passo base): $n = 2$: $\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Passo induttivo): Sia $T' \in \mathcal{M}(n, \mathbb{K})$ triangolare strettamente:

$$T' = \left(\begin{array}{ccc|c} & & & X \\ & T & & \\ \hline 0 & \dots & 0 & 0 \end{array} \right) \Rightarrow T'^{n+1} = \left(\begin{array}{ccc|c} & & & T^n X \\ & T^{n+1} & & \\ \hline 0 & \dots & 0 & 0 \end{array} \right)$$

Ma per ipotesi induttiva (poiché T è triangolare strettamente), $T^n = 0$, quindi $T'^{n+1} = 0$.

Quindi, poiché $A^n \sim T^n = 0 \Rightarrow A^n = 0$, tesi.

Osservazione: L'esempio già visto:

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

prova che gli invarianti di similitudine trovati finora non sono un sistema completo di invarianti neppure nella classe degli endomorfismi triangolabili.

DEFINIZIONE 3.4.5: $f, g \in \text{End}(V)$ si dicono **simultaneamente triangolabili** se \exists base \mathcal{B} di V a bandiera sia per f sia per g .

TEOREMA DI TRIANGOLAZIONE SIMULTANEA: Siano $f, g \in \text{End}(V)$ triangolabili tali che $f \circ g = g \circ f$. Allora:

- 1) Se $W \subseteq V$ è un sottospazio f -invariante, allora $f|_W$ è triangolabile;
- 2) f e g ammettono un autovettore comune;
- 3) f e g sono simultaneamente triangolabili.

Dimostrazione:

- 1) Sia \mathcal{B}_W una base di W ; estendiamola a base \mathcal{B} di V . Sia inoltre $A_W = \mathfrak{M}_{\mathcal{B}_W}(f|_W)$.

Allora:

$$A = \mathfrak{M}_{\mathcal{B}}(f) = \left(\begin{array}{c|c} A_W & * \\ \hline 0 & C \end{array} \right)$$

Dunque $p_f(t) = p_A(t) = \det(A - tI) = \det(A_W - tI) \cdot \det(C - tI) = p_{A_W}(t) \cdot q(t) = p_{f|_W}(t) \cdot q(t)$.

Poiché $p_f(t)$ si fattorizza completamente, anche $p_{f|_W}(t)$ si fattorizza completamente $\Rightarrow f|_W$ è triangolabile.

- 2) f triangolabile $\Rightarrow \exists$ autospazio V_λ per f di dimensione > 0 .

$g(V_\lambda) \subseteq V_\lambda \Rightarrow$ per 1) $g|_{V_\lambda}$ è triangolabile $\Rightarrow \exists v \in V_\lambda$ autovettore per $g|_{V_\lambda}$, dunque per g .
 v è l'autovettore sia di f che di g cercato.

3) Per induzione su $n = \dim(V)$:

Passo base): $n = 1$: ovvio.

Passo induttivo): Per 2) sappiamo che $\exists v$ autovettore sia per f che per g . Sia $V_1 = \text{Span}(v)$.

Estendiamo v a base $\mathcal{B} = \{v, v_2, \dots, v_n\}$ di V . Sia $W = \text{Span}(v_2, \dots, v_n)$. Evidentemente

$V = V_1 \oplus W$. Allora:

$$A_1 = \mathfrak{M}_{\mathcal{B}}(f) = \left(\begin{array}{c|ccc} \lambda & & & \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \dots \\ \\ \\ C_1 \end{array} \right); \quad A_2 = \mathfrak{M}_{\mathcal{B}}(g) = \left(\begin{array}{c|ccc} \mu & & & \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \dots \\ \\ \\ C_2 \end{array} \right).$$

Siano $\pi_{V_1}: V \rightarrow V_1$ e $\pi_W: V \rightarrow W$ le proiezioni indotte dalla somma diretta $V = V_1 \oplus W$.

Allora $\pi_W \circ f, \pi_W \circ g \in \text{End}(W)$, inoltre:

$$(\pi_W \circ f) \circ (\pi_W \circ g) \stackrel{\square}{=} \pi_W \circ f \circ g = \pi_W \circ g \circ f = (\pi_W \circ g) \circ (\pi_W \circ f)$$

dove il passaggio contrassegnato con \square segue dal fatto che $\pi_W \circ f \circ \pi_W = \pi_W \circ f$, infatti, se $v \in V \Rightarrow v = v_1 + w$, con $v_1 \in V_1$ e $w \in W$, dunque:

$$(\pi_W \circ f \circ \pi_W)(v) = (\pi_W \circ f)(w) = \pi_W(f(w));$$

$$(\pi_W \circ f)(v) = (\pi_W \circ f)(v_1 + w) = \pi_W(\lambda v_1 + f(w)) = \pi_W(f(w)).$$

Quindi si può applicare l'ipotesi induttiva a $\pi_W \circ f$ e $\pi_W \circ g$, che dunque ammettono una base $\{w_2, \dots, w_n\}$ di W a bandiera sia per $\pi_W \circ f$ che per $\pi_W \circ g$.

Sicuramente $\{v, w_2, \dots, w_n\}$ è base di V e la verifica che è evidentemente a bandiera per f e per g è analoga a quella nel teorema di triangolarità.

3.5 FORMA CANONICA DI JORDAN

DEFINIZIONE 3.5.1: $\forall f \in \text{End}(V), \forall p(t) = a_0 + \dots + a_s t^s \in \mathbb{K}[t]$, poniamo $p(f) = a_0 f^0 + \dots + a_s f^s \in \text{End}(V)$, dove $f^0 = \text{id}$ e $\forall j \in \mathbb{N}, f^j = \underbrace{f \circ \dots \circ f}_{j \text{ volte}}$.

DEFINIZIONE 3.5.2: $S \subseteq \mathbb{K}[t]$ si definisce **ideale di polinomi** in $\mathbb{K}[t]$ se:

- $\forall p(t), q(t) \in S, p + q \in S$;
- $\forall p(t) \in S, \forall q(t) \in \mathbb{K}[t], pq \in S$.

DEFINIZIONE 3.5.3: $\forall f \in \text{End}(V)$, definiamo **ideale** di f , $I(f) = \{p(t) \in \mathbb{K}[t] \mid p(f) = 0\}$.

Osservazioni: 1) Se $g = h^{-1} \circ f \circ h$, con $h \in GL(V)$, allora $\forall p(t) \in \mathbb{K}[t], p(g) = h^{-1} \circ p(f) \circ h$, poiché $p(g) = a_0(h^{-1} \circ \text{id} \circ h) + a_1(h^{-1} \circ f \circ h) + \dots + a_s \underbrace{(h^{-1} \circ f \circ h \circ \dots \circ h^{-1} \circ f \circ h)}_{s \text{ volte}} =$

$$= a_0 \text{id} + a_1(h^{-1} \circ f \circ h) + \dots + a_s(h^{-1} \circ f^s \circ h) = h^{-1} \circ p(f) \circ h.$$

Dunque $f \sim g \Rightarrow p(f) \sim p(g) \forall p(t) \in \mathbb{K}[t]$.

2) $\forall p(t), q(t) \in \mathbb{K}[t]$ si ha che:

- $(p + q)(f) = p(f) + q(f)$;
- $(pq)(f) = p(f) \circ q(f) = q(f) \circ p(f)$

(le semplici verifiche sono lasciate al lettore).

Dunque $\forall f \in \text{End}(V)$ l'applicazione:

$$F: (\mathbb{K}[t], +, \cdot) \rightarrow (\text{End}(V), +, \circ) \\ p(t) \rightarrow p(f)$$

è un omomorfismo di anelli.

3) Se \mathcal{B} è base di V e $A = \mathfrak{M}_{\mathcal{B}}(f)$, allora $p(A) = \mathfrak{M}_{\mathcal{B}}(p(f))$.

PROPOSIZIONE 3.5.1: $f, g \in \text{End}(V)$. Allora:

- 1) Se $f \sim g$ allora $I(f) = I(g)$;
- 2) $I(f)$ è un ideale di $\mathbb{K}[t]$;
- 3) Se $A = \mathfrak{M}_{\mathcal{B}}(f)$, allora $I(f) = I(A)$.

Dimostrazione:

- 1) Le due inclusioni sono ovvie grazie al fatto che $g = h^{-1} \circ f \circ h \Rightarrow p(g) = h^{-1} \circ p(f) \circ h$.
- 2) Le due verifiche sono immediate.
- 3) $p(t) \in I(f) \Leftrightarrow p(f) = 0 \Leftrightarrow \mathfrak{M}_{\mathcal{B}}(p(f)) = 0 \Leftrightarrow p(A) = 0 \Leftrightarrow p(t) \in I(A)$.

Osservazione: Quindi $I(f)$ è un invariante di coniugio.

Osservazione: $I(f)$ contiene polinomi di grado ≥ 1 . Infatti, se $f = 0$, $t^s \in I(f) \forall s \in \mathbb{N}$.

Se $f \neq 0$ e $\dim(V) = n$, allora $\{f^0, \dots, f^{n^2}\}$ sono linearmente dipendenti in $\text{End}(V)$, poiché sono $n^2 + 1$ elementi in uno spazio di dimensione n^2 .

Dunque $\exists a_0, \dots, a_{n^2} \in \mathbb{K}$ non tutti nulli tali che:

$$a_0 f^0 + \dots + a_{n^2} f^{n^2} = 0.$$

Allora $p(t) = a_0 + a_1 t + \dots + a_{n^2} t^{n^2} \in \mathbb{K}[t]$ ha grado ≥ 1 e $p(t) \in I(f)$.

TEOREMA DI HAMILTON-CAYLEY: $\forall f \in \text{End}(V)$, $p_f \in I(f)$.

Dimostrazione:

Sia \mathcal{B} una base di V e $A = \mathfrak{M}_{\mathcal{B}}(f)$.

$p_f = p_A \Rightarrow \mathfrak{M}_{\mathcal{B}}(p_f(f)) = p_f(A) = p_A(A)$, dunque per provare che $p_f(f) = 0$ mi basta provare che $p_A(A) = 0$.

Procediamo per induzione su $n = \dim(V)$:

Passo base): $n = 1$: ovvio, poiché se $A = (a)$, allora $p_A(t) = a - t$, quindi $p_A(A) = aI - A = 0$.

Passo induttivo): 1° CASO: A è triangolabile.

Allora $\exists v_1 \in \mathbb{K}^n$ autovettore per A .

Lo completo a base di \mathbb{K}^n , $\mathcal{S} = \{v_1, \dots, v_n\}$.

Allora:

$$\tilde{A} = \mathfrak{M}_{\mathcal{S}}(A) = \left(\begin{array}{c|ccc} \lambda & & & \dots \\ \hline 0 & & & \\ \vdots & & & A_1 \\ 0 & & & \end{array} \right)$$

Poiché $A \sim \tilde{A}$, allora $p_A(A) \sim p_A(\tilde{A})$, quindi la tesi è provare che $p_A(\tilde{A}) = 0$, cioè non è restrittivo supporre $A = \tilde{A}$.

Osserviamo che $\forall m \in \mathbb{N}$:

$$A^m = \left(\begin{array}{c|c} \lambda & \dots \\ \hline 0 & A_1 \\ \vdots & \\ 0 & \end{array} \right)^m = \left(\begin{array}{c|c} \lambda^m & \dots \\ \hline 0 & A_1^m \\ \vdots & \\ 0 & \end{array} \right)$$

quindi $\forall q(t)$ si ha che:

$$q(A) = \left(\begin{array}{c|c} q(\lambda) & \dots \\ \hline 0 & q(A_1) \\ \vdots & \\ 0 & \end{array} \right).$$

Ora, $p_A(t) = (\lambda - t)p_{A_1}(t)$.

Ma poiché $p_A(A)$ è completamente fattorizzabile, allora anche $p_{A_1}(t)$ è completamente fattorizzabile $\Rightarrow A_1$ è triangolabile.

Per ipotesi induttiva $p_{A_1}(A_1) = 0$, dunque:

$$p_{A_1}(A) = \left(\begin{array}{c|c} p_{A_1}(\lambda) & \dots \\ \hline 0 & p_{A_1}(A_1) \\ \vdots & \\ 0 & \end{array} \right) = \left(\begin{array}{c|c} p_{A_1}(\lambda) & \dots \\ \hline 0 & 0 \\ \vdots & \\ 0 & \end{array} \right)$$

Allora $\forall v \in \mathbb{K}^n$:

$p_A(A)(v) = (\lambda I - A) \underbrace{p_{A_1}(A)(v)}_{\in \text{Span}(v_1)} = 0$, poiché $v_1 \in \text{Ker}(\lambda I - A)$ in quanto v_1 è autovettore

relativo all'autovalore λ .

CASO GENERALE: Utilizziamo il fatto che esiste un campo \mathbb{F} estensione di \mathbb{K} | $p_A(t)$ è completamente fattorizzabile in $\mathbb{F}[t]$ (ad esempio se $\mathbb{K} = \mathbb{R}, \mathbb{F} = \mathbb{C}$).

Si procede come nel caso precedente lavorando in \mathbb{F} e si trova che $p_A(A) = 0$ in $\mathcal{M}(n, \mathbb{F})$, dunque tale relazione vale anche in $\mathcal{M}(n, \mathbb{K})$, da cui la tesi.

PROPOSIZIONE 3.5.2: Sia I un ideale $\neq \{0\}$ di $\mathbb{K}[t]$. Allora esiste un unico polinomio monico $g(t) \in I$ che genera I , ossia $I = \{g(t)q(t) | q(t) \in \mathbb{K}[t]\}$ (se g genera I , scriviamo $I = (g)$).

Dimostrazione:

Esistenza): Sia $g(t) \in I$ monico, di grado ≥ 0 e minimo (cioè il polinomio con grado più basso in I).

Sia $a(t) \in I$. Per il teorema di divisione in $\mathbb{K}[t]$, $\exists! q(t), r(t) \in \mathbb{K}[t]$ tali che:

$$\begin{cases} a(t) = g(t)q(t) + r(t) \\ \deg(r(t)) \leq \deg(g(t)) \end{cases}$$

Allora $r(t) = a(t) - g(t)q(t)$.

Ma $a(t) \in I, g(t) \in I \Rightarrow g(t)q(t) \in I \Rightarrow r(t) \in I$.

Poiché $\deg(r(t)) \leq \deg(g(t))$, avrei trovato un polinomio di grado più piccolo di g in I , assurdo $\Rightarrow r(t) = 0$.

Unicità): Se $g_1, g_2 \in I$ sono due polinomi monici di grado minimo e ≥ 0 , allora:

$I = (g_1) = (g_2)$, quindi $g_1 | g_2$ e $g_2 | g_1$, poiché sono entrambi generatori.

Allora $g_1 = k \cdot g_2$, quindi $g_1 = g_2$, in quanto sono monici.

PROPOSIZIONE 3.5.3: Siano $a(t), b(t) \in \mathbb{K}[t]$ non nulli.

Sia $I(a(t), b(t)) = \{\varphi(t)a(t) + \psi(t)b(t) \mid \varphi(t), \psi(t) \in \mathbb{K}[t]\}$. Allora:

- 1) $I(a(t), b(t))$ è un ideale di $\mathbb{K}[t]$ che contiene polinomi non nulli;
- 2) Se $d(t)$ è il generatore monico di $I(a(t), b(t))$, allora $d(t)$ è il M.C.D. di $a(t)$ e $b(t)$.

Dimostrazione:

- 1) Le semplici verifiche sono lasciate al lettore.

Sicuramente $a(t), b(t) \in I(a(t), b(t))$, dunque l'ideale contiene polinomi non nulli.

- 2) Sicuramente $d(t)$ divide $a(t)$ e $b(t)$, poiché questi ultimi appartengono all'ideale e $d(t)$ è il generatore dell'ideale.

Sia $d_1(t)$ tale che divide sia $a(t)$ che $b(t)$. Allora qualsiasi $q(t) \in I(a(t), b(t))$ è diviso da $d_1(t)$, in quanto $q(t) = \varphi_0(t)a(t) + \psi_0(t)b(t) = d_1(t)(\varphi_0(t)a'(t) + \psi_0(t)b'(t))$.

Ma $d(t) \in I(a(t), b(t))$, dunque $d_1(t) \mid d(t)$, tesi.

IDENTITÀ DI BEZOUT: Se $d(t) = M.C.D.(a(t), b(t))$, allora $\exists \varphi(t), \psi(t) \in \mathbb{K}[t]$ tali che $d(t) = \varphi(t)a(t) + \psi(t)b(t)$.

Dimostrazione:

È un corollario immediato della precedente proposizione.

DEFINIZIONE 3.5.4: Sia $f \in \text{End}(V)$. Si definisce **polinomio minimo** di f il generatore $m_f(t)$ monico dell'ideale $I(f)$.

Osservazione: dal teorema di Hamilton-Cayley segue che $m_f \mid p_f \quad \forall f \in \text{End}(V)$.

Osservazione: Non conosciamo un metodo pratico e/o algoritmico per calcolare il polinomio minimo di un endomorfismo. Quindi è utile calcolare prima il polinomio caratteristico e poi risalire (a tentativi) al polinomio minimo.

Esempi:

- 1) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$; $p_A(t) = t^3$, inoltre $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$, dunque sicuramente $m_A(t) = t^3$
- 2) $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$; $p_B(t) = t^3$ e $B^2 = 0$, dunque $m_B(t) = t^2$.

Osservazione: Se W è sottospazio vettoriale di V e $f(W) \subseteq W$, allora $m_{f|_W} \mid m_f$. Infatti $f|_W$ è un endomorfismo e poiché m_f si annulla su V , a maggior ragione si annulla su W ; inoltre evidentemente $m_{f|_W}$ è polinomio minimo di $f|_W$, dunque $m_{f|_W} \mid m_f$.

PROPOSIZIONE 3.5.4: Sia λ autovalore per f . Allora $\forall q(t) \in I(f)$, $q(\lambda) = 0$.

Dimostrazione:

Sia $v \neq 0$, $v \in V \mid f(v) = \lambda v$. Per ipotesi $q(f) = 0$ e dunque $q(f)(v) = 0$.

Se $q(t) = a_0 + a_1 t + \dots + a_s t^s$,

$0 = q(f)(v) = (a_0 \text{id} + \dots + a_s f^s)(v) = a_0 v + a_1 \lambda v + \dots + a_s \lambda^s v = q(\lambda) \cdot v$.

Ma poiché $v \neq 0 \Rightarrow q(\lambda) = 0$.

LEMMA 3.5.7: $f \in \text{End}(V)$, $g(t) \in \mathbb{K}[t]$, allora $W = \text{Ker}(g(f))$ è f -invariante.

Dimostrazione:

Dobbiamo provare che $f(W) \subseteq W$, ossia che $\forall x \in W$, $g(f)(f(x)) = 0$.

Poiché $g(t) \cdot t = t \cdot g(t)$, si ha che $g(f) \circ f = f \circ g(f)$ e dunque:

$$g(f)(f(x)) = (g(f) \circ f)(x) = (f \circ g(f))(x) = f(g(f)(x)) \stackrel{\square}{=} f(0) = 0,$$

dove il passaggio contrassegnato da \square segue dal fatto che $x \in \text{Ker}(g(f))$.

TEOREMA DI DECOMPOSIZIONE PRIMARIA: $f \in \text{End}(V)$, $q(t) \in I(f)$.

Sia $q(t) = q_1(t)q_2(t)$, con $q_1(t), q_2(t) \in \mathbb{K}[t]$ e $M.C.D(q_1, q_2) = 1$. Allora:

- 1) $V = \text{Ker}(q_1(f)) \oplus \text{Ker}(q_2(f))$;
- 2) Gli addendi $\text{Ker}(q_1(f))$ e $\text{Ker}(q_2(f))$ sono f -invarianti;
- 3) Se $f \sim g$, allora $\begin{cases} \dim(\text{Ker}(q_1(f))) = \dim(\text{Ker}(q_1(g))) \\ \dim(\text{Ker}(q_2(f))) = \dim(\text{Ker}(q_2(g))) \end{cases}$.

Dimostrazione:

- 1) Per Bezout $\exists a(t), b(t) \in \mathbb{K}[t]$ tali che $1 = a(t)q_1(t) + b(t)q_2(t)$.

Valutando in f ho:

$$id = a(f) \circ q_1(f) + b(f) \circ q_2(f) \Rightarrow \forall v \in V, v = (a(f) \circ q_1(f))(v) + (b(f) \circ q_2(f))(v).$$

Notiamo che $(a(f) \circ q_1(f))(v) \in \text{Ker}(q_2(f))$, infatti:

$$\begin{aligned} q_2(f)(a(f) \circ q_1(f))(v) &= (a(f) \circ q_1(f) \circ q_2(f))(v) = (a(f) \circ (q_1q_2)(f))(v) = \\ &= (a(f) \circ q(f))(v) = 0, \end{aligned}$$

in quanto $q(f) = 0$.

Analogamente si prova che $(b(f) \circ q_2(f))(v) \in \text{Ker}(q_1(f))$.

Dunque abbiamo dimostrato che $V = \text{Ker}(q_1(f)) + \text{Ker}(q_2(f))$.

Sia ora $z \in \text{Ker}(q_1(f)) \cap \text{Ker}(q_2(f))$; allora:

$$z = \underbrace{(a(f) \circ q_1(f))(z)}_{=0 \text{ poichè } z \in \text{Ker}(q_1(f))} + \underbrace{(b(f) \circ q_2(f))(z)}_{=0 \text{ poichè } z \in \text{Ker}(q_2(f))} = 0,$$

da cui la tesi.

- 2) Già provato.

- 3) Se $g = k^{-1} \circ f \circ k$, con $k \in GL(V)$, sappiamo che $q_1(g) = k^{-1} \circ q_1(f) \circ k$, cioè

$q_1(g) \sim q_1(f)$, per cui $rk(q_1(g)) = rk(q_1(f))$, da cui

$$\dim(\text{Ker}(q_1(f))) = \dim(\text{Ker}(q_1(g))).$$

Analogamente per q_2 .

COROLLARIO 3.5.8: $f \in \text{End}(V)$, $q(t) \in I(f)$. Sia $q = q_1 \cdot \dots \cdot q_m$, con $M.C.D(q_i, q_j) = 1$

$\forall i \neq j$.

Allora $V = \text{Ker}(q_1(f)) \oplus \dots \oplus \text{Ker}(q_m(f))$ e gli addendi sono f -invarianti.

Dimostrazione:

Per induzione su m :

Passo base): $m = 2$: già fatto.

Passo induttivo): Poiché $m > 2$, poniamo $\tilde{q} = q_2 \cdot \dots \cdot q_m$.

Allora $q = q_1\tilde{q}$ e $M.C.D(q_1, \tilde{q}) = 1$.

Per il teorema di decomposizione primaria abbiamo che:

$$V = Ker(q_1(f)) \oplus \underbrace{Ker(\tilde{q}(f))}_{=W}$$

W è f -invariante e $\tilde{q} \in I(f|_W)$. Per ipotesi induttiva:

$$W = Ker(q_2(f|_W)) \oplus \dots \oplus Ker(q_m(f|_W)).$$

Ora $\forall 2 \leq j \leq m, Ker(q_j(f|_W)) = Ker(q_j(f)|_W) = Ker(q_j(f)) \cap W$.

D'altra parte $Ker(q_j(f)) \subseteq W$, in quanto se $x \in Ker(q_j(f))$:

$$\tilde{q}(f)(x) = (q_2 \cdot \dots \cdot q_m)(f)(x) = (\dots \cdot q_j)(f)(x) = 0.$$

Dunque $W = Ker(q_2(f)) \oplus \dots \oplus Ker(q_m(f))$, da cui la tesi.

COROLLARIO 3.5.9: $f \in End(V)$ triangolabile, $Sp(f) = \{\lambda_1, \dots, \lambda_k\}$.

$$p_f = (-1)^n (t - \lambda_1)^{h_1} \cdot \dots \cdot (t - \lambda_k)^{h_k};$$

$$m_f = (t - \lambda_1)^{s_1} \cdot \dots \cdot (t - \lambda_k)^{s_k}, \text{ con } 1 \leq s_i \leq h_i \quad \forall i.$$

Allora:

$$V = \bigoplus_{j=1}^k Ker(f - \lambda_j id)^{h_j}$$

$$V = \bigoplus_{j=1}^k Ker(f - \lambda_j id)^{s_j}$$

e ogni addendo è f -invariante.

PROPOSIZIONE 3.5.10: f è diagonalizzabile $\Leftrightarrow m_f = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k)$.

Dimostrazione:

$\Rightarrow \exists \mathcal{B}$ tale che:

$$\mathfrak{M}_{\mathcal{B}}(f) = \begin{pmatrix} \boxed{\lambda_1 I_{d_1}} & & \\ & \ddots & \\ & & \boxed{\lambda_k I_{d_k}} \end{pmatrix}$$

$$m_f = m.c.m((t - \lambda_1), \dots, (t - \lambda_k)) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k)$$

\Leftarrow Per la proposizione precedente $V = \underbrace{Ker(f - \lambda_1 id)}_{=V_{\lambda_1}} \oplus \dots \oplus \underbrace{Ker(f - \lambda_k id)}_{=V_{\lambda_k}}$, dunque segue la

tesi.

LEMMA 3.5.11: Sia $\varphi \in End(V)$. Allora:

- 1) $Ker(\varphi^j) \subseteq Ker(\varphi^{j+1}) \quad \forall j \in \mathbb{N}$;
- 2) Se $\exists m \in \mathbb{N} \mid Ker(\varphi^m) = Ker(\varphi^{m+1})$, allora $Ker(\varphi^m) = Ker(\varphi^{m+t}) \quad \forall t \geq 1$;
- 3) La successione $\left\{ \dim(Ker(\varphi^j)) \right\}_j$ è un invariante di coniugio.

Dimostrazione:

1) Se $\varphi^j(x) = 0$, allora $\varphi^{j+1}(x) = \varphi(\varphi^j(x)) = \varphi(0) = 0$.

2) Basta provare che $Ker(\varphi^{m+1}) = Ker(\varphi^{m+2})$ e poi iterare.

Ovviamente per 1) $Ker(\varphi^{m+1}) \subseteq Ker(\varphi^{m+2})$, quindi resta da mostrare il contenimento opposto.

Sia $x \in Ker(\varphi^{m+2})$. Allora $\varphi^{m+2}(x) = \varphi^{m+1}(\varphi(x)) = 0$, cioè $\varphi(x) \in Ker(\varphi^{m+1})$.

Ma $Ker(\varphi^m) = Ker(\varphi^{m+1})$, dunque $\varphi(x) \in Ker(\varphi^m)$, cioè $x \in Ker(\varphi^{m+1})$, tesi.

3) Già provato.

Osservazione: Se applichiamo il lemma a $\varphi = f - \lambda_j id$:

$$Ker(f - \lambda_j id) \subseteq \dots \subseteq Ker(f - \lambda_j id)^{s_j} \subseteq \dots \subseteq Ker(f - \lambda_j id)^{h_j}.$$

Ma poiché:

$$V = \bigoplus_{j=1}^k Ker(f - \lambda_j id)^{h_j} = \bigoplus_{j=1}^k Ker(f - \lambda_j id)^{s_j}$$

allora $Ker(f - \lambda_j id)^{s_j} = Ker(f - \lambda_j id)^{h_j} \quad \forall j$, quindi le due decomposizioni primarie coincidono.

DEFINIZIONE 3.5.5: \forall autovalore λ_j , il sottospazio f -invariante $V'_j = Ker(f - \lambda_j id)^{s_j}$ è detto **autospazio generalizzato** relativo a λ_j .

Osservazione: Se $\mathfrak{M}_B(f) = A$, gli autospazi generalizzati si possono calcolare risolvendo i sistemi lineari $(A - \lambda_j I)^{h_j} X = 0$.

Si ha dunque $V = V'_{\lambda_1} \oplus \dots \oplus V'_{\lambda_k}$.

PROPOSIZIONE 3.5.12: Sia $f \in End(V)$ e $Sp(f) = \{\lambda_1, \dots, \lambda_k\}$. Allora:

- 1) $\forall j, f|_{V'_{\lambda_j}}$ ha solo l'autovalore λ_j ;
- 2) $f|_{V'_{\lambda_j}}$ ha polinomio caratteristico $= \pm(t - \lambda_j)^{h_j}$ e polinomio minimo $= (t - \lambda_j)^{s_j}$;
- 3) $\dim(V'_{\lambda_j}) = h_j = \mu_a(\lambda_j)$;
- 4) Se poniamo $d_i = \dim(Ker(f - \lambda_j id)^i) \quad \forall 1 \leq i \leq s_j$, i numeri $d_1 < \dots < d_{s_j}$ sono invarianti di coniugio (Osservazione: $d_1 = \mu_g(\lambda_j), d_{s_j} = \mu_a(\lambda_j)$).

Dimostrazione:

- 1) Sicuramente $V_{\lambda_j} \subseteq V'_{\lambda_j}$.

Se $f|_{V'_{\lambda_j}}$ avesse un altro autovalore μ , allora $V_\mu \cap V'_{\lambda_j} \neq \{0\}$ e dunque $V'_\mu \cap V'_{\lambda_j} \neq \{0\}$, assurdo, poiché sono in somma diretta.

2,3) Sia \mathcal{B}_j una base di V'_{λ_j} . Allora $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ è una base di V .

$$\mathfrak{M}_B(f) = \begin{pmatrix} \boxed{M_1} & & \\ & \ddots & \\ & & \boxed{M_k} \end{pmatrix},$$

con M_j blocco quadrato di ordine $= \dim(V'_{\lambda_j})$.

$$p_f = p_{M_1} \cdot \dots \cdot p_{M_k}.$$

Poiché ogni blocco M_j ha solo l'autovalore λ_j , segue che $p_{M_j} = p_{f|_{V'_{\lambda_j}}} = \pm(t - \lambda_j)^{h_j}$ e

$$h_j = \text{ordine di } M_j = \dim(V'_{\lambda_j}).$$

Inoltre $m_f = m.c.m(m_{M_1}, \dots, m_{M_k}) = m_{M_1} \cdot \dots \cdot m_{M_k}$, da cui $m_{f|_{V'_{\lambda_j}}} = (t - \lambda_j)^{s_j}$.

4) Già provato.

Osservazione: Dunque possiamo ricondurci a studiare la restrizione di f a ciascun autospazio generalizzato; se λ è autovalore, $W = V'_\lambda$ e $\varphi = f|_W$, allora:

- $\dim(W) = \mu_\alpha(\lambda) = h$
- $p_\varphi = \pm(t - \lambda)^h$
- $m_\varphi = (t - \lambda)^s$, con $1 \leq s \leq h$
- se $d_i = \dim(\text{Ker}(\varphi - \lambda id)^i)$, la stringa $[\lambda, s, [d_1 < \dots < d_s = h]]$ è un invariante di coniugio.

Osservazione: Possiamo in ogni caso ridurci al caso nilpotente, poiché se $\psi = \varphi - \lambda id$, allora ψ ha solo l'autovalore 0 (cioè è nilpotente).

$$p_\psi = \pm t^h; \quad m_\psi = t^s.$$

La stringa di invarianti di ψ è $[0, s, [d_1 < \dots < d_s = h]]$.

PROPOSIZIONE 3.5.13: Sia $\psi \in \text{End}(V)$, $\dim(V) = h$, ψ nilpotente. Sono fatti equivalenti:

1) $m_\psi = t^h$ (ossia $m_\psi = \pm p_\psi$)

$$2) \exists \mathcal{B} \text{ base di } V \mid \mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix}.$$

Dimostrazione:

$$\Leftrightarrow \text{Ovvio, poiché } \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix}^h = 0 \text{ e } \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix}^{h-1} \neq 0.$$

$\Rightarrow \text{Ker}(\psi) \subseteq \dots \subseteq \text{Ker}(\psi^h)$, con $\dim(\text{Ker}(\psi^h)) = h$.

$$\psi^{h-1} \neq 0 \Rightarrow \exists v \notin \text{Ker}(\psi^{h-1}), v \neq 0.$$

Mostriamo che $\mathcal{B} = \{\psi^{h-1}(v), \dots, \psi(v), v\}$ è una base di V e in questo modo giungo alla tesi, poiché nella prima colonna di $\mathfrak{M}_{\mathcal{B}}(\psi)$ ci andrebbe $\psi(\psi^{h-1}(v)) = 0$, nella seconda $\psi(\psi^{h-2}(v)) = \psi^{h-1}(v)$ e così via.

Poiché i vettori sono in numero adeguato, basta che siano linearmente indipendenti.

$$\text{Sia } a_0 v + a_1 \psi(v) + \dots + a_{h-1} \psi^{h-1}(v) = 0.$$

Se applico ψ^{h-1} , ottengo $a_0 \psi^{h-1}(v) = 0$, ma $\psi^{h-1}(v) \neq 0$, quindi $a_0 = 0$.

Allo stesso modo se applico ψ^{h-2} ottengo $a_1 = 0$; iterando il processo si ottiene $a_0 = \dots = a_{h-1} = 0$, da cui la tesi.

DEFINIZIONE 3.5.6: La matrice $r \times r$:

$$J(\lambda, r) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{pmatrix}$$

è detta **blocco di Jordan** di ordine r relativo a λ .

DEFINIZIONE 3.5.7: Si chiama **matrice di Jordan** ogni matrice diagonale a blocchi:

$$J = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_n} \end{pmatrix}$$

in cui ogni blocco J_i è un blocco di Jordan.

DEFINIZIONE 3.5.8: Se $f \in \text{End}(V)$, si chiama base di Jordan per f ogni base \mathcal{B} tale che $\mathfrak{M}_{\mathcal{B}}(f)$ è una matrice di Jordan.

Osservazione: L'ultima proposizione afferma dunque che:

$$m_{\psi} = \pm p_{\psi} = \pm t^h \Leftrightarrow \exists \mathcal{B} \mid \mathfrak{M}_{\mathcal{B}}(\psi) = J(0, h).$$

In tal caso la stringa di invarianti è $[0, h, [1, 2, \dots, h-1, h]]$.

LEMMA 3.5.14: Sia $\psi \in \text{End}(V)$ nilpotente con indice di nilpotenza s (cioè $\psi^s = 0$ e $\psi^{s-1} \neq 0$) $\dim(V) = h$.

$\forall j \mid 3 \leq j \leq s$ si consideri $\text{Ker}(\psi^{j-2}) \subseteq \text{Ker}(\psi^{j-1}) \subseteq \text{Ker}(\psi^j)$.

Sia W un sottospazio vettoriale di V tale che $\text{Ker}(\psi^j) = \text{Ker}(\psi^{j-1}) \oplus W$ e sia $\{w_1, \dots, w_k\}$ una base di W . Allora:

- 1) $\psi(w_1), \dots, \psi(w_k) \in \text{Ker}(\psi^{j-1})$ e sono linearmente indipendenti;
- 2) $\text{Span}(\psi(w_1), \dots, \psi(w_k)) \cap \text{Ker}(\psi^{j-2}) = \{0\}$.

Dimostrazione:

- 1) Sicuramente $\psi(w_1), \dots, \psi(w_k) \in \text{Ker}(\psi^{j-1})$, poiché $\psi^{j-1}(\psi(w_j)) = \psi^j(w_j) = 0$.

Sia ora $a_1\psi(w_1) + \dots + a_k\psi(w_k) = 0$. Per linearità

$$\psi(a_1w_1 + \dots + a_kw_k) = 0 \Rightarrow \underbrace{a_1w_1}_{\in W} + \dots + \underbrace{a_kw_k}_{\in W} \in \text{Ker}(\psi) \cap W \subseteq \text{Ker}(\psi^{j-1}) \cap W = \{0\},$$

perciò $a_1w_1 + \dots + a_kw_k = 0$, da cui $a_1 = \dots = a_k = 0$ poiché $\{w_1, \dots, w_k\}$ è base di W .

- 2) Sia $z = a_1\psi(w_1) + \dots + a_k\psi(w_k) \in \text{Span}(\psi(w_1), \dots, \psi(w_k)) \cap \text{Ker}(\psi^{j-2})$.

$$\begin{aligned} z &= \psi(a_1w_1 + \dots + a_kw_k) \in \text{Ker}(\psi^{j-2}) \Rightarrow \psi^{j-1}(a_1w_1 + \dots + a_kw_k) = 0 \Rightarrow \\ &\Rightarrow a_1w_1 + \dots + a_kw_k \in \text{Ker}(\psi^{j-1}) \cap W = \{0\} \Rightarrow a_i = 0 \quad \forall i \Rightarrow z = 0. \end{aligned}$$

TEOREMA 3.5.15 (forma canonica di Jordan per endomorfismi nilpotenti): $\psi \in \text{End}(V)$ nilpotente, $\dim(V) = h$, con stringa di invarianti $[0, s, [d_1 < d_2 < \dots < d_s = h]]$. Allora:

- 1) $\exists \mathcal{B}$ base di V tale che:

$$\mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} J(0, n_1) & & \\ & \ddots & \\ & & J(0, n_t) \end{pmatrix},$$

dove $n_1 + \dots + n_t = h$.

- 2) La matrice $\mathfrak{M}_{\mathcal{B}}(\psi)$, detta **forma di Jordan** di ψ , è unica a meno di permutazioni dei blocchi sulla diagonale ed è completamente determinata dalla stringa di invarianti $[0, s, [d_1 < d_2 < \dots < d_s = h]]$.
- 3) La stringa di invarianti è un sistema completo di coniugio per endomorfismi nilpotenti. (Osservazione: Se conveniamo che i blocchi J_1, \dots, J_m sulla diagonale siano in ordine decrescente, allora la forma di Jordan di ψ è unica e possiamo denotarla con $J(\psi)$).

Dimostrazione:

- 1) Per ipotesi $m_{\psi}(t) = t^s$.

$$\text{Ker}(\psi) \subsetneq \text{Ker}(\psi^2) \subsetneq \dots \subsetneq \text{Ker}(\psi^s) = V.$$

Poniamo $d_i = \dim(\text{Ker}(\psi^i))$.

Sia W_s un supplementare di $\text{Ker}(\psi^{s-1})$ in $\text{Ker}(\psi^s)$, ossia $\text{Ker}(\psi^s) = \text{Ker}(\psi^{s-1}) \oplus W_s$.

Poniamo $r_s = \dim(W_s)$. (Osservazione: $r_s = d_s - d_{s-1} = h - d_{s-1}$).

Sia $\{v_{1,s}, \dots, v_{r_s,s}\}$ una base di W_s .

Poniamo $v_{j,s-1} = \psi(v_{j,s}) \quad \forall 1 \leq j \leq r_s$.

Per il lemma, $v_{1,s-1}, \dots, v_{r_s,s-1}$ stanno in $\text{Ker}(\psi^{s-1})$ e sono linearmente indipendenti;

dunque posso completarli a una base $\{v_{1,s-1}, \dots, v_{r_s,s-1}, v_{r_s+1,s-1}, \dots, v_{r_{s-1},s-1}\}$ di un

supplementare W_{s-1} di $\text{Ker}(\psi^{s-2})$ in $\text{Ker}(\psi^{s-1})$, ossia $\text{Ker}(\psi^{s-1}) = \text{Ker}(\psi^{s-2}) \oplus W_{s-1}$.

Itero il procedimento scegliendo supplementari $W_j \mid \text{Ker}(\psi^j) = \text{Ker}(\psi^{j-1}) \oplus W_j$,

$r_j = \dim(W_j)$.

I vettori così trovati possono essere organizzati in una tabella:

$$\begin{array}{cccccccc}
 v_{1,s} & \dots & v_{r_s,s} & & & & & \text{base di } W_s \\
 \downarrow & & \downarrow & & & & & \\
 v_{1,s-1} & \dots & v_{r_s,s-1} & \dots & v_{r_{s-1},s-1} & & & \text{base di } W_{s-1} \\
 \downarrow & & \downarrow & & \downarrow & & & \\
 \vdots & & \vdots & & \vdots & & & \\
 \downarrow & & \downarrow & & \downarrow & & & \\
 v_{1,2} & \dots & v_{r_s,2} & \dots & v_{r_{s-1},2} & \dots & v_{r_2,2} & \text{base di } W_2 \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
 v_{1,1} & \dots & v_{r_s,1} & \dots & v_{r_{s-1},1} & \dots & v_{r_2,1} & \dots & v_{r_1,1} & \text{base di } \text{Ker}(\psi)
 \end{array}$$

dove $r_1 = \dim(\text{Ker}(\psi)) = d_1$.

Si ha:

$$V = \text{Ker}(\psi^s) = \text{Ker}(\psi^{s-1}) \oplus W_s = \text{Ker}(\psi^{s-2}) \oplus W_{s-1} \oplus W_s = \dots$$

$$\dots = \text{Ker}(\psi) \oplus W_2 \oplus \dots \oplus W_s,$$

dunque per costruzione i vettori presenti nella tabella formano una base di V (in quanto sono linearmente indipendenti e in numero adeguato).

Inoltre ad esempio:

$$v_{1,1} = \psi(v_{1,2}) = \psi^2(v_{1,3}) = \dots = \psi^{s-1}(v_{1,s})$$

dunque:

$$\{v_{1,1}, v_{1,2}, \dots, v_{1,s}\} = \{\psi^{s-1}(v_{1,s}), \psi^{s-2}(v_{1,s}), \dots, v_{1,s}\}$$

Pertanto la base \mathcal{B} ottenuta riordinando i vettori della tabella procedendo da sinistra verso destra e risalendo le colonne è una base di Jordan.

Ogni colonna "alta" j produce in $\mathfrak{M}_{\mathcal{B}}(\psi)$ un blocco di Jordan di ordine j .

- 2) Mostriamo che il tipo dei blocchi in $\mathfrak{M}_{\mathcal{B}}(\psi)$ è completamente determinato dalla stringa di invarianti $[0, s, [d_1 < d_2 < \dots < d_s = h]]$.

Notiamo innanzitutto che, per quanto visto nel punto 1):

- s = indice di nilpotenza = massimo ordine dei blocchi in $\mathfrak{M}_{\mathcal{B}}(\psi)$ (in quanto l'altezza di una colonna non può superare l'indice di nilpotenza);
- $d_1 = \dim(\text{Ker}(\psi))$ = numero totale dei blocchi (cioè il numero delle colonne).

Più precisamente, poniamo b_j = numero dei blocchi $j \times j$ in $\mathfrak{M}_{\mathcal{B}}(\psi)$. Allora:

$$b_s = \dim(W_s) = r_s;$$

$$b_j = \dim(W_j) - \dim(W_{j+1}) = r_j - r_{j+1}.$$

Se poniamo $r_{s+1} = 0$, allora $b_j = r_j - r_{j+1} \quad \forall 1 \leq j \leq s$.

Esprimiamo i b_j in funzione dei d_j :

$$r_j = \dim(W_j) = \dim(Ker(\psi^j)) - \dim(Ker(\psi^{j-1})) = d_j - d_{j-1};$$

$$r_1 = \dim(Ker(\psi)) = d_1.$$

Dunque poniamo $d_0 = 0$ e $d_{s+1} = d_s = h$, in modo da avere:

$$b_j = r_j - r_{j+1} = d_j - d_{j-1} - (d_{j+1} - d_j) = 2d_j - d_{j-1} - d_{j+1} \quad \forall j.$$

Dunque $\mathfrak{M}_{\mathcal{B}}(\psi)$ dipende, a meno di permutazioni dei blocchi sulla diagonale, solo dai d_i .

- 3) Se ψ_1, ψ_2 sono endomorfismi nilpotenti di V con stessa stringa di invarianti, allora, per quanto visto nei punti precedenti, $\exists \mathcal{B}_1, \mathcal{B}_2$ basi di V tale che $\mathfrak{M}_{\mathcal{B}_1}(\psi_1) = \mathfrak{M}_{\mathcal{B}_2}(\psi_2)$, quindi $\psi_1 \sim \psi_2$.

Osservazione: Dal lemma segue che $\dim(W_j) \leq \dim(W_{j-1})$, cioè $r_j \leq r_{j-1}$, ossia la successione $\{d_j - d_{j-1}\}_j$ è decrescente (in generale non strettamente).

Osservazione: Dato $f \in \text{End}(V)$, $Sp(f) = \{\lambda_1, \dots, \lambda_k\}$. Allora ci restringiamo a lavorare in ogni autospazio generalizzato V'_{λ_j} ; studiando l'endomorfismo di V'_{λ_j} , $f_j = f|_{V'_{\lambda_j}} - \lambda_j id$, troviamo la forma di Jordan $J(f_j)$ di f_j , da cui risaliamo facilmente a quella di $f|_{V'_{\lambda_j}}$, semplicemente aggiungendo a $\lambda_j I$ a $J(f_j)$. Dunque la forma canonica di Jordan dell'endomorfismo f è:

$$J(f) = \begin{pmatrix} J(f_1) + \lambda_1 I & & \\ & \ddots & \\ & & J(f_k) + \lambda_k I \end{pmatrix}$$

Da questo segue:

COROLLARIO 3.5.16 (forma canonica di Jordan per endomorfismi triangolabili): Sia $f \in \text{End}(V)$ triangolabile. Allora:

- 1) $\exists \mathcal{B}$ base di V | $\mathfrak{M}_{\mathcal{B}}(f)$ è una matrice di Jordan (\mathcal{B} è detta base di Jordan per f);
- 2) La matrice $\mathfrak{M}_{\mathcal{B}}(f)$ è unica a meno di permutazioni dei blocchi sulla diagonale ed è determinata dalla stringa di invarianti $s(\lambda_i) = [\lambda_i, s_i, [d_1(\lambda_1) < \dots < d_{s_i}(\lambda_i)]]$ associati agli autovalori di f ;
- 3) Due endomorfismi triangolabili di V sono coniugati \Leftrightarrow hanno la stessa forma canonica di Jordan (che è quindi un invariante completo di coniugio).

COROLLARIO 3.5.17: Ogni matrice triangolabile A è simile alla sua trasposta.

Dimostrazione:

$J(A)$ dipende solo dalle dimensioni dei $Ker(A - \lambda I)^j$ e dunque dai $rk(A - \lambda I)^j$.

Poiché $\forall j, rk({}^t A - \lambda I)^j = rk(A - \lambda I)^j$, allora $J({}^t A) = J(A)$, da cui ${}^t A \sim A$.

Osservazione: Per quanto visto finora siamo in grado di stabilire se due endomorfismi triangolabili sono simili. Dunque la forma canonica di Jordan è un sistema completo di invarianti in $\text{End}(V)$, con V \mathbb{K} -spazio vettoriale e \mathbb{K} algebricamente chiuso.

Però ad esempio in $\mathcal{M}(n, \mathbb{R})$ abbiamo visto che esistono matrici non triangolabili; ovviamente la teoria della forma canonica di Jordan non può essere applicata a tali matrici. Riprendendo però il teorema che dice che $A \sim_{\mathbb{R}} B \Leftrightarrow A \sim_{\mathbb{C}} B$, possiamo lavorare in questo modo:

$$A \sim_{\mathbb{R}} B \Leftrightarrow A \sim_{\mathbb{C}} B \Leftrightarrow J_{\mathbb{C}}(A) = J_{\mathbb{C}}(B)$$

Dunque possiamo considerare $A, B \in \mathcal{M}(n, \mathbb{R}) \subseteq \mathcal{M}(n, \mathbb{C})$, cioè $A, B: \mathbb{C}^n \rightarrow \mathbb{C}^n$, trovare la loro forma di Jordan complessa (che di sicuro esiste) e stabilire se sono simili o meno a partire da questa. In quest'ultima parte del capitolo vogliamo arrivare alla stessa conclusione provando che nella classe di similitudine di A non triangolabile in $\mathcal{M}(n, \mathbb{R})/\sim$ esiste un rappresentante canonico, detto **forma di Jordan reale** di A , univocamente determinata da $J_{\mathbb{C}}(A)$ e quindi essenzialmente unica.

FORMA DI JORDAN REALE: $\forall A \in \mathcal{M}(n, \mathbb{R})$ esiste un rappresentante canonico di similitudine, univocamente determinato dalla forma di Jordan complessa di A .

Dimostrazione:

Sia $A \in \mathcal{M}(n, \mathbb{R})$. Procediamo per passi dimostrando lemmi:

1) Poiché $p_A(t) \in \mathbb{R}[t]$, gli autovalori di A sono del tipo:

$$\underbrace{\lambda_1, \dots, \lambda_k}_{\in \mathbb{R}}, \underbrace{\mu_1, \dots, \mu_r, \bar{\mu}_1, \dots, \bar{\mu}_r}_{\in \mathbb{C} \setminus \mathbb{R}}$$

$$\text{e } \mu_a(\mu_j) = \mu_a(\bar{\mu}_j) \quad \forall j.$$

2) Pensiamo A come endomorfismo di \mathbb{C}^n . Allora per il teorema di Jordan:

$$\mathbb{C}^n = V'_{\lambda_1} \oplus \dots \oplus V'_{\lambda_k} \oplus V'_{\mu_1} \oplus \dots \oplus V'_{\mu_r} \oplus V'_{\bar{\mu}_1} \oplus \dots \oplus V'_{\bar{\mu}_r}$$

3) Sia λ uno degli autovalori reali di A . Una base di Jordan di $V'_\lambda \subseteq \mathbb{C}^n$ si trova prendendo basi opportune nella successione di sottospazi:

$$\text{Ker}(A - \lambda I) \subseteq \text{Ker}(A - \lambda I)^2 \subseteq \dots$$

Poiché $\forall j$ il sottospazio $\text{Ker}(A - \lambda I)^j$ ha la stessa dimensione sia come sottospazio di \mathbb{R}^n sia come sottospazio di \mathbb{C}^n (in quanto è reale), allora possiamo scegliere una base di Jordan di V'_λ formata da vettori reali.

4) Sia $\mu \in \mathbb{C} \setminus \mathbb{R}$ uno degli autovalori complessi non reali di A . Se $\{z_1, \dots, z_t\}$ è una base di Jordan di V'_μ , allora $\{\bar{z}_1, \dots, \bar{z}_t\}$ è una base di Jordan di $V'_{\bar{\mu}}$.

Dimostrazione:

Poiché $\overline{\text{Ker}(A - \mu I)^j} = \text{Ker}(A - \bar{\mu} I)^j \quad \forall j$, in quanto A è reale e quindi $A = \bar{A}$, allora i vettori $\bar{z}_1, \dots, \bar{z}_t \in V'_{\bar{\mu}} = \bigcup_j \text{Ker}(A - \bar{\mu} I)^j$.

Dimostriamo che gli \bar{z}_i sono linearmente indipendenti:

$$a_1 \bar{z}_1 + \dots + a_t \bar{z}_t = 0, \text{ con } a_i \in \mathbb{C} \quad \forall i.$$

Coniugando:

$$\bar{a}_1 z_1 + \dots + \bar{a}_t z_t = 0 \Rightarrow \bar{a}_i = 0 \quad \forall i, \text{ poiché gli } z_i \text{ sono linearmente indipendenti} \Rightarrow a_i = 0 \quad \forall i.$$

Inoltre $\dim(V'_\mu) = \mu_a(\mu) = \mu_a(\bar{\mu}) = \dim(V'_{\bar{\mu}}) = t$, quindi $\{\bar{z}_1, \dots, \bar{z}_t\}$ è una base di $V'_{\bar{\mu}}$.

Ci resta da mostrare che è una base di Jordan di $V'_{\bar{\mu}}$.

Poiché $\{z_1, \dots, z_t\}$ è una base di Jordan di V'_μ , allora:

$$Az_j = \mu z_j \vee Az_j = \mu z_j + z_{j-1} \quad \forall j$$

$$\text{Se } Az_j = \mu z_j \Rightarrow A\bar{z}_j = \overline{Az_j} = \overline{\mu z_j} = \bar{\mu} \bar{z}_j;$$

$$\text{se } Az_j = \mu z_j + z_{j-1} \Rightarrow A\bar{z}_j = \overline{\mu z_j + z_{j-1}} = \bar{\mu} \bar{z}_j + \bar{z}_{j-1},$$

dunque $\{\bar{z}_1, \dots, \bar{z}_t\}$ è effettivamente una base di Jordan di $V'_{\bar{\mu}}$.

(Osservazione: Abbiamo quindi mostrato che, se in $J_{\mathbb{C}}(A)$ ci sono b blocchi relativi all'autovalore μ di ordine m , allora in $J_{\mathbb{C}}(A)$ ci sono b blocchi di ordine m relativi a $\bar{\mu}$.)

5) Sia $\mu \in \mathbb{C} \setminus \mathbb{R}$ uno degli autovalori complessi non reali di A . Allora esiste una base di $V'_\mu \oplus V'_{\bar{\mu}}$ formata da vettori reali.

Dimostrazione:

Sia $\{z_1, \dots, z_t\}$ una base di Jordan di V'_μ . Per 4) sappiamo che $\{\bar{z}_1, \dots, \bar{z}_t\}$ è una base di Jordan di V'_μ . Quindi $\dim(V'_\mu \oplus V'_\mu) = 2t$.

Poniamo $x_j = \Re(z_j)$ e $y_j = \Im(z_j) \quad \forall j$.

Poiché $x_j = \frac{z_j + \bar{z}_j}{2}$ e $y_j = \frac{z_j - \bar{z}_j}{2i}$, i vettori $x_1, y_1, \dots, x_t, y_t \in V'_\mu \oplus V'_\mu$, in quanto combinazioni lineari di z_j e \bar{z}_j . Inoltre, poiché gli z_j e \bar{z}_j generano $V'_\mu \oplus V'_\mu$, allora anche gli x_j, y_j generano, in quanto ciascuno è combinazione lineare di z_j e \bar{z}_j .

Infine, poiché $\dim(V'_\mu \oplus V'_\mu) = 2t$, allora $\{x_1, y_1, \dots, x_t, y_t\}$ è una base di vettori reali di $V'_\mu \oplus V'_\mu$.

6) Scriviamo la matrice associata a $A|_{V'_\mu \oplus V'_\mu}$ rispetto alla base $\{x_1, y_1, \dots, x_t, y_t\}$.

• Se $Az_j = \mu z_j$, allora:

$$Ax_j = A \frac{z_j + \bar{z}_j}{2} = \frac{\mu z_j + \bar{\mu} \bar{z}_j}{2} = \Re(\mu z_j) \stackrel{\square}{=} \Re(\mu) \Re(z_j) - \Im(\mu) \Im(z_j) = \\ = \Re(\mu) x_j - \Im(\mu) y_j$$

$$Ay_j = A \frac{z_j - \bar{z}_j}{2i} = \frac{\mu z_j - \bar{\mu} \bar{z}_j}{2i} = \Im(\mu z_j) \stackrel{\square}{=} \Im(\mu) \Re(z_j) + \Re(\mu) \Im(z_j) = \\ = \Im(\mu) x_j + \Re(\mu) y_j$$

I passaggi contrassegnati con \square derivano dalle relazioni:

$$z, w \in \mathbb{C}, \quad \Re(zw) = \Re(z) \Re(w) - \Im(z) \Im(w);$$

$$\Im(zw) = \Im(z) \Re(w) + \Re(z) \Im(w).$$

• Analogamente, se $Az_j = \mu z_j + z_{j-1}$:

$$Ax_j = A \frac{z_j + \bar{z}_j}{2} = \frac{\mu z_j + z_{j-1} + \bar{\mu} \bar{z}_j + \bar{z}_{j-1}}{2} = \Re(\mu z_j) + \Re(z_{j-1}) = \\ = \Re(\mu) x_j - \Im(\mu) y_j + x_{j-1}$$

$$Ay_j = \Im(\mu) x_j + \Re(\mu) y_j + y_{j-1}$$

Pertanto, se la matrice associata ad $A|_{V'_\mu}$ rispetto alla base $\{z_1, \dots, z_t\}$ era:

$$\begin{pmatrix} \boxed{J_1} & & & & \\ & \ddots & & & \\ & & \boxed{J_r} & & \end{pmatrix}$$

con J_i blocco di Jordan di ordine m_i relativo a μ , allora la matrice associata a $A|_{V'_\mu \oplus V'_\mu}$ rispetto alla base $\{x_1, y_1, \dots, x_t, y_t\}$ è una matrice di ordine $2t$ e del tipo:

$$\begin{pmatrix} \boxed{\tilde{J}_1} & & & & \\ & \ddots & & & \\ & & \boxed{\tilde{J}_r} & & \end{pmatrix}$$

dove \tilde{J}_i è un blocco di ordine $2m_i$ della forma:

$$\begin{pmatrix} H_\mu & I & & & \\ & H_\mu & I & & \\ & & \ddots & & \\ & & & H_\mu & I \\ & & & & H_\mu \end{pmatrix}$$

$$\text{con } H_\mu = \begin{pmatrix} \Re(\mu) & \Im(\mu) \\ -\Im(\mu) & \Re(\mu) \end{pmatrix} \text{ e } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Osservazione: Se per la coppia di autovalori μ e $\bar{\mu}$ si fosse scelto di lavorare con $\bar{\mu}$, al posto del blocchetto $\begin{pmatrix} \Re(\mu) & \Im(\mu) \\ -\Im(\mu) & \Re(\mu) \end{pmatrix}$ avremmo avuto il blocchetto $\begin{pmatrix} \Re(\mu) & -\Im(\mu) \\ \Im(\mu) & \Re(\mu) \end{pmatrix}$, simile ma non uguale a quello associato a μ .

Per evitare ambiguità nella forma $J_{\mathbb{R}}(A)$ conveniamo di scegliere l'autovalore $|\Im(\mu)| > 0$.

A questo punto la forma di Jordan reale di A è unica a meno di permutazioni dei blocchetti.

Esempio: Trovare la forma di Jordan reale di $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

$$p_A(t) = \det \begin{pmatrix} -t & 0 & 1 \\ 1 & -t & 0 \\ 0 & 1 & -t \end{pmatrix} = -t(t^2) + 1(1) = -t^3 + 1 = (1-t)(t^2 + t + 1).$$

Poiché su \mathbb{C} il polinomio caratteristico ha tre radici distinte, $1, \mu = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \bar{\mu} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$, allora su \mathbb{C} A è diagonalizzabile:

$$J_{\mathbb{C}}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \bar{\mu} \end{pmatrix}$$

Per quanto visto prima, sostituiamo i due blocchetti (μ) e $(\bar{\mu})$ con il blocchetto

$$\begin{pmatrix} \Re(\mu) & \Im(\mu) \\ -\Im(\mu) & \Re(\mu) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}. \text{ Dunque:}$$

$$J_{\mathbb{R}}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

PROPOSIZIONE 3.5.18: Siano $\mathbb{K} \subseteq \mathbb{K}'$ campi. $A \in \mathcal{M}(n, \mathbb{K})$. Allora $m_{\mathbb{K}}(A) = m_{\mathbb{K}'}(A)$ (dove $m_{\mathbb{F}}(A)$ è il polinomio minimo di A su \mathbb{F}).

Dimostrazione:

Lavoriamo in $\mathbb{K}'[t]$. In questo anello, poiché $m_{\mathbb{K}'}(A)$ genera l'ideale dei polinomi che si annullano in A , e $m_{\mathbb{K}}(A)$ appartiene a questo ideale, ho che $m_{\mathbb{K}'}(A) | m_{\mathbb{K}}(A)$.

Per concludere mi basta mostrare che $\deg(m_{\mathbb{K}}(A)) \leq \deg(m_{\mathbb{K}'}(A))$, poiché, essendo sicuramente $\deg(m_{\mathbb{K}}(A)) \geq \deg(m_{\mathbb{K}'}(A))$, avrei che $\deg(m_{\mathbb{K}}(A)) = \deg(m_{\mathbb{K}'}(A))$, ma essendo i due polinomi monici, avrei la tesi.

Notiamo innanzitutto che, dati $v_1, \dots, v_k \in \mathbb{K}^n \subseteq (\mathbb{K}')^n$, sono linearmente indipendenti su $\mathbb{K}^n \Leftrightarrow$ lo sono su $(\mathbb{K}')^n$, infatti entrambe le condizioni sono equivalenti alla condizione:

“Se M è la matrice $n \times k$ avente i v_i come colonne, \exists minore $k \times k$ di M con determinante $\neq 0$ ”.

Mostriamo ora che, date $A_1, \dots, A_k \in \mathcal{M}(n, \mathbb{K})$, con $k \leq n^2$, sono linearmente indipendenti su $\mathbb{K} \Leftrightarrow$ lo sono su \mathbb{K}' .

Infatti, considerando che $\mathcal{M}(n, \mathbb{K}') \cong (\mathbb{K}')^{n^2}$, per la precedente osservazione se esiste una combinazione lineare non nulla su $\mathcal{M}(n, \mathbb{K}')$ delle A_i che dà 0, allora esiste anche su $(\mathbb{K}')^{n^2}$, quindi anche su \mathbb{K}^{n^2} , quindi su $\mathcal{M}(n, \mathbb{K})$.

Inoltre il viceversa è ovvio, in quanto se le A_i sono indipendenti su \mathbb{K}' lo sono anche su \mathbb{K} .

A questo punto sia $d = \deg(m_{\mathbb{K}'}(A))$. Allora I, A, \dots, A^d sono linearmente dipendenti su \mathbb{K}' per definizione di polinomio minimo, ma per quanto visto I, A, \dots, A^d sono linearmente dipendenti anche su \mathbb{K} , quindi \exists polinomio $q(t) \in \mathbb{K}[t], q \neq 0, \deg(q) \leq d$ tale che $q(A) = 0$. Per cui $\deg(m_{\mathbb{K}}(A)) \leq d$, da cui la tesi.

Osservazione: Quindi $m_{\mathbb{R}}(A) = m_{\mathbb{C}}(A) \forall A \in \mathcal{M}(n, \mathbb{R})$. Inoltre è evidente che $p_{\mathbb{R}}(A) = p_{\mathbb{C}}(A)$, poiché il polinomio caratteristico, essendo un determinante, dipende solo da A .

3.6 BASI CICLICHE PER ENDOMORFISMI

Nel seguito sia V spazio vettoriale su \mathbb{K} , $\text{char}(\mathbb{K}) = 0$.

DEFINIZIONE 3.6.1: $\dim(V) = n, f \in \text{End}(V)$. Una base \mathcal{B} di V si dice **ciclica** per f se $\exists v \in V$ tale che $\mathcal{B} = \{v, f(v), \dots, f^{n-1}(v)\}$.

DEFINIZIONE 3.6.2: $v \in V. I(f, v) = \{q \in \mathbb{K}[t] | q(f)(v) = 0\}$ è un ideale di $\mathbb{K}[t]$, dunque esiste un generatore $m_{f,v}$ di $I(f, v)$, detto **polinomio minimo di v rispetto a f** .

Osservazione: Visto che $I_f \subseteq I(f, v)$, si ha che $m_{f,v} | m_f$.

LEMMA 3.6.1: \mathbb{K} campo con $\text{char}(\mathbb{K}) = 0$. Sia V un \mathbb{K} -spazio vettoriale e siano W_1, \dots, W_n sottospazi di V tali che $V = W_1 \cup \dots \cup W_n$. Allora $\exists i$ tale che $V = W_i$.

Dimostrazione:

Supponiamo che l'unione $V = W_1 \cup \dots \cup W_n$ sia minimale, cioè $W_j \not\subseteq W_1 \cup \dots \cup W_{j-1} \cup W_{j+1} \cup \dots \cup W_n \forall j$. Supponiamo per assurdo che $n \geq 2$; allora per minimalità $W_n \not\subseteq W_1 \cup \dots \cup W_{n-1}$.

Sia $u \notin W_n$ e $v \in W_n \setminus (W_1 \cup \dots \cup W_{n-1})$ e denotiamo $S = \{v + tu | t \in \mathbb{K}\}$.

$u \neq 0$ e $\#\mathbb{K} = +\infty$, dunque $\#S = +\infty$; visto che $S \subseteq V = W_1 \cup \dots \cup W_n$, dovrà esistere un W_i tale che $\#(S \cap W_i) = +\infty$. Vediamo che ciò è assurdo.

Se $v + tu \in W_n$ per $t \neq 0$, si avrebbe $W_n \ni (v + tu) - v = tu$, cioè $u \in W_n$, assurdo.

Se invece $v + t_1 u, v + t_2 u \in W_i$ con $t_1 \neq t_2$ e $i < n$, si avrebbe $(t_2 - t_1)v = t_2(v + t_1 u) - t_1(v + t_2 u) \in W_i$, cioè $v \in W_i$, assurdo.

Dunque $\#(S \cap W_i) \leq 1 \forall i$, da cui l'assurdo e la tesi.

Osservazione: La precedente proposizione è falsa se $\text{char}(\mathbb{K}) > 0$; non è infatti difficile trovare un controesempio con $\mathbb{K} = \mathbb{F}_2$.

LEMMA 3.6.2: $f \in \text{End}(V)$. Allora $\exists v \in V$ tale che $m_{f,v} = m_f$.

Dimostrazione:

Visto che $\forall v \in V m_{f,v} | m_f$, l'insieme $\{m_{f,v} | v \in V\}$ è finito, pertanto coincide con

$\{m_{f,v_1}, \dots, m_{f,v_p}\}$ per alcuni $v_i \in V$.

$\forall j = 1, \dots, p$ considero i sottospazi $W_j = \text{Ker}(m_{f,v_j}(f)) = \{z \in V | m_{f,v_j}(f)(z) = 0\}$.

$\forall z \in V, \exists j$ tale che $m_z = m_{v_j}$ e quindi $m_{v_j}(f)(z) = m_z(f)(z) = 0$, cioè $z \in W_j$.

Allora $V = W_1 \cup \dots \cup W_p$ e dunque $\exists i_0$ tale che $W_{i_0} = V$, ossia $\text{Ker}(m_{v_{i_0}}(f)) = V$, cioè $m_{v_{i_0}} \in I_f$ quindi $m_f | m_{v_{i_0}}$.

TEOREMA 3.6.3: f ammette una base ciclica $\Leftrightarrow m_f = \pm p_f$.

Dimostrazione:

\Rightarrow) Ovvvia, in quanto se $\{v, f(v), \dots, f^{n-1}(v)\}$ sono linearmente indipendenti, $\deg(m_f) \geq n$ e dunque $m_f = \pm p_f$.

\Leftarrow) Per il lemma $\exists v \in V$ tale che $m_{f,v} = m_f$. Per ipotesi $\deg(m_{f,v}) = \deg(m_f) = n$.

Ma allora $\{v, f(v), \dots, f^{n-1}(v)\}$ sono linearmente indipendenti; infatti se $b_0v + b_1f(v) + \dots + b_{n-1}f^{n-1}(v) = 0$, il polinomio $g(t) = b_0 + b_1t + \dots + b_{n-1}t^{n-1}$ deve soddisfare $g(t) \in I(f, v)$, ma $m_{f,v} = m_f$, dunque tutti i polinomi in $I(f, v)$ hanno grado $\geq n$, quindi $g(t) \equiv 0$ e $b_i = 0 \forall i$.

4 FORME BILINEARI

4.1 FORME BILINEARI E FORME QUADRATICHE

DEFINIZIONE 4.1.1: Sia V un \mathbb{K} -spazio vettoriale. $\phi: V \times V \rightarrow \mathbb{K}$ è detta **applicazione** (o **forma**) **bilineare** se:

- 1) $\forall x, y, z \in V, \phi(x + y, z) = \phi(x, z) + \phi(y, z)$
- 2) $\forall x, y, z \in V, \phi(x, y + z) = \phi(x, y) + \phi(x, z)$
- 3) $\forall x, y \in V, \forall \alpha \in \mathbb{K}, \phi(\alpha x, y) = \alpha\phi(x, y) = \phi(x, \alpha y)$.

PROPOSIZIONE 4.1.1: Le seguenti applicazioni sono bilineari:

- 1) $\phi \equiv 0$;
- 2) $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \mid \phi(X, Y) = {}^tXY = \sum_{i=1}^n x_i y_i$, detto **prodotto scalare standard su \mathbb{R}^n** ;
- 3) $A \in \mathcal{M}(n, \mathbb{K}), \phi: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K} \mid \phi(X, Y) = {}^tXAY$;
- 4) $\phi: \mathcal{M}(n, \mathbb{K}) \times \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K} \mid \phi(A, B) = \text{tr}({}^tAB)$;
- 5) $\phi: \mathcal{M}(n, \mathbb{K}) \times \mathcal{M}(n, \mathbb{K}) \rightarrow \mathbb{K} \mid \phi(A, B) = \text{tr}(AB)$;
- 6) $a_1, \dots, a_r \in \mathbb{K}, \phi: \mathbb{K}_n[x] \times \mathbb{K}_n[x] \rightarrow \mathbb{K} \mid \phi(p(x), q(x)) = \sum_{i=1}^r p(a_i)q(a_i)$, cioè la valutazione ($\mathbb{K}_n[x]$ è l'insieme dei polinomi di grado $\leq n$);

$$7) \phi: \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R} \mid \phi \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4, \text{ detto } \mathbf{prodotto scalare di}$$

Minkowski.

Dimostrazione:

- 1) Ovvvia.
- 2) Segue immediatamente dal fatto che la trasposizione e il prodotto fra matrici sono lineari.
- 3) Analoga alla 2).
- 4) La traccia, così come la trasposizione e il prodotto fra matrici, è lineare.
- 5) Analoga alla 4).
- 6) Segue dal fatto che la valutazione è lineare.
- 7) Lasciata al lettore.

DEFINIZIONE 4.1.2: Una forma bilineare $\phi: V \times V \rightarrow \mathbb{K}$ si dice **prodotto scalare** se è **simmetrica**, cioè $\phi(v, w) = \phi(w, v) \forall v, w \in V$.

PROPOSIZIONE 4.1.2: Delle precedenti forme bilineari, 1), 2), 4), 5), 6), 7) sono prodotti scalari, mentre la 3) è un prodotto scalare $\Leftrightarrow A$ è simmetrica.

Dimostrazione:

La 3) è un prodotto scalare $\Leftrightarrow {}^tXAY = {}^tYAX = {}^t({}^tX{}^tAY) \forall X, Y \in \mathbb{K}^n$, ma essendo numeri, ${}^tXAY = {}^t({}^tX{}^tAY) \forall X, Y \in \mathbb{K}^n \Leftrightarrow {}^tXAY = {}^tX{}^tAY \forall X, Y \in \mathbb{K}^n$. Poiché l'uguaglianza deve valere $\forall X, Y \in \mathbb{K}^n$, in particolare varrà per $X = e_i, Y = e_j$, con $1 \leq i, j \leq n$.

Quindi ${}^t e_i A e_j = {}^t e_i {}^t A e_j \forall i, j \Leftrightarrow [A]_{ij} = [{}^t A]_{ij} \forall i, j \Leftrightarrow A = {}^t A$.

La 4) è un prodotto scalare perché $tr({}^tAB) = tr({}^t({}^tAB)) = tr({}^tBA)$, mentre la 5) è un prodotto scalare perché abbiamo dimostrato che $tr(AB) = tr(BA)$. Le altre sono verifiche immediate.

Osservazione: Durante tutta la trattazione delle forme bilineari lavoreremo solo in campi \mathbb{K} con $char(\mathbb{K}) \neq 2$, per poter dividere per 2.

DEFINIZIONE 4.1.3: Sia $\phi: V \times V \rightarrow \mathbb{K}$ una forma bilineare. Si definisce **forma quadratica indotta da ϕ** l'applicazione $q_\phi: V \rightarrow \mathbb{K}$ definita da $q_\phi(v) = \phi(v, v) \quad \forall v \in V$.

Esempio: $\phi(X, Y) = {}^tXY$ su \mathbb{K}^n induce $q_\phi(x) = {}^tXX = \sum_{i=1}^n x_i^2$ (detta **forma quadratica standard**).

DEFINIZIONE 4.1.4: Una applicazione $q: V \rightarrow \mathbb{K}$ si dice **forma quadratica** se $\exists \phi$ forma bilineare su $V \mid q = q_\phi$.

Osservazione: Più forme bilineari possono definire la stessa forma quadratica, ad esempio $\phi_1 \equiv 0$ e $\phi_2(X, Y) = {}^tXAY$, con A antisimmetrica, inducono la stessa $q_\phi = 0$, in quanto: ${}^tXAX = {}^t({}^tX{}^tAX) \equiv {}^tX(-A)X = -{}^tXAX \Rightarrow {}^tXAX = 0$ (il passaggio \equiv deriva dal fatto che la trasposizione lascia invariato un numero).

PROPOSIZIONE 4.1.3: Sia $q: V \rightarrow \mathbb{K}$ una forma quadratica. Allora esiste uno e un solo prodotto scalare che induce q .

Dimostrazione:

Per definizione $\exists \phi$ forma bilineare tale che $q = q_\phi$, cioè tale che $q(v) = \phi(v, v) \quad \forall v \in V$.

Allora $\phi'(u, v) = \frac{\phi(u, v) + \phi(v, u)}{2}$ è un prodotto scalare che induce q (in quanto è evidentemente simmetrico e $\phi'(u, u) = \phi(u, u) = q(u)$).

Inoltre se ψ è un prodotto scalare che induce q , allora:

$$q(u + v) - q(u) - q(v) = \psi(u + v, u + v) - \psi(u, u) - \psi(v, v) = \psi(u, v) + \psi(v, u) = 2\psi(u, v)$$

da cui $\psi(u, v) = \frac{q(u+v) - q(u) - q(v)}{2}$ (detta formula di polarizzazione), dunque ψ è univocamente determinato e quindi unico.

DEFINIZIONE 4.1.5: Definiamo:

- $Bil(V) = \{\phi: V \times V \rightarrow \mathbb{K} \mid \phi \text{ è bilineare}\}$
- $PS(V) = \{\phi: V \times V \rightarrow \mathbb{K} \mid \phi \text{ è prodotto scalare}\}$
- $Q(V) = \{q: V \rightarrow \mathbb{K} \mid q \text{ è forma quadratica}\}$.

DEFINIZIONE 4.1.6: Definiamo una somma e un prodotto per scalari in $Bil(V)$:

- $\forall \phi, \psi \in Bil(V), (\phi + \psi)(v, w) \stackrel{\text{def}}{=} \phi(v, w) + \psi(v, w)$;
- $\forall \lambda \in \mathbb{K}, \forall \phi \in Bil(V), (\lambda\phi)(v, w) \stackrel{\text{def}}{=} \lambda\phi(v, w)$.

PROPOSIZIONE 4.1.4: 1) $Bil(V)$ è uno spazio vettoriale;

2) $PS(V)$ è un sottospazio vettoriale di $Bil(V)$;

3) $Q(V)$ è un sottospazio vettoriale di $\mathcal{F}(V, \mathbb{K}) = \{f: V \rightarrow \mathbb{K}\}$.

Osservazione: $PS(V) \cong Q(V)$, in quanto l'applicazione $F: PS(V) \rightarrow Q(V) \mid F(\phi) = q_\phi$ è un isomorfismo (sappiamo che è bigettiva e si vede con un'immediata verifica che è lineare).

Notazione: Indicheremo con (V, ϕ) lo spazio vettoriale V dotato del prodotto scalare ϕ .

DEFINIZIONE 4.1.7: Siano (V, ϕ) e (W, ψ) \mathbb{K} -spazi vettoriali. $f: V \rightarrow W$ lineare si dice **isometria** se:

- f è isomorfismo di spazi vettoriali;
- $\forall x, y \in V, \phi(x, y) = \psi(f(x), f(y))$.

DEFINIZIONE 4.1.8: (V, ϕ) e (W, ψ) si dicono **isometrici** se $\exists f: V \rightarrow W$ isometria (in tal caso ϕ e ψ si dicono **prodotti scalari isometrici**).

Osservazione: L'essere isometrici è una relazione di equivalenza (la verifica è lasciata al lettore).

Osservazione: Se $f: V \rightarrow (W, \psi)$ è un isomorfismo, l'applicazione $f_\psi^*: V \times V \rightarrow \mathbb{K} \mid f_\psi^*(v, w) = \psi(f(v), f(w))$ è evidentemente un prodotto scalare su V e $f: (V, f_\psi^*) \rightarrow (W, \psi)$ è un'isometria per costruzione.

Dunque in generale basterà studiare $\{(V, \phi)\} / \text{isometrie}$, con V fissato.

Denoteremo con \langle, \rangle il prodotto scalare standard di \mathbb{R}^n .

Esempi:

- 1) Le rotazioni di centro l'origine sono isometrie lineari di \mathbb{R}^2 dotato del prodotto scalare standard, infatti sia $\begin{pmatrix} x \\ y \end{pmatrix}$ un vettore di \mathbb{R}^2 e sia $\begin{pmatrix} x' \\ y' \end{pmatrix}$ il vettore ottenuto ruotando l'altro di un angolo α ; allora:

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos(\alpha) - y \sin(\alpha) \\ x \sin(\alpha) + y \cos(\alpha) \end{pmatrix}; \\ \left\langle \begin{pmatrix} x_1 \cos(\alpha) - y_1 \sin(\alpha) \\ x_1 \sin(\alpha) + y_1 \cos(\alpha) \end{pmatrix}, \begin{pmatrix} x_2 \cos(\alpha) - y_2 \sin(\alpha) \\ x_2 \sin(\alpha) + y_2 \cos(\alpha) \end{pmatrix} \right\rangle &= \\ &= x_1 x_2 (\cos^2(\alpha) + \sin^2(\alpha)) + y_1 y_2 (\sin^2(\alpha) + \cos^2(\alpha)) = x_1 x_2 + y_1 y_2 = \\ &= \left\langle \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right\rangle \end{aligned}$$

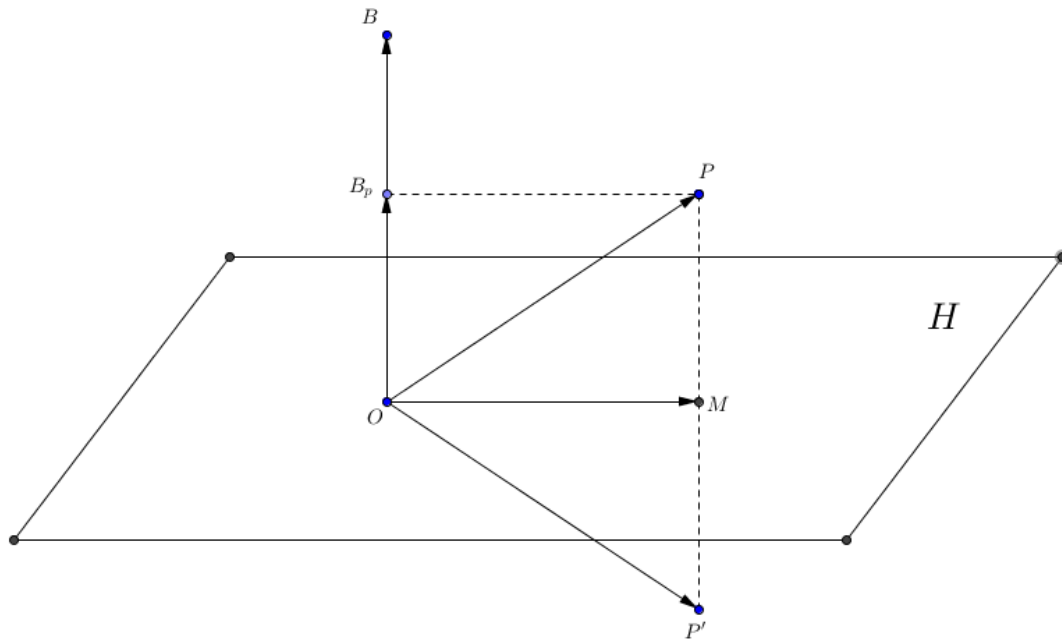
- 2) Sia $H = \{b_1 x_1 + \dots + b_n x_n = 0\}$ un iperpiano in $(\mathbb{R}^n, \langle, \rangle)$ passante per O .

Sia $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Allora $H = \{X \mid \langle B, X \rangle = 0\}$.

Sia ora $\rho: \mathbb{R}^n \rightarrow \mathbb{R}^n$ la riflessione ortogonale rispetto ad H che associa ad ogni $P \in \mathbb{R}^n$ il simmetrico rispetto ad H .

Grazie alla figura (che è in \mathbb{R}^3 ma che comunque rende l'idea della situazione), vediamo che: $\rho(P) = P - 2B_P$; ma avendo l'iperpiano dimensione $n - 1$, su di esso giaceranno $n - 1$ componenti di P , che chiamiamo M_1, \dots, M_{n-1} , mentre l'ultima componente sarà esattamente B_P . Allora:

$$\langle P, B \rangle = \langle B_P + \sum_{i=1}^{n-1} M_i, B \rangle = \langle B_P, B \rangle + \sum_{i=1}^{n-1} \langle M_i, B \rangle = \langle B_P, B \rangle.$$



Sia ora $B_p = k \cdot B$; quindi:

$$\langle P, B \rangle = \langle B_p, B \rangle = k \langle B, B \rangle \Rightarrow k = \frac{\langle P, B \rangle}{\langle B, B \rangle}.$$

Dunque $\rho(P) = P - 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B$.

Notiamo che $\rho^2 = id$, infatti:

$$\begin{aligned} \rho(\rho(P)) &= P - 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B - 2 \frac{\langle P - 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B, B \rangle}{\langle B, B \rangle} B = \\ &= P - 2 \frac{B}{\langle B, B \rangle} \left(\langle P, B \rangle + \langle P, B \rangle - 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} \langle B, B \rangle \right) = P \end{aligned}$$

Inoltre ρ è un'isometria, in quanto:

- è lineare (la verifica è immediata);
- è iniettiva, poiché:

$$\begin{aligned} Ker(\rho) &= \left\{ P \mid P = 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B \right\} = \left\{ P \mid P = 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B = 2 \frac{\langle 2 \frac{\langle P, B \rangle}{\langle B, B \rangle} B, B \rangle}{\langle B, B \rangle} B \right\} = \\ &= \{ P \mid \langle P, B \rangle = 2 \langle P, B \rangle \} = \{0\} \end{aligned}$$

- è surgettiva, poiché la controimmagine di P è $\rho(P)$, in quanto $\rho^2 = id$;
- mantiene i prodotti scalari, in quanto:

$$\begin{aligned} \langle \rho(X), \rho(Y) \rangle &= \left\langle X - 2 \frac{\langle X, B \rangle}{\langle B, B \rangle} B, Y - 2 \frac{\langle Y, B \rangle}{\langle B, B \rangle} B \right\rangle \\ &= \langle X, Y \rangle - 2 \frac{\langle X, B \rangle \langle Y, B \rangle}{\langle B, B \rangle} - 2 \frac{\langle X, B \rangle \langle Y, B \rangle}{\langle B, B \rangle} + 4 \frac{\langle X, B \rangle \langle Y, B \rangle \langle B, B \rangle}{\langle B, B \rangle \langle B, B \rangle} = \langle X, Y \rangle. \end{aligned}$$

DEFINIZIONE 4.1.9: Si definisce **gruppo ortogonale** di (V, ϕ) l'insieme $\{f \in GL(V) \mid f: (V, \phi) \rightarrow (V, \phi) \text{ è isometria}\}$, cioè:
 $f \in O(V, \phi) \Leftrightarrow \phi(x, y) = \phi(f(x), f(y)) \quad \forall x, y \in V$.

$$O(V, \phi) =$$

PROPOSIZIONE 4.1.5: $(O(V, \phi), \circ)$ è un gruppo ed è sottogruppo di $GL(V)$.

DEFINIZIONE 4.1.10: $x, y \in V$ si dicono **ortogonali** rispetto a ϕ se $\phi(x, y) = 0$.

Osservazione: Se x, y sono ortogonali rispetto a ϕ e $f \in O(V, \phi)$, allora $f(x), f(y)$ sono ortogonali rispetto a ϕ .

DEFINIZIONE 4.1.11: $\phi \in Bil(V)$, $\mathcal{B} = \{v_1, \dots, v_n\}$ base di V . Si definisce **matrice associata** a ϕ rispetto a \mathcal{B} la matrice $\mathfrak{M}_{\mathcal{B}}(\phi) \in \mathcal{M}(n, \mathbb{K})$ definita da $[\mathfrak{M}_{\mathcal{B}}(\phi)]_{ij} = \phi(v_i, v_j)$.

PROPOSIZIONE 4.1.6: Se $\mathfrak{M}_{\mathcal{B}}(\phi) = A$, allora $\forall v, w \in V$, $\phi(v, w) = {}^t[v]_{\mathcal{B}}A[w]_{\mathcal{B}}$.

Dimostrazione:

$$\phi(v, w) = \phi\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \phi(v_i, v_j) = {}^t[v]_{\mathcal{B}}A[w]_{\mathcal{B}}.$$

PROPOSIZIONE 4.1.7: Sia \mathcal{B} base di V . Allora l'applicazione:

$$\mathfrak{M}_{\mathcal{B}}: \begin{array}{l} Bil(V) \rightarrow \mathcal{M}(n, \mathbb{K}) \\ \phi \rightarrow \mathfrak{M}_{\mathcal{B}}(\phi) \end{array}$$

è un isomorfismo di spazi vettoriali.

Dimostrazione:

$\mathfrak{M}_{\mathcal{B}}$ è evidentemente una bigezione, poiché $\forall A \in \mathcal{M}(n, \mathbb{K})$, $\phi(v_i, v_j) = {}^t[v]_{\mathcal{B}}A[w]_{\mathcal{B}}$, dove $\mathcal{B} = \{v_1, \dots, v_n\}$ è fissata. Essendo ϕ definita su vettori di base, è univoca l'estensione a tutti i $v \in V$. Inoltre dalla definizione segue la linearità, dunque ho la tesi.

COROLLARIO 4.1.8: $\dim(Bil(V)) = n^2$.

Osservazione: $\phi \in Bil(V)$, \mathcal{B} base di V . Allora $\phi \in PS(V) \Leftrightarrow \mathfrak{M}_{\mathcal{B}}(\phi)$ è simmetrica.

COROLLARIO 4.1.9: $\dim(PS(V)) = \dim(Q(V)) = \frac{n(n+1)}{2}$.

Dimostrazione:

$\mathfrak{M}_{\mathcal{B}}|_{PS(V)}: PS(V) \rightarrow \mathcal{S}(n, \mathbb{K})$ è un isomorfismo.

Osservazione: \mathcal{B} base di V , $\phi \in PS(V)$, $A = \mathfrak{M}_{\mathcal{B}}(\phi)$. Sia ψ_A il prodotto scalare su \mathbb{K}^n associato alla matrice simmetrica A , cioè $\psi_A(X, Y) = {}^tXAY$.

Allora $[\]_{\mathcal{B}}: (V, \phi) \rightarrow (\mathbb{K}^n, \psi_A)$ è un'isometria, infatti $\phi(v, w) = {}^t[v]_{\mathcal{B}}A[w]_{\mathcal{B}} = \psi_A([v]_{\mathcal{B}}, [w]_{\mathcal{B}})$.

PROPOSIZIONE 4.1.10: $f: (V, \phi) \rightarrow (W, \psi)$ isomorfismo, \mathcal{B} base di V , \mathcal{S} base di W .

Siano $M = \mathfrak{M}_{\mathcal{B}}(\phi)$, $N = \mathfrak{M}_{\mathcal{S}}(\psi)$, $A = \mathfrak{M}_{\mathcal{B}, \mathcal{S}}(f)$. Allora f è isometria $\Leftrightarrow M = {}^tANA$.

Dimostrazione:

Siano $X = [v]_{\mathcal{B}}$, $Y = [w]_{\mathcal{B}}$. Allora $\phi(v, w) = {}^tXMY$.

$$\psi(f(v), f(w)) = {}^t[f(v)]_{\mathcal{S}}N[f(w)]_{\mathcal{S}} = {}^t(AX)N(AY) = {}^tX{}^tANAY.$$

Dunque f è isometria $\Leftrightarrow {}^tXMY = {}^tX{}^tANAY \ \forall X, Y \in \mathbb{K}^n \Leftrightarrow M = {}^tANA$ (per il solito discorso che quell'uguaglianza deve valere $\forall X, Y$ e quindi in particolare per gli $X = e_i, Y = e_j$).

Osservazione: f isomorfismo. $f \in O(V, \phi) \Leftrightarrow M = {}^tAMA$, dove $M = \mathfrak{M}_{\mathcal{B}}(\phi)$ e $A = \mathfrak{M}_{\mathcal{B}}(f)$.

4.2 CONGRUENZA E DECOMPOSIZIONE DI WITT

DEFINIZIONE 4.2.1: $A, B \in \mathcal{M}(n, \mathbb{K})$ si dicono **congruenti** se $\exists M \in GL(n, \mathbb{K}) \mid B = {}^tMAM$.

PROPOSIZIONE 4.2.1: Sia $\phi \in Bil(V)$ e $\mathcal{B}, \mathcal{B}'$ basi di V . Poniamo $A = \mathfrak{M}_{\mathcal{B}}(\phi)$ e $A' = \mathfrak{M}_{\mathcal{B}'}(\phi)$. Allora A e A' sono congruenti.

Dimostrazione:

Sia M la matrice del cambiamento di base da \mathcal{B} a \mathcal{B}' ; quindi $[v]_{\mathcal{B}} = M[v]_{\mathcal{B}'} \quad \forall v \in V$.

Allora:

$$\phi(u, v) = {}^t[u]_{\mathcal{B}}A[v]_{\mathcal{B}} = {}^t[u]_{\mathcal{B}'}{}^tMAM[v]_{\mathcal{B}'};$$

$$\phi(u, v) = {}^t[u]_{\mathcal{B}'}A'[v]_{\mathcal{B}'}.$$

Poiché ${}^t[u]_{\mathcal{B}'}{}^tMAM[v]_{\mathcal{B}'} = {}^t[u]_{\mathcal{B}'}A'[v]_{\mathcal{B}'} \quad \forall u, v \in V$, allora $A' = {}^tMAM$.

Osservazione: La congruenza è una relazione di equivalenza

Osservazione: Il rango è un invariante di congruenza (poiché si moltiplica per matrici invertibili).

DEFINIZIONE 4.2.2: $rk(\phi) = rk(\mathfrak{M}_{\mathcal{B}}(\phi))$.

Osservazione: La definizione è ben posta perché il rango non dipende da \mathcal{B} .

PROPOSIZIONE 4.2.2: Sia V un \mathbb{K} -spazio vettoriale, $\phi, \psi \in PS(V)$. Sono fatti equivalenti:

- 1) (V, ϕ) e (V, ψ) sono isometrici;
- 2) $\forall \mathcal{B}$ base di V , $\mathfrak{M}_{\mathcal{B}}(\phi)$ e $\mathfrak{M}_{\mathcal{B}}(\psi)$ sono congruenti;
- 3) $\exists \mathcal{B}, \mathcal{B}'$ basi di $V \mid \mathfrak{M}_{\mathcal{B}}(\phi) = \mathfrak{M}_{\mathcal{B}'}(\psi)$.

Dimostrazione:

Analoga a quella per gli endomorfismi.

Osservazione: Gli invarianti rispetto all'isometria in $PS(V)$ corrispondono agli invarianti rispetto alla congruenza in $\mathcal{S}(n, \mathbb{K})$.

Osservazione: Se $B = {}^tMAM$, con $M \in GL(n, \mathbb{K})$, allora $\det(B) = \det(A) \cdot (\det(M))^2$. Quindi:

- il determinante non è invariante di congruenza;
- se $\mathbb{K} = \mathbb{R}$ il segno del determinante è invariante per congruenza.

DEFINIZIONE 4.2.3: Sia $\phi \in PS(V)$. Si definisce **radicale** di ϕ l'insieme $Rad(\phi) = \{v \in V \mid \phi(v, w) = 0 \quad \forall w \in V\}$.

PROPOSIZIONE 4.2.3: $Rad(\phi)$ è sottospazio di V e $\dim(Rad(\phi)) = \dim(V) - rk(\phi)$.

Dimostrazione:

La verifica che sia sottospazio è lasciata.

Sia $\dim(V) = n$ e \mathcal{B} base di V ; sia $A = \mathfrak{M}_{\mathcal{B}}(\phi)$.

Poniamo $X = [v]_{\mathcal{B}}$ e $Y = [w]_{\mathcal{B}}$.

$$v \in \text{Rad}(\phi) \Leftrightarrow \phi(v, w) = 0 \quad \forall w \in V \Leftrightarrow {}^t XAY = 0 \quad \forall Y \in \mathbb{K}^n \Leftrightarrow {}^t XA = 0 \Leftrightarrow {}^t({}^t XA) = 0 \\ \Leftrightarrow AX = 0$$

dove il passaggio $\boxed{\Leftrightarrow}$ segue dal fatto che deve essere ${}^t XAY = 0 \quad \forall Y \in \mathbb{K}^n$, dunque in particolare per $Y = e_i$, con $1 \leq i \leq n$.

Quindi l'immagine di $\text{Rad}(\phi)$ tramite l'isomorfismo $[\]_{\mathcal{B}}: V \rightarrow \mathbb{K}^n$ è $\text{Ker}(A)$, per cui $\dim(\text{Rad}(\phi)) = \dim(\text{Ker}(A)) = n - \text{rk}(A) = n - \text{rk}(\phi)$.

Osservazione: In particolare $\text{Rad}(\phi)$ si calcola risolvendo il sistema lineare $AX = 0$.

DEFINIZIONE 4.2.4: Diciamo che $\phi \in \text{PS}(V)$ è **non degenere** se $\text{Rad}(\phi) = \{0\}$, ossia se $\phi(v, w) = 0 \quad \forall w \in V \Rightarrow v = 0$ (e **degenere** altrimenti).

COROLLARIO 4.2.4: ϕ è non degenere $\Leftrightarrow \text{rk}(\phi) = \dim(V)$.

Osservazione: Se $A = \mathfrak{M}_{\mathcal{B}}(\phi)$ allora ϕ è non degenere $\Leftrightarrow \det(A) \neq 0$.

Notazione: Se W è sottospazio di V e $\phi \in \text{PS}(V)$, denoteremo con $\phi|_W$ la restrizione di ϕ a $W \times W$.

Osservazione: Sia $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ prodotto scalare su \mathbb{K}^2 e $W = \text{Span}(e_1)$.

$\phi|_W \equiv 0$, dunque la restrizione di un prodotto scalare non degenere può essere degenere.

DEFINIZIONE 4.2.5: (V, ϕ) e (W, ψ) si dicono **canonicamente isometrici** se esiste un'isometria fra di essi che non dipende da nessuna base.

PROPOSIZIONE 4.2.5: $\phi \in \text{PS}(V)$.

- 1) Se $V = \text{Rad}(\phi) \oplus U$, allora $\phi|_U$ è non degenere.
- 2) Se $V = \text{Rad}(\phi) \oplus U_1 = \text{Rad}(\phi) \oplus U_2$, allora U_1 e U_2 sono canonicamente isometrici.

Dimostrazione:

- 1) Sia $n = \dim(V)$, $p = \dim(\text{Rad}(\phi))$.

Sia $\mathcal{B}_1 = \{v_1, \dots, v_p\}$ base di $\text{Rad}(\phi)$; sia $\mathcal{B}_2 = \{v_{p+1}, \dots, v_n\}$ base di U .

Sia $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ (che ovviamente è base di V); allora:

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & | & 0 \\ \hline 0 & | & A \end{pmatrix},$$

dove $A = \mathfrak{M}_{\mathcal{B}_2}(\phi|_U) \in \mathcal{M}(n-p, \mathbb{K})$.

Poiché $\text{rk}(\mathfrak{M}_{\mathcal{B}}(\phi)) = n - \dim(\text{Rad}(\phi)) = n - p$, allora A è invertibile e dunque $\phi|_U$ è non degenere.

- 2) $\forall u \in U_1 \subseteq V \exists! z_u \in \text{Rad}(\phi), u' \in U_2 \mid u = z_u + u'$. Definiamo $L: U_1 \rightarrow U_2 \mid L(u) = u'$.

L è sicuramente lineare, in quanti restrizione di $\pi_{U_2}: V \rightarrow U_2$.

Poiché $\dim(U_1) = \dim(U_2)$, per provare che L è isomorfismo basta provare che è iniettiva.

Sia $u \in \text{Ker}(L) \Rightarrow u' = L(u) = 0 \Rightarrow u = z_u \in \text{Rad}(\phi) \Rightarrow u \in U_1 \cap \text{Rad}(\phi) = \{0\}$.

Mostriamo che L è un'isometria:

$$\phi(u, v) = \phi(z_u + u', z_v + v') = \phi(z_u, z_v) + \phi(z_u, v') + \phi(u', z_v) + \phi(u', v') = \phi(u', v') = \\ = \phi(L(u), L(v)), \text{ da cui la tesi.}$$

DEFINIZIONE 4.2.6: Se $\phi \in PS(V)$ e $S \subseteq V$, definiamo ortogonale di S :
 $S^\perp = \{v \in V \mid \phi(v, s) = 0 \ \forall s \in S\}$.

Osservazione: $Rad(\phi) = V^\perp$.

PROPOSIZIONE 4.2.6: Siano $S, T \subseteq V$.

- 1) S^\perp è sottospazio di V
- 2) $S \subseteq T \Rightarrow T^\perp \subseteq S^\perp$
- 3) $S^\perp = (Span(S))^\perp$
- 4) $S \subseteq S^{\perp\perp}$

Siano U, W sottospazi di V .

- 5) $(U + W)^\perp = U^\perp \cap W^\perp$
- 6) $U^\perp + W^\perp \subseteq (U \cap W)^\perp$

Dimostrazione:

- 1) Verifica immediata.
- 2) $v \in T^\perp \Rightarrow \phi(v, t) = 0 \ \forall t \in T \supseteq S \Rightarrow \phi(v, s) = 0 \ \forall s \in S \Rightarrow v \in S^\perp$.
- 3) Poiché $S \subseteq Span(S)$, per 2) ho che $(Span(S))^\perp \subseteq S^\perp$.

Inoltre se $v \in S^\perp \Rightarrow \phi(v, s) = 0 \ \forall s \in S$.

Ma se $S = \{s_1, \dots, s_k\}$, allora:

$$\phi(v, \sum_{i=1}^k a_i s_i) = \sum_{i=1}^k \phi(v, s_i) = 0, \text{ da cui la tesi.}$$

- 4) $v \in S \Rightarrow \phi(v, s) = 0 \ \forall s \in S^\perp \Rightarrow v \in S^{\perp\perp}$.
- 5) $v \in (U + W)^\perp \Leftrightarrow \phi(v, u + w) = 0 \ \forall u \in U, w \in W \Leftrightarrow \phi(v, u) + \phi(v, w) = 0 \Leftrightarrow$
 $\Leftrightarrow \phi(v, u) = 0 \wedge \phi(v, w) = 0 \ \forall u \in U, w \in W \Leftrightarrow v \in U^\perp \cap W^\perp,$
 dove il passaggio contrassegnato con \Leftrightarrow si ottiene ponendo prima $u = 0$ e poi $w = 0$.
- 6) $v \in U^\perp + W^\perp \Rightarrow v = v_U + v_W$, con $v_U \in U^\perp, v_W \in W^\perp \Rightarrow \forall h \in U \cap W, \phi(v, h) =$
 $= \phi(v_U + v_W, h) = \phi(v_U, h) + \phi(v_W, h) = 0 + 0 = 0 \Rightarrow v \in (U \cap W)^\perp$.

Osservazione: Se W è sottospazio di V , allora $W \cap W^\perp = Rad(\phi|_W)$, poiché in $W \cap W^\perp$ ci sono i vettori di W ortogonali a tutti i vettori di W , cioè i vettori di $Rad(\phi|_W)$.

Esempio: Sia $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ e $W = Span(e_1)$. $W^\perp = Span(e_2) \Rightarrow W \cap W^\perp = Rad(\phi|_W) = Span(e_1)$.

PROPOSIZIONE 4.2.7: Se W è sottospazio di V e $W \cap Rad(\phi) = \{0\}$, allora:
 $\dim(W^\perp) = \dim(V) - \dim(W)$.

(Osservazione: Non è detto che W e W^\perp siano in somma diretta, nonostante abbiamo dimensioni adeguate; un controesempio può essere l'esempio precedente).

Dimostrazione:

Sia $\dim(W) = k$ e $\dim(Rad(\phi)) = h$.

Esiste una base \mathcal{B} di V del tipo:

$$\mathcal{B} = \left\{ \underbrace{w_1, \dots, w_k}_{\text{base di } W}, v_{k+1}, \dots, v_{n-h}, \underbrace{v_{n-h+1}, \dots, v_n}_{\text{base di } Rad(\phi)} \right\}$$

Allora:

$$A = \mathfrak{M}_B(\phi) = \left(\begin{array}{c|c|c} M_1 & M_2 & 0 \\ \hline {}^t M_2 & M_3 & 0 \\ \hline 0 & 0 & 0 \end{array} \right)$$

e $rk(A) = n - h$.

$$W^\perp = \{v \in V \mid \phi(v, w) = 0 \ \forall w \in W\} = \{v \in V \mid \phi(w_1, v) = \dots = \phi(w_k, v) = 0\}.$$

Infatti il contenimento \subseteq è ovvio, mentre l'altro segue dal fatto che ogni $w \in W$ si può scrivere come combinazione lineare dei w_i .

Ora:

$$\begin{aligned} \phi(w_1, v) &= (1 \ 0 \ \dots \ 0)AX \\ &\vdots \\ \phi(w_k, v) &= (0 \ \dots \ 1 \ 0 \ \dots \ 0)AX \end{aligned}, \text{ con } X = [v]_B.$$

Dunque attraverso l'isomorfismo $[\]_B$ i vettori di W^\perp corrispondono alle soluzioni del sistema lineare:

$$\underbrace{(I_k \mid 0)}_{=(M_1 \mid M_2 \mid 0)} AX = 0.$$

Poiché $rk(A) = n - h$, la matrice $\left(\begin{array}{c|c} M_1 & M_2 \\ \hline {}^t M_2 & M_3 \end{array} \right)$ è invertibile e dunque $rk(M_1 \mid M_2) = k$.

Allora lo spazio delle soluzioni del sistema $(M_1 \mid M_2 \mid 0)X = 0$ ha dimensione $n - k$ e dunque $\dim(W^\perp) = n - k$.

COROLLARIO 4.2.8: 1) $\dim(W^\perp) = \dim(V) - \dim(W) + \dim(W \cap Rad(\phi))$

2) In generale $\dim(W^\perp) + \dim(W) \geq \dim(V)$

3) Se ϕ è non degenere, allora $\dim(W^\perp) = \dim(V) - \dim(W)$

4) $\phi|_W$ è non degenere $\Leftrightarrow V = W \oplus W^\perp$.

Dimostrazione:

1) Sia $W = (W \cap Rad(\phi)) \oplus W_1$.

Allora $W^\perp = W_1^\perp$, infatti sicuramente $W^\perp \subseteq W_1^\perp$ in quanto $W_1 \subseteq W$, e:

$$\begin{aligned} v \in W_1^\perp &\Rightarrow \phi(v, w') = 0 \ \forall w' \in W_1 \Rightarrow \phi(v, w) = \phi(v, w_0 + w'), w_0 \in Rad(\phi), w' \in W_1 \\ &\Rightarrow \phi(v, w) = \phi(v, w_0) + \phi(v, w') = 0 + 0 = 0 \Rightarrow v \in W^\perp \end{aligned}$$

Inoltre $W_1 \cap Rad(\phi) = \{0\}$, poiché:

$$\{0\} = W_1 \cap (W \cap Rad(\phi)) = (W_1 \cap W) \cap Rad(\phi) \stackrel{\square}{=} W_1 \cap Rad(\phi),$$

dove il passaggio contrassegnato con \square segue dal fatto che $W_1 \subseteq W$.

Per la proposizione precedente:

$$\dim(W^\perp) = \dim(W_1^\perp) = \dim(V) - \dim(W_1) = \dim(V) - (\dim(W) - \dim(W \cap Rad(\phi))),$$

da cui la tesi.

2) Segue dal punto 1).

3) Poiché $Rad(\phi) = \{0\}$ e per il punto 1) si ha la tesi.

4) $\phi|_W$ è non degenere $\Leftrightarrow Rad(\phi|_W) = W \cap W^\perp = \{0\}$.

Dunque $\dim(W \oplus W^\perp) \leq \dim(V)$.

D'altra parte $\dim(W \oplus W^\perp) = \dim(W) + \dim(W^\perp) \geq \dim(V)$, dunque ho la tesi.

Osservazione: Se ϕ è non degenere, poiché $U \subseteq U^{\perp\perp}$ e $\dim(U) = \dim(U^{\perp\perp})$, allora segue che

$U = U^{\perp\perp}$. Inoltre, sapendo che $(U + W)^\perp = U^\perp \cap W^\perp$, se ϕ è non degenere segue che $U^\perp + W^\perp = (U \cap W)^\perp$ (con un ragionamento analogo al precedente).

DEFINIZIONE 4.2.7: Se $V = W \oplus W^\perp$, la proiezione $\pi_W: V \rightarrow W$ è detta **proiezione ortogonale** su W .

Osservazione: $\forall v \in V, v - \pi_W(v) \in W^\perp$.

DEFINIZIONE 4.2.8: $v \in V$ si dice **isotropo** se $\phi(v, v) = 0$.
Denotiamo con $\mathcal{I}(\phi)$ l'insieme dei vettori isotropi per ϕ .

Osservazione: Se ogni $v \in V$ è isotropo per ϕ , allora $\phi \equiv 0$ (per la formula di polarizzazione).

Osservazione: Se v non è isotropo, quindi $V = \text{Span}(v) \oplus \text{Span}(v)^\perp$, allora ogni $w \in V$ si scrive come $w = w_1 + w_2$, con $w_1 \in \text{Span}(v)$ e $w_2 \in \text{Span}(v)^\perp$.

Inoltre $w_1 = c \cdot v$, quindi $w_2 = w - cv$.

$$w_2 = w - cv \in \text{Span}(v)^\perp \Leftrightarrow \phi(w - cv, v) = 0 \Leftrightarrow \phi(w, v) = c\phi(v, v) \Leftrightarrow c = \frac{\phi(v, w)}{\phi(v, v)}$$

Il numero $c = \frac{\phi(v, w)}{\phi(v, v)}$ prende il nome di **coefficiente di Fourier** di w rispetto a v .

Dunque $\pi|_{\text{Span}(v)}(w) = \frac{\phi(v, w)}{\phi(v, v)}v$.

DEFINIZIONE 4.2.9: Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V si dice **ortogonale** se $\phi(v_i, v_j) = 0 \quad \forall i \neq j$.

Osservazione: \mathcal{B} è ortogonale $\Leftrightarrow \mathfrak{M}_{\mathcal{B}}(\phi)$ è diagonale.

PROPOSIZIONE 4.2.9: $\forall \phi \in \text{PS}(V)$ esiste una base di V ortogonale rispetto a ϕ .

Dimostrazione 1:

Per induzione su $n = \dim(V)$:

Passo base): $n = 1$: ogni base è ortogonale.

Passo induttivo): Se $\phi(v, v) = 0 \quad \forall v \in V$, allora $\phi \equiv 0$ e quindi ogni base è ortogonale.

Altrimenti $\exists v_1 \mid \phi(v_1, v_1) \neq 0$; allora $V = \text{Span}(v_1) \oplus \text{Span}(v_1)^\perp$.

Per ipotesi induttiva $\exists \{v_2, \dots, v_n\}$ base di $\text{Span}(v_1)^\perp$ ortogonale per la restrizione di ϕ (e dunque per ϕ).

Allora $\{v_1, \dots, v_n\}$ è base di V ortogonale per ϕ .

Dimostrazione 2 – **Algoritmo di Lagrange**:

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base qualsiasi di V . Sia $A = \mathfrak{M}_{\mathcal{B}}(\phi)$.

• Supponiamo $[A]_{11} = \phi(v_1, v_1) \neq 0$, cioè $v_1 \notin \mathcal{I}(\phi)$. Poniamo:

$$v'_1 = v_1;$$

$$v'_2 = v_2 - \frac{\phi(v_2, v'_1)}{\phi(v'_1, v'_1)}v'_1;$$

⋮

$$v'_n = v_n - \frac{\phi(v_n, v'_1)}{\phi(v'_1, v'_1)}v'_1.$$

Allora $\phi(v'_j, v'_1) = 0 \quad \forall j \geq 2$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ è una base di V , infatti, se mettiamo i vettori v'_i per colonna in una matrice, otteniamo:

$$M = \begin{pmatrix} 1 & -k_2 & -k_3 & \dots & -k_n \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ 0 & 0 & 0 & & 1 \end{pmatrix}$$

che ha evidentemente $\det M = 1 \neq 0$.

Inoltre abbiamo che:

$$\mathfrak{M}_{\mathcal{B}'}(\phi) = \begin{pmatrix} * & 0 & \dots & 0 \\ 0 & \boxed{C} & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$$

- Se $[A]_{11} = 0$ guardo se $\exists j \in \{2, \dots, n\} | [A]_{jj} \neq 0$.
Se lo trovo, permuto la base \mathcal{B} in modo che v_j sia il primo vettore e procedo come prima.
Altrimenti:
- Se $[A]_{ii} = 0 \ \forall i$ ci sono due casi:
 - a) ϕ è nullo, quindi ogni base è ortogonale;
 - b) $\exists i \neq j | [A]_{ij} = [A]_{ji} \neq 0$. In tal caso $\phi(v_i + v_j, v_i + v_j) = 2[A]_{ij} \neq 0$.
Allora scelgo una base di V in cui $v_i + v_j$ è il primo vettore e poi applico il primo caso.

Dopo aver ortogonalizzato i vettori rispetto al primo, itero il procedimento sulla matrice C e così via.

COROLLARIO 4.2.10: Ogni matrice simmetrica è congruente ad una matrice diagonale.

Osservazione: Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortogonale.

Sia $v = \alpha_1 v_1 + \dots + \alpha_n v_n$.

$\phi(v, v_1) = \alpha_1 \phi(v_1, v_1)$, quindi, se v_1 è non isotropo, la coordinata di v_1 coincide con il coefficiente di Fourier di v rispetto a v_1 .

Quindi, se ϕ è non degenere:

$$v = \frac{\phi(v, v_1)}{\phi(v_1, v_1)} v_1 + \dots + \frac{\phi(v, v_n)}{\phi(v_n, v_n)} v_n$$

Inoltre, se W è sottospazio di V , $\dim(W) = k$, $\phi|_W$ è non degenere (per cui $V = W \oplus W^\perp$) e $\{w_1, \dots, w_k\}$ è una base ortogonale di W , allora:

$$v = \frac{\phi(v, w_1)}{\phi(w_1, w_1)} w_1 + \dots + \frac{\phi(v, w_k)}{\phi(w_k, w_k)} w_k + z, z \in W^\perp$$

Quindi:

$$\pi_W(v) = \frac{\phi(v, w_1)}{\phi(w_1, w_1)} w_1 + \dots + \frac{\phi(v, w_k)}{\phi(w_k, w_k)} w_k$$

Osservazione: Sia V un \mathbb{C} -spazio vettoriale, $\dim(V) = n$, $\phi \in PS(V)$, $rk(\phi) = r$.

Allora $\exists \mathcal{S} = \{v_1, \dots, v_n\}$ base ortogonale di V tale che:

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} a_{11} & & & & & \\ & \ddots & & & & \\ & & a_{rr} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

con $a_{ii} \neq 0 \ \forall i \leq r$.

Sia $\mathcal{B} = \left\{ \frac{v_1}{\sqrt{\phi(v_1, v_1)}}, \dots, \frac{v_r}{\sqrt{\phi(v_r, v_r)}}, v_{r+1}, \dots, v_n \right\}$.

\mathcal{B} è detta **base ortogonale normalizzata** per ϕ e:

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}.$$

TEOREMA DI SYLVESTER COMPLESSO: Sia V un \mathbb{C} -spazio vettoriale.

Allora (V, ϕ) e (V, ψ) sono isometrici $\Leftrightarrow rk(\phi) = rk(\psi)$.

(Osservazione: Quindi il rango è un sistema completo di invarianti per l'isometria nel caso $\mathbb{K} = \mathbb{C}$).

Dimostrazione:

(V, ϕ) e (V, ψ) sono isometrici $\Leftrightarrow \exists \mathcal{B}, \mathcal{S}$ basi di $V \mid \mathfrak{M}_{\mathcal{B}}(\phi) = \mathfrak{M}_{\mathcal{S}}(\psi)$.

Poiché abbiamo visto nell'osservazione precedente che se $rk(\phi) = rk(\psi)$ allora $\exists \mathcal{B}, \mathcal{S}$ basi di

$V \mid \mathfrak{M}_{\mathcal{B}}(\phi) = \mathfrak{M}_{\mathcal{S}}(\psi) = \begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}$, segue la tesi.

COROLLARIO 4.2.11: Tutti i prodotti scalari non degeneri su un \mathbb{C} -spazio vettoriale sono isometrici.

COROLLARIO 4.2.12: Ogni matrice simmetrica complessa di rango r è congruente a $\begin{pmatrix} I_r & | & 0 \\ \hline 0 & | & 0 \end{pmatrix}$ e dunque il rango è un invariante completo di congruenza su \mathbb{C} .

Osservazione: Sia V un \mathbb{R} -spazio vettoriale, $\dim(V) = n$, $\phi \in PS(V)$, $rk(\phi) = r$.

Allora $\exists \mathcal{S} = \{v_1, \dots, v_n\}$ base ortogonale di V tale che:

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} a_{11} & & & & & \\ & \ddots & & & & \\ & & a_{rr} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

con $a_{ii} \neq 0 \ \forall i \leq r$.

Supponiamo che $a_{ii} > 0$ per $1 \leq i \leq p$ e $a_{ii} < 0$ per $p+1 \leq i \leq r$.

Sia $\mathcal{B} = \left\{ \frac{v_1}{\sqrt{a_{11}}}, \dots, \frac{v_p}{\sqrt{a_{pp}}}, \frac{v_{p+1}}{\sqrt{-a_{(p+1)(p+1)}}}, \frac{v_r}{\sqrt{-a_{rr}}}, v_{r+1}, \dots, v_n \right\}$.

\mathcal{B} è detta **base ortogonale normalizzata** per ϕ e:

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} \boxed{I_p} & & & & & \\ & \boxed{-I_{r-p}} & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \boxed{0} \end{pmatrix}.$$

DEFINIZIONE 4.2.10: Sia V un \mathbb{R} -spazio vettoriale, $\phi \in PS(V)$.

- ϕ si dice **definito positivo** (o **negativo**) se $\phi(v, v) > 0 \forall v \neq 0$ (oppure $\phi(v, v) < 0 \forall v \neq 0$);
- ϕ si dice **definito** se è definito positivo o definito negativo;
- ϕ si dice **semidefinito positivo** (o **negativo**) se $\phi(v, v) \geq 0 \forall v$ (oppure $\phi(v, v) \leq 0 \forall v$);
- ϕ si dice **semidefinito** se è semidefinito positivo o semidefinito negativo.

Osservazione: ϕ definito $\Rightarrow \phi$ non degenerare; inoltre se W è sottospazio di V e ϕ è (semi)definito, allora $\phi|_W$ è (semi)definito.

DEFINIZIONE 4.2.11: • Il numero $i_+(\phi) = \max\{\dim(W) \mid W \text{ ssv di } V, \phi|_W \text{ def. positivo}\}$ prende il nome di **indice di positività**;

- Il numero $i_-(\phi) = \max\{\dim(W) \mid W \text{ ssv di } V, \phi|_W \text{ def. negativo}\}$ prende il nome di **indice di negatività**;
- Il numero $i_0(\phi) = \dim(\text{Rad}(\phi))$ prende il nome di **indice di nullità**.

Osservazione: Questi tre numeri sono invarianti per isometria, poiché le isometrie mantengono il prodotto scalare e dunque anche i segni dei prodotti scalari.

DEFINIZIONE 4.2.12: La terna $\sigma(\phi) = (i_+(\phi), i_-(\phi), i_0(\phi))$ è detta **segnatura** di ϕ .

TEOREMA DI SYLVESTER REALE: Sia V un \mathbb{R} -spazio vettoriale, $\dim(V) = n$, $\phi \in PS(V)$. Sia \mathcal{B} una base ortogonale di V tale che:

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} \boxed{I_p} & & \\ & \boxed{-I_q} & \\ & & \boxed{0} \end{pmatrix}.$$

Allora $p = i_+(\phi)$ e $q = i_-(\phi)$.

Dimostrazione:

Sia $\mathcal{B} = \{v_1, \dots, v_p, v_{p+1}, \dots, v_{p+q}, v_{p+q+1}, \dots, v_n\}$.

La restrizione di ϕ a $\text{Span}(v_1, \dots, v_p)$ è definita positiva, dunque $i_+(\phi) \geq p$.

Sia ora W un sottospazio di V tale che $\dim(W) = i_+(\phi) \wedge \phi|_W$ è definito positivo.

Sia $Z = \text{Span}(v_{p+1}, \dots, v_n)$. Dunque $\forall z \in Z$, $z = a_{p+1}v_{p+1} + \dots + a_nv_n$.

$\phi(z, z) = -a_{p+1}^2 - \dots - a_{p+q}^2 \leq 0$, dunque $\phi|_Z$ è semidefinito negativo.

Notiamo che $W \cap Z = \{0\}$, infatti, se $v \in W \cap Z$, allora $\phi(v, v) \geq 0 \wedge \phi(v, v) \leq 0$, quindi $\phi(v, v) = 0$, cioè $v = 0$ perché $v \in W$ ed essendo W definito positivo il suo unico vettore isotropo è il vettore nullo.

Allora esiste $W \oplus Z$ sottospazio di V , per cui $\dim(W \oplus Z) = \dim(W) + \dim(Z) \leq \dim(V) = n$, ossia $i_+(\phi) + n - p \leq n$, cioè $i_+(\phi) \leq p$.

Segue dunque che $i_+(\phi) = p$ e con questo che $i_+(\phi)$ non dipende dalla scelta della base.

Poiché $p + q = rk(\phi)$ e $i_+(\phi) + i_-(\phi) = rk(\phi)$, allora $q = i_-(\phi)$, da cui la tesi.

COROLLARIO 4.2.13: Se $\mathbb{K} = \mathbb{R}$, (V, ϕ) e (V, ψ) sono isometrici $\Leftrightarrow \sigma(\phi) = \sigma(\psi)$, cioè la segnatura è un invariante completo di isometria nel caso $\mathbb{K} = \mathbb{R}$.

COROLLARIO 4.2.14: $A, B \in \mathcal{S}(n, \mathbb{R})$ sono congruenti $\Leftrightarrow \sigma(A) = \sigma(B)$, cioè la segnatura è un invariante completo di congruenza nel caso reale.

DEFINIZIONE 4.2.13: V \mathbb{K} -spazio vettoriale, $\phi \in PS(V)$. Una base \mathcal{B} di V si dice **ortonormale** per $\phi \Leftrightarrow \mathfrak{M}_{\mathcal{B}}(\phi) = I$.

Osservazione: Se $\mathbb{K} = \mathbb{C}$ \exists una base ortonormale per $\phi \Leftrightarrow \phi$ è non degenere.
 Se $\mathbb{K} = \mathbb{R}$ \exists una base ortonormale per $\phi \Leftrightarrow \phi$ è definito positivo.

DEFINIZIONE 4.2.14: Sia $A \in \mathcal{S}(n, \mathbb{R})$. A si dice **definita positiva (negativa)** se ${}^tXAX > 0$ (${}^tXAX < 0$) $\forall X \neq 0$.

Sia $A \in \mathcal{S}(n, \mathbb{R})$. A si dice **semidefinita positiva (negativa)** se ${}^tXAX \geq 0$ (${}^tXAX \leq 0$) $\forall X$.

Osservazione: A è definita positiva $\Leftrightarrow \psi_A: (X, Y) \rightarrow {}^tXAY$ è definito positivo.
 Analogamente se è definita negativa.

Osservazione: A è definita positiva $\Leftrightarrow A$ è congruente a $I \Leftrightarrow \exists M \in GL(n, \mathbb{R}) \mid A = {}^tMM$ (dunque $\det(A) = \det({}^tMM) = (\det(M))^2 > 0$).

DEFINIZIONE 4.2.15: $A \in \mathcal{S}(n, \mathbb{R})$. $\forall 1 \leq i \leq n$ si definisce i -esimo minore principale $M_i(A)$ il minore formato dalle prime i righe e dalle prime i colonne.

CRITERIO DEI MINORI PRINCIPALI: $A \in \mathcal{S}(n, \mathbb{R})$.

A è definita positiva $\Leftrightarrow \det(M_i(A)) > 0 \forall i$.

Dimostrazione:

\Rightarrow) $M_i(A)$ è la matrice associata alla restrizione di ψ_A al sottospazio $Span(e_1, \dots, e_i)$. Tale restrizione è definita positiva e quindi $\det(M_i(A)) > 0 \forall i$.

\Leftarrow) Per induzione su n :

Passo base): $n = 1$: ovvio.

Passo induttivo): Per ipotesi tutti i minori principali della matrice $M_{n-1}(A)$ hanno determinante positivo.

Allora per ipotesi induttiva la restrizione di ψ_A a $Span(e_1, \dots, e_{n-1})$ è definita positiva e quindi $i_+(\psi_A) \geq n - 1$.

Allora, se $i_+(\psi_A) = n - 1$, esisterebbe per Sylvester una base \mathcal{B} tale che:

$$\mathfrak{M}_{\mathcal{B}}(\psi_A) = \begin{pmatrix} \boxed{I_{n-1}} & \\ & -1 \end{pmatrix}$$

e $\det(\mathfrak{M}_{\mathcal{B}}(\psi_A)) = -1 < 0$, assurdo, quindi $i_+(\psi_A) = n$, cioè A è definita positiva.

PROPOSIZIONE 4.2.15: Le trasformazioni di base con l'algoritmo di Lagrange nel caso in cui il vettore non sia isotropo non alterano i determinanti dei minori principali.

Dimostrazione:

$$A = \mathfrak{M}_{\mathcal{B}}(\psi) \xrightarrow{\text{trasf. di base}} A' = \mathfrak{M}_{\mathcal{B}'}(\psi); \mathcal{B} = \{v_1, \dots, v_n\}, \mathcal{B}' = \{v'_1, \dots, v'_n\}.$$

$$M = \mathfrak{M}_{\mathcal{B}'\mathcal{B}}(id) = \begin{pmatrix} 1 & * & * & & * \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ 0 & 0 & 0 & & 1 \end{pmatrix}$$

Notiamo che $A' = {}^tMAM$; sia inoltre $W_k = \text{Span}(v_1, \dots, v_k)$.

Allora $A_k = \mathfrak{M}_{\{v_1, \dots, v_k\}}(\psi|_{W_k})$, $W_k = \text{Span}(v'_1, \dots, v'_k)$ e $\mathfrak{M}_{\{v'_1, \dots, v'_k\}, \{v_1, \dots, v_k\}}(id) = M_k$.

Quindi $A'_k = {}^tM_k A_k M_k$.

Ma $\det(M_k) = 1 \ \forall k$, quindi $D'_k = \det(A'_k) = \det(A_k) \cdot (\det(M_k))^2 = \det(A_k) = D_k$, tesi.

CRITERIO DI JACOBI: $A \in \mathcal{S}(n, \mathbb{R})$, $rk(A) = r$. Supponiamo che $D_i \neq 0 \ \forall i \leq r$.

Allora $\exists T$ triangolare superiore, $\det(T) = 1$, tale che:

$${}^tTAT = \begin{pmatrix} D_1 & & & & & \\ & \frac{D_2}{D_1} & & & & \\ & & \ddots & & & \\ & & & \frac{D_r}{D_{r-1}} & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

Dimostrazione:

$D_1 \neq 0 \Rightarrow [A]_{11} = D_1 \neq 0$, quindi posso effettuare una trasformazione di base.

$A' = {}^tMAM$, M triangolare superiore, $\det(M) = 1$.

$$A' = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & a'_{22} & \dots & * \\ \vdots & \vdots & \ddots & \\ 0 & * & & * \end{pmatrix}.$$

Ma si conservano i determinanti dei minori principali $\Rightarrow D_2 = D'_2 = D_1 \cdot a'_{22} \Rightarrow a'_{22} = \frac{D_2}{D_1}$.

Itero perché $\frac{D_2}{D_1} \neq 0 \Rightarrow a'_{33} \cdot \frac{D_2}{D_1} \cdot D_1 = D_3 = D'_3 \Rightarrow a'_{33} = \frac{D_3}{D_2}$.

In generale $a'_{kk} \cdot \frac{D_{k-1}}{D_{k-2}} \cdot \dots \cdot \frac{D_2}{D_1} \cdot D_1 = D_k \Rightarrow a'_{kk} = \frac{D_k}{D_{k-1}}$.

Quindi, chiamando \tilde{A} la matrice A dopo r trasformazioni di base:

$$\tilde{A} = {}^tM_r {}^tM_{r-1} \dots {}^tM_1 A M_1 \dots M_r = {}^t(M_1 \dots M_r) A M_1 \dots M_r,$$

quindi se $T = M_1 \dots M_r$, allora T è la matrice cercata, in quanto è triangolare superiore (perché prodotto di matrici triangolari superiori) e ha $\det(T) = \det(M_1 \dots M_r) = 1$.

COROLLARIO 4.2.16: Si deduce il criterio dei minori principali.

Dimostrazione:

ψ è definito positivo $\Leftrightarrow D_1 > 0, \frac{D_2}{D_1} > 0, \dots, \frac{D_r}{D_{r-1}} > 0 \Leftrightarrow D_1 > 0, \dots, D_r > 0$.

COROLLARIO 4.2.17: A è definita negativa $\Leftrightarrow D_1 < 0, D_2 > 0, D_3 < 0, \dots$

Dimostrazione:

ψ è definito negativo $\Leftrightarrow D_1 < 0, \frac{D_2}{D_1} < 0, \dots, \frac{D_r}{D_{r-1}} < 0 \Leftrightarrow D_1 < 0, D_2 > 0, D_3 < 0, \dots$

DEFINIZIONE 4.2.16: (P, ψ) si dice **piano iperbolico** se P è uno spazio vettoriale di dimensione 2 e ψ è un prodotto scalare di P non degenerare per cui esiste un vettore isotropo non nullo.

PROPOSIZIONE 4.2.18: (P, ψ) piano iperbolico, $v \neq 0$ isotropo. Allora v si estende ad una base $\mathcal{B} = \{v, w\}$ di P tale che $\mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (\mathcal{B} è detta **base iperbolica**).

Dimostrazione:

Sia $\mathcal{S} = \{v, z\}$ una base di P . Allora $\mathfrak{M}_{\mathcal{S}}(\psi) = \begin{pmatrix} 0 & a \\ a & b \end{pmatrix}$.

ψ non degenerare $\Rightarrow a \neq 0$.

Ora cerco $\lambda, \mu \mid w = \lambda v + \mu z$, $\mathcal{B} = \{v, w\}$ sia base di P e $\mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Notiamo che \mathcal{B} è base $\Leftrightarrow \mu \neq 0$.

$\psi(v, w) = \psi(v, \lambda v + \mu z) = \mu a$; $\psi(w, w) = \psi(\lambda v + \mu z, \lambda v + \mu z) = 2\lambda\mu a + \mu^2 b$.

Allora $\mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} \mu a = 1 \\ 2\lambda\mu a + \mu^2 b \end{cases} \Leftrightarrow \begin{cases} \mu = a^{-1} \\ \lambda = -2^{-1}b\mu^2 = -b(2a^2)^{-1} \end{cases}$$

Poiché abbiamo trovato tali λ, μ , ho la tesi.

Osservazione: Se (P, ϕ) è un piano iperbolico, allora $\sigma(\phi) = (1, 1, 0)$. Infatti sia $\{v, w\}$ una base iperbolica per (P, ϕ) . Sia $\mathcal{S} = \left\{ \frac{v+w}{\sqrt{2}}, \frac{v-w}{\sqrt{2}} \right\}$, che quindi è base di P . Si può facilmente vedere che:

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

da cui segue che $\sigma(\phi) = (1, 1, 0)$.

DEFINIZIONE 4.2.17: (V, ϕ) si dice **anisotropo** se V non contiene vettori isotropi non nulli.

PROPOSIZIONE 4.2.19: V \mathbb{K} -spazio vettoriale, ϕ prodotto scalare non degenerare. Allora:

- 1) Se $\mathbb{K} = \mathbb{C}$, (V, ϕ) è anisotropo $\Leftrightarrow \dim(V) = 1$;
- 2) Se $\mathbb{K} = \mathbb{R}$, (V, ϕ) è anisotropo $\Leftrightarrow \phi$ è definito.

Dimostrazione:

1) \Leftarrow) Ovvio.

\Rightarrow) Se per assurdo $\dim(V) \geq 2$ e $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di $V \mid \mathfrak{M}_{\mathcal{B}}(\phi) = I$, allora $\phi(v_1 + iv_2, v_1 + iv_2) = 1 - 1 = 0$ e dunque $v_1 + iv_2 \in \mathcal{J}(\phi)$, assurdo.

2) \Leftarrow) Ovvio per definizione.

\Rightarrow) Se per assurdo ϕ non è definito, $\exists \mathcal{B} = \{v_1, \dots, v_p, v_{p+1}, \dots, v_n\}$ base di V tale che

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} \boxed{I_p} & & \\ & & \\ & & \boxed{-I_{n-p}} \end{pmatrix}. \text{ Allora } \phi(v_1 + v_{p+1}, v_1 + v_{p+1}) = 1 - 1 = 0, \text{ assurdo.}$$

Notazione: Denoteremo con $W_1 \oplus^\perp W_2$ la **somma diretta ortogonale** di W_1 e W_2 , che sta ad indicare la somma diretta dei sottospazi W_1 e W_2 tali che $\phi(w_1, w_2) = 0 \quad \forall w_1 \in W_1, w_2 \in W_2$.

FORMA NORMALE DI WITT: Sia (V, ϕ) , ϕ non degenera.

Caso 1): $\mathbb{K} = \mathbb{C}$.

a) $\dim(V) = n = 2m$.

Sappiamo che $\exists \mathcal{B} = \{v_1, w_1, \dots, v_m, w_m\}$ base di V | $\mathfrak{M}_{\mathcal{B}}(\phi) = I$.

Allora $\mathfrak{M}_{\{v_j, w_j\}}(\phi|_{\text{Span}(v_j, w_j)}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e $v_j + iw_j$ è isotropo, dunque

$P_j = \text{Span}(v_j, w_j)$ è un piano iperbolico e:

$$V = P_1 \oplus^\perp \dots \oplus^\perp P_m,$$

detta **decomposizione di Witt** di V .

Prendendo una base iperbolica in ogni P_j , $\exists \mathcal{S}$ base di V tale che:

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & & & \\ & \ddots & & \\ & & \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & \\ & & & \end{pmatrix},$$

detta **forma normale di Witt** di V .

b) $\dim(V) = n = 2m + 1$.

Con un procedimento analogo al precedente si trova la decomposizione di Witt:

$$V = P_1 \oplus^\perp \dots \oplus^\perp P_m \oplus^\perp \text{Span}(z),$$

Inoltre come prima $\exists \mathcal{S}$ base di V |

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & & & \\ & \ddots & & \\ & & \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & \\ & & & 1 \end{pmatrix},$$

che è la forma normale di Witt di V .

Caso 2): $\mathbb{K} = \mathbb{R}$.

1) $i_+(\phi) \leq i_-(\phi)$. Sia $p = i_+(\phi)$.

Allora $\exists \mathcal{B} = \{v_1, \dots, v_p, w_1, \dots, w_{n-p}\}$ base di V |

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} \boxed{I_p} & \\ & \boxed{-I_{n-p}} \end{pmatrix}.$$

Si definisca $P_j = \text{Span}(v_j, w_j) \forall j$ e $A = \text{Span}(w_{p+1}, \dots, w_{n-p})$. Allora P_j è un piano iperbolico $\forall j$ e ϕ_A è definito negativo; inoltre:

$$V = P_1 \oplus^\perp \dots \oplus^\perp P_p \oplus^\perp A,$$

che è la decomposizione di Witt di V e $\exists \mathcal{S}$ base di V |

$$\mathfrak{M}_{\mathcal{S}}(\phi) = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & & & \\ & \ddots & & \\ & & \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} & \\ & & & \boxed{-I_{n-2p}} \end{pmatrix},$$

che è la forma normale di Witt di V .

2) $i_-(\phi) \leq i_+(\phi)$.

È del tutto analogo al caso precedente, tranne che $\phi|_A$ è definito positivo.

In generale:

DEFINIZIONE 4.2.18: Se ϕ è non degenera, si chiama **decomposizione di Witt** di (V, ϕ) una decomposizione:

$$V = P_1 \oplus^\perp \dots \oplus^\perp P_h \oplus^\perp A,$$

dove ogni P_j è un piano iperbolico e $\phi|_A$ è anisotropo.

Dunque, grazie a quello che abbiamo visto, segue:

TEOREMA 4.2.20: Se $\mathbb{K} = \mathbb{C} \vee \mathbb{K} = \mathbb{R}$ e ϕ è non degenera:

- 1) Se $\mathbb{K} = \mathbb{C}$ e $\dim(V) = 2m \Rightarrow \begin{cases} \# \text{piani iperbolici} = m \\ A = \{0\} \end{cases}$
- 2) Se $\mathbb{K} = \mathbb{C}$ e $\dim(V) = 2m + 1 \Rightarrow \begin{cases} \# \text{piani iperbolici} = m \\ \dim(A) = 1 \end{cases}$
- 3) Se $\mathbb{K} = \mathbb{R} \Rightarrow \begin{cases} \# \text{piani iperbolici} = \min(i_+(\phi), i_-(\phi)) \\ \phi|_A \text{ è definito} \end{cases}$

Osservazione: Da questo teorema segue che sia nel caso complesso sia in quello reale, il numero dei piani iperbolici è invariante per isometria.

DEFINIZIONE 4.2.19: $\phi \in PS(V)$. Si definisce **indice di Witt** di (V, ϕ) il numero naturale $w(\phi) = \max\{\dim(W) \mid W \text{ è ssv di } V, \phi|_W \equiv 0\}$.

Osservazioni: 1) $w(\phi) = 0 \Leftrightarrow \phi$ è anisotropo;

2) $w(\phi)$ è invariante per isometria;

3) Se ϕ non degenera, $w(\phi) \leq \frac{\dim(V)}{2}$

Dimostrazione:

Sia W un sottospazio di V tale che $\dim(W) = w(\phi) \wedge \phi|_W \equiv 0$.

Allora $\exists B$ base di V

$$\mathfrak{M}_B(\phi) = \left(\begin{array}{c|c} 0 & A \\ \hline \underbrace{{}^t A}_{w(\phi)} & \underbrace{C}_{n-w(\phi)} \end{array} \right)$$

Quindi $n = rk(\phi) \leq \underbrace{(n - w(\phi))}_{=\max(rk({}^t A|C))} + \underbrace{(n - w(\phi))}_{=\max(rk(A))} \Rightarrow n \geq 2w(\phi)$, da cui la tesi.

4) Se $V = P_1 \oplus^\perp \dots \oplus^\perp P_h \oplus^\perp A$ è una decomposizione di Witt, allora $h \leq w(\phi)$.

Dimostrazione:

Se $\{v_j, w_j\}$ è una base iperbolica per P_j , allora $Z = \text{Span}(v_1, \dots, v_h)$ è un sottospazio di V $\dim(Z) = h$ e $\phi|_Z \equiv 0$, dunque $w(\phi) \geq h$.

TEOREMA 4.2.21: Sia (V, ϕ) , ϕ non degenera, $\dim(V) = n$. Allora:

1) Se $\mathbb{K} = \mathbb{C}$, $w(\phi) = \lfloor \frac{n}{2} \rfloor$;

2) Se $\mathbb{K} = \mathbb{R}$, $w(\phi) = \min(i_+(\phi), i_-(\phi))$.

Dimostrazione:

1) Se $\mathbb{K} = \mathbb{C}$, abbiamo visto che $V = P_1 \oplus^\perp \dots \oplus^\perp P_{\lfloor \frac{n}{2} \rfloor} \oplus^\perp A$, con $\dim(A) \leq 1$.

Allora, per le osservazioni 3) e 4), abbiamo che $\lfloor \frac{n}{2} \rfloor \leq w(\phi) \leq \frac{n}{2} \Rightarrow w(\phi) = \lfloor \frac{n}{2} \rfloor$.

2) Sia $V = P_1 \oplus^\perp \dots \oplus^\perp P_m \oplus^\perp A$ una decomposizione di Witt di V .

Vogliamo provare che $m = w(\phi)$, ma per l'osservazione 4), $m \leq w(\phi)$.

Supponiamo per assurdo che $m < w(\phi)$.

Sia Z sottospazio di V tale che $\dim(Z) = w(\phi)$ e $\phi|_Z \equiv 0$.

$\phi|_A$ è definito, supponiamo che sia definito negativo (altrimenti la dimostrazione è analoga).

Allora $\exists W$ sottospazio di V tale che $\dim(W) = n - m$ e $\phi|_W$ è definito negativo (ad esempio

$W = A \oplus \text{Span}(w_1, \dots, w_m)$, con $w_i \in P_i$ $\phi(w_i, w_i) < 0$).

Poiché $\dim(Z) + \dim(W) > n \Rightarrow Z \cap W \neq \{0\} \Rightarrow \exists v \neq 0, v \in Z \cap W$.

Ma $\phi(v, v) = 0$ in quanto $v \in Z$ e $\phi(v, v) < 0$ in quanto $v \in W$, assurdo.

Quindi abbiamo mostrato che tutte le decomposizioni di Witt in V contengono lo stesso numero di piani iperbolici, in particolare $w(\phi)$ piani iperbolici.

Poiché abbiamo trovato una decomposizione di Witt di V con $\min(i_+(\phi), i_-(\phi))$ piani iperbolici, segue la tesi.

DEFINIZIONE 4.2.20: Definiamo segno di (V, ϕ) , con ϕ definito, il numero $\text{sgn}(V)$ che è 1 se (V, ϕ) è definito positivo, -1 se è definito negativo.

COROLLARIO 4.2.22: V \mathbb{R} -spazio vettoriale, ϕ non degenera.

La coppia $(w(\phi), \text{sgn}(\phi|_A))$ è un sistema completo di invarianti per isometria su \mathbb{R} .

Dimostrazione:

Sappiamo che $w(\phi) = \min(i_+(\phi), i_-(\phi))$, inoltre:

- se $\text{sgn}(\phi|_A) = 1 \Rightarrow i_+(\phi) \geq i_-(\phi)$ (questo segue facilmente dalla dimostrazione della forma normale di Witt);
- se $\text{sgn}(\phi|_A) = -1 \Rightarrow i_+(\phi) \leq i_-(\phi)$.

Dunque la conoscenza di $(w(\phi), \text{sgn}(\phi|_A))$ porta immediatamente alla conoscenza di $(i_+(\phi), i_-(\phi))$. Da questo segue la tesi.

PROPOSIZIONE 4.2.20: Se $V = \text{Rad}(\phi) \oplus U$, allora $w(\phi) = \dim(\text{Rad}(\phi)) + w(\phi|_U)$.

Dimostrazione:

Sia Z sottospazio di U tale che $\dim(Z) = w(\phi|_U)$ e $\phi|_Z \equiv 0$.

Allora $\phi|_{Z \oplus \text{Rad}(\phi)} \equiv 0$, quindi $w(\phi) \geq \dim(\text{Rad}(\phi)) + w(\phi|_U)$.

Supponiamo per assurdo che $w(\phi) > \dim(\text{Rad}(\phi)) + w(\phi|_U)$.

Allora sia H sottospazio di V tale che $\dim(H) = w(\phi)$ e $\phi|_H \equiv 0$.

Allora $H \cap U$ è un sottospazio di U in cui ϕ si annulla, quindi:

$$\begin{aligned} w(\phi|_U) &\geq \dim(H \cap U) = \dim(H) + \dim(U) - \dim(H + U) > \\ &> \left(\dim(\text{Rad}(\phi)) + w(\phi|_U) \right) + \dim(U) - \dim(H + U) = \\ &= \dim(V) + w(\phi|_U) - \dim(H + U) \geq w(\phi|_U), \end{aligned}$$

assurdo.

4.3 ISOMETRIE

PROPOSIZIONE 4.3.1: V spazio vettoriale e $\phi \in PS(V)$, $f \in O(V, \phi)$.

Se W è sottospazio di V tale che $f(W) \subseteq W$, allora $f(W^\perp) \subseteq W^\perp$.

Dimostrazione:

Essendo f isomorfismo, ho che $f(W) = W$.

Allora, se $w \in W$, $\exists y \in W \mid f(y) = w$.

La tesi è mostrare che $\forall x \in W^\perp$, $f(x) \in W^\perp$, cioè che $\forall x \in W^\perp$, $\forall w \in W$, $\phi(f(x), w) = 0$.

Ora:

$\phi(f(x), w) = \phi(f(x), f(y)) = \phi(x, y) = 0$, da cui la tesi.

Osservazione: Sia $\phi \in PS(V)$ non degenera. Se $v \in V$ è non isotropo, allora:

$$V = \text{Span}(v) \oplus \text{Span}(v)^\perp$$

$\forall w \in V$, $w = c(w) \cdot v + (w - c(w) \cdot v)$.

Se $c(w) = \frac{\phi(w, v)}{\phi(v, v)}$, allora $w - c(w) \cdot v \in \text{Span}(v)^\perp$.

Poniamo $z(w) = w - c(w) \cdot v$. Allora w si scrive in modo unico come $w = c(w) \cdot v + z(w)$, con $c(w) \cdot v \in \text{Span}(v)$ e $w - c(w) \cdot v \in \text{Span}(v)^\perp$.

DEFINIZIONE 4.3.1: Definiamo **riflessione parallela a un vettore v** l'applicazione

$$\rho_v: V \rightarrow V \mid \rho_v(w) = \rho(c(w) \cdot v + z(w)) = -c(w) \cdot v + z(w).$$

DEFINIZIONE 4.3.2: Definiamo **luogo dei punti fissi** di un'isometria f l'insieme

$$\text{Fix}(f) = \{v \in V \mid f(v) = v\}.$$

PROPOSIZIONE 4.3.2: $\rho_v^2 = id$ e $\rho_v \in O(V, \phi)$.

Dimostrazione:

Che $\rho_v^2 = id$ è evidente; inoltre ρ_v è sicuramente un isomorfismo.

Con una semplice verifica si vede che $\phi(\rho_v(w_1), \rho_v(w_2)) = \phi(w_1, w_2)$.

Osservazione: $\rho_v(v) = -v$ e $\text{Fix}(\rho_v) = \text{Span}(v)^\perp$.

TEOREMA 4.3.3: $\phi \in PS(V)$ non degenera. Allora $O(V, \phi)$ è generato dalle riflessioni, cioè ogni $f \in O(V, \phi)$ è composizione di un numero finito di riflessioni parallele a vettori non isotropi.

(Osservazione: Conveniamo che id è composizione di 0 riflessioni).

Dimostrazione:

Per induzione su $n = \dim(V)$:

Passo base): $n = 1$: $V = \text{Span}(v)$, con v non isotropo perché ϕ non degenera.

Sia $f \in O(V, \phi)$. Allora $f(v) = \lambda v$, $\lambda \neq 0$.

$$\phi(f(v), f(v)) = \lambda^2 \phi(v, v) \Rightarrow \lambda = \pm 1.$$

Se $\lambda = 1 \Rightarrow f = id$, che è composizione di 0 riflessioni.

Se $\lambda = -1 \Rightarrow f(v) = -v = \rho_v(v) \Rightarrow f = \rho_v$, poiché coincidono su una base.

Passo induttivo): Sia $w \in V$ non isotropo.

Caso 1): $f(w) = w$.

In questo caso per la prima proposizione si ha che $f(Z_w) = Z_w$, con

$Z_w = \text{Span}(w)^\perp$, quindi posso applicare l'ipotesi induttiva a $f|_{Z_w}$ e $\phi|_{Z_w}$. Allora $\exists \widetilde{\rho}_1, \dots, \widetilde{\rho}_k$ riflessioni di Z_w | $f|_{Z_w} = \widetilde{\rho}_1 \circ \dots \circ \widetilde{\rho}_k$.

Ogni $\widetilde{\rho}_i$ si estende ad una riflessione ρ_i di V (parallela allo stesso vettore) ponendo $\rho_i(w) = w$. Allora $f = \rho_1 \circ \dots \circ \rho_k$, da cui la tesi.

Caso 2): $f(w) \neq w$.

Notiamo che $w = \frac{f(w)+w}{2} - \frac{f(w)-w}{2}$. Inoltre $f(w) + w$ e $f(w) - w$ sono ortogonali e non contemporaneamente isotropi, infatti:

$$\phi(f(w) + w, f(w) + w) = 2\phi(w, w) + 2\phi(f(w), w)$$

$$\phi(f(w) - w, f(w) - w) = 2\phi(w, w) - 2\phi(f(w), w).$$

Se fossero entrambi isotropi, sommando avremmo che $4\phi(w, w) = 0$, cioè w isotropo, assurdo. Quindi:

Se $f(w) - w = u$ è non isotropo, allora:

$$\rho_u(w) = \rho_u \left(\underbrace{-\frac{f(w)-w}{2}}_{\in \text{Span}(u)} + \underbrace{\frac{f(w)+w}{2}}_{\in \text{Span}(u)^\perp} \right) = \frac{f(w)-w}{2} + \frac{f(w)+w}{2} = f(w),$$

quindi, applicando ρ_u a entrambi i membri, $(\rho_u \circ f)(w) = \rho_u^2(w) = w$.

Dunque w è punto fisso per $\rho_u \circ f$.

Allora, per il caso 1), $\exists \rho_1, \dots, \rho_k$ di V tali che

$$\rho_u \circ f = \rho_1 \circ \dots \circ \rho_k, \text{ e perciò } \rho_u^2 \circ f = f = \rho_u \circ \rho_1 \circ \dots \circ \rho_k.$$

Se invece è $f(w) + w = u$ a essere non isotropo, possiamo comunque ricondurci al caso precedente in quanto $-u = (-f)(w) - w$.

Dunque $\exists \rho_0, \dots, \rho_k$ riflessioni tali che

$$-f = \rho_0 \circ \dots \circ \rho_k.$$

Ma $f = (-f) \circ (-id)$, quindi se $\{v_1, \dots, v_n\}$ è una base ortogonale per V :

$$-id = \rho_{v_1} \circ \dots \circ \rho_{v_n},$$

poiché ciascuna riflessione cambia di segno la coordinata corrispondente.

Dunque $f = \rho_{v_1} \circ \dots \circ \rho_{v_n} \circ \rho_0 \circ \dots \circ \rho_k$, da cui la tesi.

LEMMA 4.3.4: (V, ϕ) anisotropo, $f \in O(V, \phi)$.

Se $\text{Fix}(f) = \{0\}$, allora $\exists \rho$ riflessione tale che $\dim(\text{Fix}(\rho \circ f)) = 1$.

Dimostrazione:

Sia $w \in V$, $w \neq 0$. Per ipotesi $f(w) \neq w$.

(V, ϕ) anisotropo $\Rightarrow u = f(w) - w$ è non isotropo.

Come provato prima, $(\rho_u \circ f)(w) = w$, cioè $w \in \text{Fix}(\rho_u \circ f)$ e quindi $\dim(\text{Fix}(\rho_u \circ f)) \geq 1$.

Proviamo che $\text{Fix}(\rho_u \circ f) \cap f^{-1}(Z_u) = \{0\}$, con $Z_u = \text{Span}(u)^\perp$.

Infatti $\rho_u|_{Z_u} = id$, quindi $(\rho_u \circ f)|_{f^{-1}(Z_u)} = f|_{f^{-1}(Z_u)}$ e poiché f non ha punti fissi, a maggior ragione non li ha $f|_{f^{-1}(Z_u)}$.

Perciò $\text{Fix}(\rho_u \circ f) \cap f^{-1}(Z_u) = \{0\}$, e poiché $\dim(f^{-1}(Z_u)) = n - 1$ (in quanto f è isomorfismo), per la formula di Grassmann $\dim(\text{Fix}(\rho_u \circ f)) \leq 1$, tesi.

TEOREMA 4.3.5: (V, ϕ) anisotropo, $\dim(V) = n$. Allora ogni $f \in O(V, \phi)$ è composizione di $n - k$ riflessioni, dove $k = \dim(\text{Fix}(f))$.

Dimostrazione:

Se $k = n$, allora $f = id$, quindi f è composizione di 0 riflessioni, e $n - k = 0$.

Se $k < n$, allora sia $H = \text{Fix}(f)^\perp$, così $V = \text{Fix}(f) \oplus^\perp H$.

Allora $\dim(H) = n - k$, H è f -invariante (poiché $f(\text{Fix}(f)) \subseteq \text{Fix}(f)$), $f|_H \in O(H, \phi|_H)$ e $O(H, \phi|_H)$ è anisotropo.

Inoltre evidentemente $\text{Fix}(f|_H) = \{0\}$.

Per il lemma $\exists u \in H$ tale che, detta $\widetilde{\rho}_u$ la riflessione in H parallela a u , si ha che

$$\dim(\text{Fix}(\widetilde{\rho}_u \circ f|_H)) = 1.$$

$\widetilde{\rho}_u$ si estende naturalmente alla riflessione ρ_u di V , ponendo $\rho_u(w) = w \ \forall w \notin H$.

Allora $\text{Fix}(f) \subseteq \text{Span}(u)^\perp = \text{Fix}(\rho_u)$, infatti:

$$v \in \text{Fix}(f) \Rightarrow v \notin H \Rightarrow \rho_u(v) = v \Rightarrow v \in \text{Fix}(\rho_u).$$

Quindi $\text{Fix}(f) \oplus \text{Fix}(\widetilde{\rho}_u \circ f|_H) \subseteq \text{Fix}(\rho_u \circ f)$, in quanto ovviamente $\text{Fix}(f) \subseteq \text{Fix}(\rho_u \circ f)$ e $\text{Fix}(\widetilde{\rho}_u \circ f|_H) \subseteq \text{Fix}(\rho_u \circ f)$ e $\text{Fix}(f) \cap \underbrace{\text{Fix}(\widetilde{\rho}_u \circ f|_H)}_{\subseteq H} \subseteq \text{Fix}(f) \cap H = \{0\}$.

Inoltre vale l'uguaglianza $\text{Fix}(f) \oplus \text{Fix}(\widetilde{\rho}_u \circ f|_H) = \text{Fix}(\rho_u \circ f)$, infatti sia $x \in \text{Fix}(\rho_u \circ f)$.

Scrivo $x = a + b$, con $a \in \text{Fix}(f)$ e $b \in H$; allora:

$$x = (\rho_u \circ f)(x) = (\rho_u \circ f)(a) + (\rho_u \circ f)(b) = a + (\rho_u \circ f)(b), \text{ ma per la scrittura unica}$$

$$b = (\rho_u \circ f)(b) \Rightarrow b \in \text{Fix}(\widetilde{\rho}_u \circ f|_H), \text{ in quanto } b \in H.$$

$$\text{Allora } \dim(\text{Fix}(\rho_u \circ f)) = k + 1.$$

Iterando, trovo che $\exists \rho_1, \dots, \rho_{n-k}$ riflessioni tali che $\dim(\text{Fix}(\rho_{n-k} \circ \dots \circ \rho_1 \circ f)) = n$, quindi $\rho_{n-k} \circ \dots \circ \rho_1 \circ f = \text{id}$, da cui $f = \rho_1 \circ \dots \circ \rho_{n-k}$.

4.4 AGGIUNTO

DEFINIZIONE 4.4.1: Sia $\phi \in PS(V)$. $\forall y \in V$ sia $\phi_y: V \rightarrow \mathbb{K} \mid v \rightarrow \phi_y(v) \stackrel{\text{def}}{=} \phi(v, y)$.

Osservazione: ϕ_y è lineare, quindi $\phi_y \in V^*$.

DEFINIZIONE 4.4.2: Definiamo $F_\phi: V \rightarrow V^* \mid y \rightarrow \phi_y$.

PROPOSIZIONE 4.4.1: 1) F_ϕ è lineare;

$$2) \text{Ker}(F_\phi) = \text{Rad}(\phi);$$

$$3) \text{Im}(F_\phi) = \text{Ann}(\text{Rad}(\phi));$$

$$4) F_\phi \text{ è un isomorfismo } \Leftrightarrow \phi \text{ è non degenere.}$$

Dimostrazione:

1) Semplice verifica.

$$2) y \in \text{Ker}(F_\phi) \Leftrightarrow \phi_y = 0 \Leftrightarrow \phi_y(x) = \phi(x, y) = 0 \ \forall x \in V \Leftrightarrow y \in \text{Rad}(\phi).$$

$$3) \text{Im}(F_\phi) \subseteq \text{Ann}(\text{Rad}(\phi)), \text{ infatti } \forall y \in V, F_\phi(y) = \phi_y \in \text{Ann}(\text{Rad}(\phi)), \text{ perché } \forall x \in \text{Rad}(\phi), \phi_y(x) = \phi(x, y) = 0.$$

$$\text{Inoltre } \dim(\text{Im}(F_\phi)) = \dim(V) - \dim(\text{Ker}(F_\phi)) = \dim(V) - \dim(\text{Rad}(\phi)) = \dim(\text{Ann}(\text{Rad}(\phi))).$$

$$4) \text{Ovvia per quanto visto nei punti precedenti (infatti se } \phi \text{ non degenere, allora } \text{Ker}(F_\phi) = \text{Rad}(\phi) = \{0\} \text{ e } \text{Im}(F_\phi) = \text{Ann}(\text{Rad}(\phi)) = V).$$

DEFINIZIONE 4.4.3: $g \in V^*$ è detto **ϕ -rappresentabile** se $g \in \text{Im}(F_\phi)$ (ossia se $\exists y \in V$ tale che $g = F_\phi(y)$, ossia se $\exists y \in V$ tale che $g(x) = \phi(x, y) \forall x \in V$)

TEOREMA DI RAPPRESENTAZIONE DI RIESZ: Se ϕ è non degenere, ogni $g \in V^*$ è ϕ -rappresentabile in modo unico (cioè $\exists! y \in V \mid g(x) = \phi(x, y) \forall x \in V$).

Dimostrazione:

Segue dal fatto che F_ϕ è un isomorfismo.

Osservazione: Grazie alla teoria del duale, avevamo trovato un isomorfismo canonico fra V e V^{**} , mentre ora, grazie al teorema di rappresentazione di Riesz, abbiamo trovato un isomorfismo canonico F_ϕ fra V e V^* .

Osservazione: Sia W^* un sottospazio fissato di V^* .

W^* coincide con l'insieme dei funzionali ϕ -rappresentabili $\Leftrightarrow W^* = \text{Im}(F_\phi) = \text{Ann}(\text{Rad}(\phi))$.

Dunque se $W^* = \text{Ann}(S)$, cioè $S = \text{Ann}(W^*)$ (grazie all'isomorfismo canonico ψ_V), allora $W^* = \text{Im}(F_\phi) \Leftrightarrow S = \text{Rad}(\phi)$.

PROPOSIZIONE 4.4.2: Siano U, W sottospazi di V . Allora:

1) $\text{Ann}(U + W) = \text{Ann}(U) \cap \text{Ann}(W)$;

2) $\text{Ann}(U \cap W) = \text{Ann}(U) + \text{Ann}(W)$.

Dimostrazione:

1) \subseteq $U \subseteq U + W \Rightarrow \text{Ann}(U + W) \subseteq \text{Ann}(U)$ e analogamente $\text{Ann}(U + W) \subseteq \text{Ann}(W)$.

Dunque $\text{Ann}(U + W) \subseteq \text{Ann}(U) \cap \text{Ann}(W)$.

\supseteq Sia $f \in \text{Ann}(U) \cap \text{Ann}(W)$; $f(u + w) = f(u) + f(w) = 0 + 0 = 0 \forall u \in U, w \in W$, dunque $f \in \text{Ann}(U + W)$.

2) $\text{Ann}(\text{Ann}(U) + \text{Ann}(W)) = \text{Ann}(\text{Ann}(U)) \cap \text{Ann}(\text{Ann}(W)) = U \cap W$.

Passando all'annullatore:

$\text{Ann}(\text{Ann}(\text{Ann}(U) + \text{Ann}(W))) = \text{Ann}(U) + \text{Ann}(W) = \text{Ann}(U \cap W)$, da cui la tesi.

Esempio: $V = \mathbb{R}^5$. In V^* si considerino:

$f_1(x) = x_1 + x_2 + x_3 + x_4 + x_5$;

$f_2(x) = x_1 + x_2 - x_3 - x_4 - x_5$;

$f_3(x) = x_1 + x_2 + 3x_3 + 3x_4 + 3x_5$;

e sia $W^* = \text{Span}(f_1, f_2, f_3)$.

Si costruisca in V un prodotto scalare ϕ tale che W^* sia il sottospazio dei funzionali ϕ -rappresentabili.

Dimostrazione:

Cerco $\phi \in \text{PS}(V)$ tale che $W^* = \text{Ann}(\text{Rad}(\phi))$.

Vediamo che $(2f_1 - f_2)(x) = f_3(x) \forall x \in V$, quindi $W^* = \text{Span}(f_1, f_2)$.

Allora, grazie all'isomorfismo canonico ψ_V :

$$\begin{aligned} \text{Ann}(W^*) &= \text{Ann}(\text{Span}(f_1, f_2)) = \text{Ann}(\text{Span}(f_1) + \text{Span}(f_2)) = \text{Ann}(f_1) \cap \text{Ann}(f_2) = \\ &= \text{Ker}(f_1) \cap \text{Ker}(f_2) \end{aligned}$$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_1 + x_2 - x_3 - x_4 - x_5 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_3 + x_4 + x_5 = 0 \end{cases} \Leftrightarrow \begin{cases} x_3 = -x_4 - x_5 \\ x_1 = -x_2 \end{cases}, \text{ quindi:}$$

$$Sol = \left\{ \left(\begin{array}{c} -x_2 \\ x_2 \\ -x_4 - x_5 \\ x_4 \\ x_5 \end{array} \right) \middle| x_2, x_4, x_5 \in \mathbb{R} \right\}, \text{ cioè, se } W^* = Ann(S) \Rightarrow S = Ann(W^*), \text{ e:}$$

$$S = Span \left(\underbrace{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=v_1}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}}_{=v_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}}_{=v_3} \right).$$

Dunque $S = Ann(W^*) = Rad(\phi)$, perciò costruisco $\phi \in PS(V) \mid Rad(\phi) = Span(v_1, v_2, v_3)$.
 Completo v_1, v_2, v_3 a base $\mathcal{B} = \{v_1, v_2, v_3, w_1, w_2\}$ di V . Ad esempio $w_1 = e_1, w_2 = e_3$.
 Ottengo la tesi con:

$$\mathfrak{M}_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

PROPOSIZIONE 4.4.3: Sia U sottospazio di V . Se ϕ è non degenera, allora $F_{\phi}(U^{\perp}) = Ann(U)$.
 Dimostrazione:

Sicuramente $F_{\phi}(U^{\perp}) \subseteq Ann(U)$, poiché $\forall y \in U^{\perp}, \forall u \in U, \phi_y(u) = \phi(y, u) = 0$.

Inoltre $\dim(F_{\phi}(U^{\perp})) = \dim(U^{\perp}) = \dim(V) - \dim(U) = \dim(Ann(U))$, da cui la tesi.

PROPOSIZIONE 4.4.4: $\forall \phi \in PS(V)$ non degenera, $\forall f \in End(V), \exists f^*: V \rightarrow V$ tale che $\phi(f(x), y) = \phi(x, f^*(y)) \quad \forall x, y \in V$.

Dimostrazione:

Fissiamo $y \in V$.

Poniamo inoltre $g(x) = \phi(f(x), y)$; $g(x)$ è evidentemente lineare, cioè $g(x) \in V^*$.

Dunque per il teorema di rappresentazione di Riesz $\exists! w \in V \mid g(x) = \phi(x, w) \quad \forall x \in V$.

Quindi $\phi(f(x), y) = \phi(x, w) \quad \forall x \in V$.

Poiché w dipende da y , poniamo $w = f^*(y)$.

Dunque è ben definita l'applicazione $f^*: V \rightarrow V \mid \phi(f(x), y) = \phi(x, f^*(y)) \quad \forall x, y \in V$.

DEFINIZIONE 4.4.4: L'applicazione f^* trovata è detta l'**aggiunto** di f rispetto a ϕ .

PROPOSIZIONE 4.4.5: 1) f^* è lineare;

2) $f^{**} = f$, cioè è un'involuzione;

3) $Ker(f^*) = (Im(f))^{\perp}$;

4) $Im(f^*) = (Ker(f))^{\perp}$;

5) Se \mathcal{B} è base di $V, A = \mathfrak{M}_{\mathcal{B}}(f), A^* = \mathfrak{M}_{\mathcal{B}}(f^*), M = \mathfrak{M}_{\mathcal{B}}(\phi)$, allora $A^* = M^{-1t}AM$.

Dimostrazione:

1) $\forall x, y_1, y_2 \in V$ si ha:

$$\phi(x, f^*(y_1 + y_2)) = \phi(f(x), y_1 + y_2) = \phi(f(x), y_1) + \phi(f(x), y_2) =$$

$$= \phi(x, f^*(y_1)) + \phi(x, f^*(y_2)) = \phi(x, f^*(y_1) + f^*(y_2)).$$

$$\text{Quindi } \phi(x, f^*(y_1 + y_2) - f^*(y_1) - f^*(y_2)) = 0 \quad \forall x \in V \Rightarrow$$

$$\Rightarrow (x, f^*(y_1 + y_2) - f^*(y_1) - f^*(y_2)) \in \text{Rad}(\phi) = \{0\} \Rightarrow f^*(y_1 + y_2) = f^*(y_1) + f^*(y_2).$$

Analogamente per il multiplo \Rightarrow tesi.

$$2) \quad \forall x, y \in V, \phi(f(x), y) = \phi(x, f^*(y)) = \phi(f^*(y), x) = \phi(y, f^{**}(x)) = \phi(f^{**}(x), y).$$

Quindi come prima $f(x) - f^{**}(x) \in \text{Rad}(\phi) = \{0\}$, da cui la tesi.

$$3) \quad \subseteq) \text{ Sia } v \in \text{Ker}(f^*). \text{ Allora } \forall f(w) \in \text{Im}(f), \phi(v, f(w)) = \phi(f^*(v), w) = \phi(0, w) = 0.$$

$$\supseteq) \text{ Sia } v \in (\text{Im}(f))^\perp. \text{ Allora } \forall w \in V, \phi(f^*(v), w) = \phi(v, f(w)) = 0, \text{ quindi}$$

$$f^*(v) \in \text{Rad}(\phi) = \{0\}, \text{ da cui } v \in \text{Ker}(f^*)$$

$$4) \text{ Dalla 3) segue che } \text{Ker}(f^{**}) = (\text{Im}(f^*))^\perp; \text{ applicando l'ortogonale si ottiene la tesi.}$$

$$5) \text{ Poich\'e } \forall v, w \in V, \phi(f(v), w) = \phi(v, f^*(w)), \text{ allora:}$$

$${}^t X {}^t A M Y = {}^t X M A^* Y \quad \forall X, Y \in \mathbb{K}^n.$$

$$\text{Poich\'e vale } \forall X, Y \in \mathbb{K}^n, \text{ allora } {}^t A M = M A^*.$$

$$\text{Ma } \phi \text{ non degenerare } \Rightarrow M \text{ invertibile } \Rightarrow A^* = M^{-1} {}^t A M.$$

Osservazioni: 1) Se $\exists \mathcal{B}$ base ortonormale, $A^* = {}^t A$.

2) Il diagramma:

$$\begin{array}{ccc} V & \xrightarrow{f^*} & V \\ F_\phi \downarrow & & \downarrow F_\phi \\ V^* & \xrightarrow{{}^t f} & V^* \end{array}$$

commuta, cio\`e $F_\phi \circ f^* = {}^t f \circ F_\phi$. Infatti $\forall v, x \in V$:

$$(F_\phi \circ f^*)(v)(x) = F_\phi(f^*(v))(x) = \phi(f^*(v), x);$$

$$({}^t f \circ F_\phi)(v)(x) = ({}^t f(\phi_v))(x) = (\phi_v \circ f)(x) = \phi(f(x), v) = \phi(f^*(v), x).$$

Da questo si ricavano di nuovo i punti 3) e 4) della proposizione precedente, infatti:

$$\begin{aligned} \text{Ker}(f^*) &= \text{Ker}(F_\phi^{-1} \circ {}^t f \circ F_\phi) \equiv \text{Ker}(F_\phi^{-1} \circ {}^t f) \equiv F_\phi^{-1}(\text{Ker}({}^t f)) = F_\phi^{-1}(\text{Ann}(\text{Im}(f))) = \\ &= (\text{Im}(f))^\perp, \end{aligned}$$

dove i passaggi contrassegnati con \equiv derivano dal fatto che F_ϕ \u00e8 un isomorfismo.

Analogamente per $\text{Im}(f^*)$.

$$3) \quad \Psi: (x, y) \rightarrow \phi(f(x), y) \text{ \u00e8 un prodotto scalare } \Leftrightarrow \forall x, y, \phi(f(x), y) = \phi(f(y), x),$$

$$\text{cio\`e } \Leftrightarrow \phi(x, f^*(y)) = \phi(x, f(y)), \text{ ossia } \Leftrightarrow f^*(y) - f(y) \in \text{Rad}(\phi) = \{0\} \Leftrightarrow f^* = f.$$

DEFINIZIONE 4.4.5: $f \in \text{End}(V)$ si dice **autoaggiunto** $\Leftrightarrow f = f^*$.

Osservazione: Supponiamo che $\exists \mathcal{B}$ base di V ortonormale. Allora:

$$f \text{ \u00e8 autoaggiunto } \Leftrightarrow \mathfrak{M}_{\mathcal{B}}(f) = {}^t(\mathfrak{M}_{\mathcal{B}}(f)) \Leftrightarrow \mathfrak{M}_{\mathcal{B}}(f) \in \mathcal{S}(n, \mathbb{K}).$$

4.5 SPAZI EUCLIDEI

DEFINIZIONE 4.5.1: (V, ϕ) si dice **spazio euclideo** se V è un \mathbb{R} -spazio vettoriale e ϕ è definito positivo.

Esempi: 1) $(\mathbb{R}^n, \langle, \rangle)$, con $\langle X, Y \rangle = {}^tXY$, cioè il prodotto scalare standard su \mathbb{R}^n ;

2) $(\mathcal{M}(n, \mathbb{R}), \phi)$, con $\phi(A, B) = \text{tr}({}^tAB)$.

Infatti $\phi(A, A) = \text{tr}({}^tAA) = \sum_{i=1}^n \sum_{j=1}^n [A]_{ji}^2 > 0$ se $A \neq 0$.

DEFINIZIONE 4.5.2: Sia (V, ϕ) spazio euclideo. Si definisce **norma** su V l'applicazione

$$\| \cdot \|: V \rightarrow \mathbb{R} \quad \|v\| = \sqrt{\phi(v, v)}.$$

PROPOSIZIONE 4.5.1: Valgono le seguenti proprietà:

- 1) $\|v\| \geq 0 \quad \forall v \in V$ e $\|v\| = 0 \Leftrightarrow v = 0$;
- 2) $\forall \lambda \in \mathbb{R}, \forall v \in V, \|\lambda v\| = |\lambda| \cdot \|v\|$;
- 3) $\forall v, w \in V, |\phi(v, w)| \leq \|v\| \cdot \|w\|$ (disuguaglianza di Schwarz);
- 4) $\forall v, w \in V, \|v + w\| \leq \|v\| + \|w\|$ (disuguaglianza triangolare).

Dimostrazione:

1) Ovvvia.

2) $\|\lambda v\| = \sqrt{\phi(\lambda v, \lambda v)} = \sqrt{\lambda^2 \phi(v, v)} = |\lambda| \cdot \|v\|.$

3) Se $v = 0 \vee w = 0$ la tesi è ovvia.

Altrimenti $\phi(tv + w, tv + w) \geq 0 \quad \forall t \in \mathbb{R}$, cioè $t^2 \phi(v, v) + 2t\phi(v, w) + \phi(w, w) \geq 0 \quad \forall t$.

Quindi il discriminante dell'equazione di secondo grado è ≤ 0 , cioè:

$$\phi(v, w)^2 - \phi(v, v) \cdot \phi(w, w) \leq 0, \text{ da cui la tesi.}$$

4) $\|v + w\|^2 = \phi(v + w, v + w) = \phi(v, v) + 2\phi(v, w) + \phi(w, w) \leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2,$

dove il passaggio contrassegnato con \leq segue da 3).

DEFINIZIONE 4.5.3: Una applicazione $d: V \times V \rightarrow \mathbb{R}$ si dice **distanza** se verifica le proprietà:

- 1) $d(x, y) \geq 0 \quad \forall x, y \in V$;
- 2) $d(x, y) = 0 \Leftrightarrow x = y$;
- 3) $d(x, y) = d(y, x) \quad \forall x, y \in V$;
- 4) $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in V$.

DEFINIZIONE 4.5.4: Definiamo $d: V \times V \rightarrow \mathbb{R} \mid d(x, y) = \|x - y\|$.

Osservazione: L'applicazione d appena definita è una distanza.

Osservazione: In uno spazio euclideo non esistono vettori isotropi non nulli. Già sappiamo che in ogni spazio euclideo esistono basi ortonormali. Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è ortonormale, allora \mathcal{B} induce una isometria $[\]_{\mathcal{B}}: (V, \phi) \rightarrow (\mathbb{R}^n, \langle, \rangle)$ data da:

$$v \rightarrow [v]_{\mathcal{B}} = \begin{pmatrix} \phi(v, v_1) \\ \vdots \\ \phi(v, v_n) \end{pmatrix}.$$

Osservazione: Per trovare una base ortonormale in uno spazio euclideo si può procedere applicando a una qualsiasi base dello spazio l'algoritmo di Lagrange e poi normalizzare. Esiste però un metodo più rapido:

METODO DI ORTONORMALIZZAZIONE DI GRAM-SCHMIDT: \mathcal{B} base di (V, ϕ) euclideo.

Sia $v'_1 = v_1$ e $v'_2 = v_2 - \frac{\phi(v_2, v_1)}{\phi(v_1, v_1)} v_1$. Allora $\phi(v'_1, v'_2) = 0$.

Inoltre $\text{Span}(v_1, v_2) = \text{Span}(v'_1, v'_2)$ e v'_1, v'_2 sono linearmente indipendenti.

Ora cerco v'_3 in modo che $\text{Span}(v_1, v_2, v_3) = \text{Span}(v'_1, v'_2, v'_3)$ e $\{v'_1, v'_2, v'_3\}$ sia base ortogonale di $\text{Span}(v_1, v_2, v_3)$.

Notiamo che un tale v'_3 si ottiene sottraendo a v_3 la sua proiezione ortogonale su $\text{Span}(v'_1, v'_2) = V'_2$.

Poiché v'_1 e v'_2 sono ortogonali, conosciamo già l'espressione analitica della proiezione ortogonale su V'_2 :

$$\pi_{V'_2}: V \rightarrow V'_2 \mid x \rightarrow \frac{\phi(x, v'_1)}{\phi(v'_1, v'_1)} v'_1 + \frac{\phi(x, v'_2)}{\phi(v'_2, v'_2)} v'_2.$$

Dunque basta porre:

$$v'_3 = v_3 - \frac{\phi(v_3, v'_1)}{\phi(v'_1, v'_1)} v'_1 - \frac{\phi(v_3, v'_2)}{\phi(v'_2, v'_2)} v'_2.$$

$v'_3 \in \text{Span}(v'_1, v'_2, v'_3) = \text{Span}(v_1, v_2, v_3)$.

Inoltre la matrice che contiene per colonna le coordinate di v'_1, v'_2, v'_3 rispetto a $\{v_1, v_2, v_3\}$ è:

$$M = \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix}, \text{ dunque } \{v'_1, v'_2, v'_3\} \text{ è una base ortogonale di } \text{Span}(v_1, v_2, v_3).$$

Itero il procedimento ponendo $\forall j$:

$$v'_j = v_j - \sum_{i=1}^{j-1} \frac{\phi(v_j, v'_i)}{\phi(v'_i, v'_i)} v'_i$$

Alla fine ottengo una base ortogonale $\{v'_1, \dots, v'_n\}$; basta normalizzare ponendo $w_i = \frac{v'_i}{\|v'_i\|}$.

Abbiamo così ottenuto con una procedura algoritmica il seguente risultato:

TEOREMA DI ORTONORMALIZZAZIONE DI GRAM-SCHMIDT: (V, ϕ) spazio euclideo, $\{v_1, \dots, v_n\}$ base di V . Allora esiste una base ortonormale $\{w_1, \dots, w_n\}$ di V tale che $\text{Span}(w_1, \dots, w_j) = \text{Span}(v_1, \dots, v_j) \quad \forall j$.

COROLLARIO 4.5.2: (V, ϕ) euclideo, $f \in \text{End}(V)$ triangolabile.

Allora esiste \mathcal{B} base di V ortonormale e a bandiera per f .

Dimostrazione:

$\exists \mathcal{S} = \{v_1, \dots, v_n\}$ base a bandiera per f .

Applico Gram-Schmidt a \mathcal{S} e trovo $\mathcal{B} = \{w_1, \dots, w_n\}$ ortonormale.

Poiché $\forall j, \text{Span}(w_1, \dots, w_j) = \text{Span}(v_1, \dots, v_j)$, \mathcal{B} è a bandiera per f .

Osservazione: $f \in \text{End}(V)$, con (V, ϕ) spazio euclideo. Se $\phi(f(x), f(y)) = \phi(x, y) \forall x, y \in V$, allora f è un isomorfismo.

Dimostrazione:

Se $x \in \text{Ker}(f)$, $\phi(x, x) = \phi(f(x), f(x)) = \phi(0, 0) = 0$, dunque $x = 0$.

Poiché f è lineare e iniettiva, allora è un isomorfismo.

Dunque $f \in O(V, \phi) \Leftrightarrow \phi(f(x), f(y)) = \phi(x, y) \forall x, y \in V$ (se $\phi \in \text{End}(V)$ e V euclideo).

ISOMETRIE DI UNO SPAZIO EUCLIDEO: (V, ϕ) euclideo, $f: V \rightarrow V$. Sono fatti equivalenti:

- 1) $f \in O(V, \phi)$
- 2) $f \in \text{End}(V)$ e $\|f(v)\| = \|v\| \forall v \in V$
- 3) $f(0) = 0$ e $d(f(x), f(y)) = d(x, y) \forall x, y \in V$
- 4) $f \in \text{End}(V)$ e $\forall \{v_1, \dots, v_n\}$ base ortonormale di V , $\{f(v_1), \dots, f(v_n)\}$ è base ortonormale di V
- 5) $f \in \text{End}(V)$ e $\exists \{v_1, \dots, v_n\}$ base ortonormale di V | $\{f(v_1), \dots, f(v_n)\}$ è base ortonormale di V
- 6) $f \in \text{End}(V)$ e $f^* \circ f = \text{id}$.

Dimostrazione:

1) \Rightarrow 2): Ovvvia perché $\|f(v)\|^2 = \phi(f(v), f(v)) = \phi(v, v) = \|v\|^2$.

2) \Rightarrow 1): Usando la formula di polarizzazione:

$$\begin{aligned} 2\phi(v, w) &= \phi(v+w, v+w) - \phi(v, v) - \phi(w, w) = \|v+w\|^2 - \|v\|^2 - \|w\|^2 = \\ &= \|f(v+w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2 = \phi(f(v+w), f(v+w)) - \phi(f(v), f(v)) - \\ &\quad - \phi(f(w), f(w)) = \phi(f(v)+f(w), f(v)+f(w)) - \phi(f(v), f(v)) - \phi(f(w), f(w)) = \\ &= 2\phi(f(v), f(w)). \end{aligned}$$

2) \Rightarrow 3): $d(f(x), f(y)) = \|f(x) - f(y)\| = \|f(x-y)\| = \|x-y\| = d(x, y)$.

3) \Rightarrow 1): • f conserva la norma:

$$\|f(x)\| = \|f(x) - 0\| = \|f(x) - f(0)\| = d(f(x), f(0)) = d(x, 0) = \|x\|.$$

• f conserva il prodotto scalare:

Per ipotesi $\forall x, y \in V$, $\|f(x) - f(y)\| = \|x - y\|$. Elevando al quadrato ho:

$$\begin{aligned} \phi(f(x) - f(y), f(x) - f(y)) &= \phi(x - y, x - y) \Rightarrow \\ \Rightarrow \underbrace{\|f(x)\|^2}_{=\|x\|^2} + \underbrace{\|f(y)\|^2}_{=\|y\|^2} - 2\phi(f(x), f(y)) &= \|x\|^2 + \|y\|^2 - 2\phi(x, y) \Rightarrow \\ \Rightarrow \phi(f(x), f(y)) &= \phi(x, y) \quad \forall x, y \in V. \end{aligned}$$

• f manda basi ortonormali in basi ortonormali:

Sia $\{v_1, \dots, v_n\}$ una base ortonormale di V . Poiché f conserva il prodotto scalare, l'insieme $\{f(v_1), \dots, f(v_n)\}$ è un insieme ortonormale (cioè sono a due a due ortogonali e hanno tutti norma 1).

Allora $f(v_1), \dots, f(v_n)$ sono linearmente indipendenti, infatti sia $\sum_{i=1}^n a_i f(v_i) = 0$.

$\forall j, 0 = \phi\left(\sum_{i=1}^n a_i f(v_i), f(v_j)\right) = \sum_{i=1}^n a_i \phi(f(v_i), f(v_j)) = a_j \phi(f(v_j), f(v_j)) = a_j$
dunque $\forall j, a_j = 0$, quindi $\{f(v_1), \dots, f(v_n)\}$ è una base ortonormale di V .

• f è lineare:

$\forall v \in V, v = \sum_{i=1}^n x_i v_i$, con $x_i = \phi(v, v_i)$ coefficiente di Fourier.

Ma abbiamo visto che anche $\{f(v_1), \dots, f(v_n)\}$ è una base ortonormale di V , quindi:

$$f(v) = \sum_{i=1}^n \phi(f(v), f(v_i)) f(v_i) = \sum_{i=1}^n \phi(v, v_i) f(v_i) = \sum_{i=1}^n x_i f(v_i),$$

dunque f è lineare.

• f è bigettiva in quanto è lineare e manda basi in basi.

Poiché f è isomorfismo e mantiene il prodotto scalare, allora $f \in O(V, \phi)$.

- 1) \Rightarrow 4): Ovvvia perché $\phi(f(v_i), f(v_j)) = \phi(v_i, v_j) = \delta_{ij}$ e f manda basi in basi.
 4) \Rightarrow 5): Ovvvia.
 5) \Rightarrow 1): Per ipotesi $\exists \mathcal{B} = \{v_1, \dots, v_n\}$ base ortonormale tale che $\{f(v_1), \dots, f(v_n)\}$ è base ortonormale.

Siano $v = \sum_{i=1}^n x_i v_i$ e $w = \sum_{i=1}^n y_i v_i$. Allora:

$$\begin{aligned} \phi(f(v), f(w)) &= \phi(f(\sum_{i=1}^n x_i v_i), f(\sum_{i=1}^n y_i v_i)) = \phi(\sum_{i=1}^n x_i f(v_i), \sum_{i=1}^n y_i f(v_i)) = \\ &= \sum_{i,j} x_i y_j \phi(f(v_i), f(v_j)) = \sum_{i,j} x_i y_j \delta_{ij} = \phi(v, w). \end{aligned}$$

- 1) \Leftrightarrow 6): $f \in O(V, \phi) \Leftrightarrow \phi(f(x), f(y)) = \phi(x, y) \forall x, y \in V$, ma
 $\phi(f(x), f(y)) = \phi(x, f^*(f(y)))$, quindi $f \in O(V, \phi) \Leftrightarrow \phi(x, f^*(f(y))) =$
 $= \phi(x, y) \forall x, y \in V \Leftrightarrow f^*(f(y)) - y \in \text{Rad}(\phi) = \{0\} \forall y \Leftrightarrow f^* \circ f = \text{id}.$

PROPOSIZIONE 4.5.3: \mathcal{B} base ortonormale di (V, ϕ) , $f \in \text{End}(V)$. Sia $A = \mathfrak{M}_{\mathcal{B}}(f)$.

Allora $f \in O(V, \phi) \Leftrightarrow {}^t A A = I$.

Dimostrazione:

Poiché $\mathfrak{M}_{\mathcal{B}}(\phi) = I$, $\phi(f(v), f(w)) = {}^t X^t A A Y \forall v, w \in V, X = [v]_{\mathcal{B}}, Y = [w]_{\mathcal{B}}$.

$\phi(v, w) = {}^t X Y$.

Quindi $f \in O(V, \phi) \Leftrightarrow {}^t X^t A A Y = {}^t X Y \forall X, Y \in \mathbb{K}^n \Leftrightarrow {}^t A A = I$.

DEFINIZIONE 4.5.5: $M \in \mathcal{M}(n, \mathbb{R})$ si dice **ortogonale** se ${}^t M M = M^t M = I$.

Denotiamo con $O(n) = \{M \in \mathcal{M}(n, \mathbb{R}) \mid M \text{ è ortogonale}\}$.

Osservazioni: 1) $M \in O(n) \Rightarrow M$ invertibile e $M^{-1} = {}^t M$

2) $M \in O(n) \Leftrightarrow$ le righe (e le colonne) di M formano una base ortonormale di \mathbb{R}^n . Infatti:

\Rightarrow $[{}^t M M]_{ij} = [I]_{ij} = \delta_{ij}$, dunque se $v_j = M^j$, allora $\phi(v_i, v_j) = \delta_{ij}$

\Leftarrow Se $v_j = M^j$ e $\mathcal{B} = \{v_1, \dots, v_n\} \Rightarrow \mathfrak{M}_{\mathcal{B}}(\phi) = I \in O(n)$.

3) $O(n)$ dotato del prodotto è un gruppo, detto **gruppo ortogonale**

4) Se $A \in O(n) \Rightarrow \det(A) = \pm 1$ (poiché $\det({}^t A A) = \det(I) = 1 \Rightarrow (\det(A))^2 = 1$).

In particolare denotiamo con $SO(n) = \{A \in O(n) \mid \det(A) = 1\}$ il **gruppo ortogonale speciale** (che è un gruppo con il prodotto).

5) Nel caso $n = 2$:

$$O(2) = \underbrace{\left\{ \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}}_{=SO(2), \text{non diagonalizzabili (rotazioni)}} \cup \underbrace{\left\{ \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}}_{\det=-1, \text{diagonalizzabili (riflessioni)}}$$

Dimostreremo fra poco che in $O(2)$ esistono solo questi due tipi di matrici.

PROPOSIZIONE 4.5.4: $A \in O(n)$, $\lambda \in \mathbb{C}$ autovalore per $A \Rightarrow |\lambda| = 1$.

(Osservazione: $|a + bi| = \sqrt{a^2 + b^2} \Rightarrow$ se $\lambda \in \mathbb{R}$ autovalore per A , allora $\lambda = \pm 1$).

Dimostrazione:

Pensiamo $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$.

Sia $X \in \mathbb{C}^n \setminus \{0\}$ autovettore per A relativo a λ , cioè $AX = \lambda X$.

Poiché A è reale, allora $\overline{AX} = \overline{A} \overline{X} = A \overline{X} = \overline{\lambda} \overline{X}$.

$${}^t (\lambda X) \overline{\lambda X} = \lambda \overline{\lambda} {}^t X \overline{X} \\ {}^t (AX) A \overline{X} = {}^t X^t A A \overline{X} = {}^t X \overline{X}$$

Dunque $(\lambda\bar{\lambda} - 1)^t X\bar{X} = 0$. Ma $\lambda\bar{\lambda} = |\lambda|^2$, perciò:

$$(|\lambda|^2 - 1)^t X\bar{X} = 0.$$

Ora ${}^t X\bar{X} = x_1\bar{x}_1 + \dots + x_n\bar{x}_n = |x_1|^2 + \dots + |x_n|^2 \in \mathbb{R}^+$, poiché $X \neq 0$.

Allora $|\lambda|^2 = 1 \Rightarrow |\lambda| = 1$.

PROPOSIZIONE 4.5.5: (V, ϕ) euclideo, $\mathcal{B} = \{v_1, \dots, v_n\}$ base ortonormale di V .

$\mathcal{B}' = \{w_1, \dots, w_n\}$ base di V ; $M = \mathfrak{M}_{\mathcal{B}', \mathcal{B}}(id)$. Allora:

\mathcal{B}' è ortonormale $\Leftrightarrow M$ è ortogonale.

Dimostrazione:

$$\forall i, [w_i]_{\mathcal{B}} = M[w_i]_{\mathcal{B}'} = M^i.$$

$$\phi(w_i, w_j) = {}^t [w_i]_{\mathcal{B}} [w_j]_{\mathcal{B}} = {}^t (M^i) M^j = ({}^t M)_i M^j = [{}^t M M]_{ij}.$$

Dunque \mathcal{B}' è ortonormale $\Leftrightarrow \phi(w_i, w_j) = \delta_{ij} \forall i, j \Leftrightarrow [{}^t M M]_{ij} = \delta_{ij} \forall i, j \Leftrightarrow M \in O(n)$.

Dunque possiamo migliorare il teorema di triangolabilità per matrici:

PROPOSIZIONE 4.5.6: $A \in \mathcal{M}(n, \mathbb{R})$ triangolabile. Allora $\exists M \in O(n) \mid M^{-1} A M = T$ triangolare.

Dimostrazione:

Interpretando $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, allora $\mathfrak{M}_{\mathcal{C}}(A) = A$, dove \mathcal{C} è la base canonica di \mathbb{R}^n .

Come conseguenza dell'algoritmo di Gram-Schmidt, $\exists \mathcal{B}$ base di \mathbb{R}^n ortonormale e a bandiera per A .

Sia $M = \mathfrak{M}_{\mathcal{B}, \mathcal{C}}(id)$; allora $M^{-1} A M = \mathfrak{M}_{\mathcal{B}}(A) = T$ triangolare.

Poiché \mathcal{B} e \mathcal{C} sono basi ortonormali, $M \in O(n)$.

Osservazione: Se \mathcal{B} è una base ortonormale, la restrizione di $\mathfrak{M}_{\mathcal{B}}: GL(V) \rightarrow GL(n, \mathbb{R})$ a $O(V, \phi)$ identifica $O(V, \phi)$ con il sottogruppo $O(n)$ di $GL(n, \mathbb{R})$.

$$\mathfrak{M}_{\mathcal{B}}: O(V, \phi) \rightarrow O(n) \mid f \rightarrow \mathfrak{M}_{\mathcal{B}}(f)$$

Per quanto visto, se f è isometria, allora:

- $\det(f) = \pm 1$;
- gli unici autovalori reali di f sono ± 1 .

Denotiamo con $SO(V, \phi) = \{f \in O(V, \phi) \mid \det(f) = 1\}$ il **gruppo ortogonale speciale**.

Gli elementi di $SO(V, \phi)$ sono detti **isometrie dirette** o **rotazioni**.

$f \in O(V, \phi) \setminus SO(V, \phi)$ si dice **isometria inversa**.

Osservazione: Lavoriamo in $(\mathbb{R}^n, \langle, \rangle)$.

Ogni isometria di \mathbb{R}^n è composizione di al più n riflessioni.

Ogni isometria diretta è composizione di un numero pari di riflessioni, poiché se ρ è una riflessione, $\det(\rho) = -1$, quindi $\det(\rho_1 \circ \dots \circ \rho_k) = 1 \Leftrightarrow k$ pari.

Analogamente le isometrie inverse sono composizione di un numero dispari di riflessioni.

Se $f \in O(\mathbb{R}^n, \langle, \rangle)$ e $\dim(\text{Fix}(f)) = k$, allora f è composizione di $n - k$ riflessioni.

PROPOSIZIONE 4.5.7: Ogni matrice in $O(2)$ è della forma:

$$R_{\theta} = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R} \right\} \vee \rho_{\theta} = \left\{ \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R} \right\}.$$

Dimostrazione:

Se $A = (v_1 | v_2)$, $\{v_1, v_2\}$ è base ortonormale di \mathbb{R}^2 .

In particolare, se $v_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, allora $\langle v_1, v_1 \rangle = x_1^2 + x_2^2 = 1$, quindi $\exists \theta_1 | x_1 = \cos(\theta_1)$, $x_2 = \sin(\theta_1)$.

Analogamente $v_2 = \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix}$.

Ma:

$$0 = \langle v_1, v_2 \rangle = \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2) = \cos(\theta_1 - \theta_2).$$

Dunque $\theta_2 = \theta_1 + \frac{\pi}{2} + 2k\pi \vee \theta_2 = \theta_1 - \frac{\pi}{2} + 2k\pi$.

Se vale la prima, allora $\begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} = \begin{pmatrix} -\sin(\theta_1) \\ \cos(\theta_1) \end{pmatrix}$.

Se vale la seconda, allora $\begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} = \begin{pmatrix} \sin(\theta_1) \\ -\cos(\theta_1) \end{pmatrix}$, tesi.

PROPOSIZIONE 4.5.8 (Forma canonica per un'isometria in uno spazio euclideo): Sia $A \in O(n)$. Allora $\exists M \in O(n)$ tale che:

$${}^t M A M = M^{-1} A M = \begin{pmatrix} \boxed{I} & & & & \\ & \boxed{-I} & & & \\ & & \boxed{R_{\theta_1}} & & \\ & & & \ddots & \\ & & & & \boxed{R_{\theta_k}} \end{pmatrix},$$

con $R_{\theta_i} = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$, con $\theta_i \neq k\pi \forall i$.

FORMULAZIONE ALTERNATIVA: Sia (V, ϕ) uno spazio euclideo di dimensione n . Sia $\psi \in O(V, \phi)$. Allora $\exists \mathcal{B}$ base ortonormale di V tale che:

$$\mathfrak{M}_{\mathcal{B}}(\psi) = \begin{pmatrix} \boxed{I} & & & & \\ & \boxed{-I} & & & \\ & & \boxed{R_{\theta_1}} & & \\ & & & \ddots & \\ & & & & \boxed{R_{\theta_k}} \end{pmatrix}.$$

(Osservazione: Le due formulazioni sono equivalenti, infatti:

Prima \Rightarrow Seconda: Scelgo \mathcal{B}' base ortonormale per (V, ϕ) . Allora $\mathfrak{M}_{\mathcal{B}'}(\phi) = I$.

$$\psi \in O(V, \phi) \Rightarrow {}^t(\mathfrak{M}_{\mathcal{B}'}(\psi))\mathfrak{M}_{\mathcal{B}'}(\phi)\mathfrak{M}_{\mathcal{B}'}(\psi) = \mathfrak{M}_{\mathcal{B}'}(\phi), \text{ cioè}$$

$${}^t(\mathfrak{M}_{\mathcal{B}'}(\psi))\mathfrak{M}_{\mathcal{B}'}(\psi) = I.$$

Dunque per la prima formulazione $\exists M \in O(n)$ tale che:

$${}^t M \mathfrak{M}_{\mathcal{B}'}(\psi) M = \begin{pmatrix} \boxed{I} & & & & \\ & \boxed{-I} & & & \\ & & \boxed{R_{\theta_1}} & & \\ & & & \ddots & \\ & & & & \boxed{R_{\theta_k}} \end{pmatrix}$$

Ma $M \in O(n)$, per cui rappresenta un cambio di base da \mathcal{B}' a un'altra base ortonormale, cioè $M = \mathfrak{M}_{\mathcal{B}, \mathcal{B}'}(id)$, dunque ${}^t M \mathfrak{M}_{\mathcal{B}'}(\psi) M = \mathfrak{M}_{\mathcal{B}}(\psi)$, tesi.

Il viceversa è analogo.)

Dimostrazione:

Per induzione su n :

Passo base): $n = 1$: Basta osservare che un autovalore reale di A può essere solo ± 1 .

Passo induttivo): $n \geq 2$.

Caso a) A ammette l'autovalore reale $\lambda = \pm 1$.

Sia v un autovettore relativo a λ . Allora:

$\mathbb{R}^n = \text{Span}(v) \oplus \text{Span}(v)^\perp$, $Av \in \text{Span}(v) \Rightarrow$ (poiché $A \in O(n)$) $\Rightarrow Av^\perp \subseteq v^\perp$, dunque $A|_{v^\perp} \in O(v^\perp, \langle, \rangle|_{v^\perp})$.

Per ipotesi induttiva della seconda formulazione $\exists \mathcal{B}$ base ortonormale di v^\perp | $\mathfrak{M}_{\mathcal{B}}(A|_{v^\perp}) = C$ della forma voluta.

Se scelgo v normalizzato, $\mathcal{B}^* = \{v\} \cup \mathcal{B}$ è una base ortonormale di \mathbb{R}^n

e $\mathfrak{M}_{\mathcal{B}^*}(A) = \left(\begin{array}{c|c} \pm 1 & 0 \\ \hline 0 & C \end{array} \right)$, dunque a meno di permutare i vettori di \mathcal{B}^*

ho la tesi.

Caso b) A non ha autovalori reali. Sia λ un autovalore complesso.

$|\lambda| = 1 \Rightarrow \lambda = \cos(\theta) + i \sin(\theta)$. Sia $v = v_{\mathbb{R}} + i v_{\mathbb{I}} \in \mathbb{C}^n$ un autovettore per λ ($v_{\mathbb{R}}, v_{\mathbb{I}} \in \mathbb{R}^n$).

So dalla forma di Jordan reale che $v_{\mathbb{R}}, v_{\mathbb{I}}$ sono indipendenti.

Mostro che $\|v_{\mathbb{R}}\| = \|v_{\mathbb{I}}\|$ e che $\langle v_{\mathbb{R}}, v_{\mathbb{I}} \rangle = 0$.

Se denoto con \langle, \rangle il prodotto scalare standard su \mathbb{C}^n ,

${}^tAA = I \Rightarrow A \in O(\mathbb{C}^n, \langle, \rangle)$, dunque:

$\langle v, v \rangle = \langle Av, Av \rangle = \langle \lambda v, \lambda v \rangle = \lambda^2 \langle v, v \rangle$.

Ma $\lambda^2 \neq 1$ e dunque $\langle v, v \rangle = 0$, cioè:

$0 = \langle v_{\mathbb{R}} + i v_{\mathbb{I}}, v_{\mathbb{R}} + i v_{\mathbb{I}} \rangle = \langle v_{\mathbb{R}}, v_{\mathbb{R}} \rangle - \langle v_{\mathbb{I}}, v_{\mathbb{I}} \rangle + 2i \langle v_{\mathbb{I}}, v_{\mathbb{R}} \rangle \Rightarrow$
 \Rightarrow uguagliando a 0 parte reale e immaginaria:

$\langle v_{\mathbb{R}}, v_{\mathbb{R}} \rangle = \langle v_{\mathbb{I}}, v_{\mathbb{I}} \rangle \Rightarrow \|v_{\mathbb{R}}\| = \|v_{\mathbb{I}}\|$ e $\langle v_{\mathbb{R}}, v_{\mathbb{I}} \rangle = 0$.

A meno di riscalarlo v , posso supporre che $v_{\mathbb{R}}, v_{\mathbb{I}}$ siano ortonormali.

Dalla forma di Jordan reale so che $A(\text{Span}(v_{\mathbb{R}}, v_{\mathbb{I}})) \subseteq \text{Span}(v_{\mathbb{R}}, v_{\mathbb{I}})$ e

$A|_{\text{Span}(v_{\mathbb{R}}, v_{\mathbb{I}})}$ si rappresenta rispetto a $\{v_{\mathbb{R}}, v_{\mathbb{I}}\}$ come:

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Concludo come prima sfruttando l'invarianza di $\text{Span}(v_{\mathbb{R}}, v_{\mathbb{I}})^\perp$.

4.6 IL TEOREMA SPETTRALE REALE

In tutto il capitolo lavoreremo in uno spazio euclideo (V, ϕ) .

DEFINIZIONE 4.6.1: $f \in \text{End}(V)$ si dice **ortogonalmente diagonalizzabile** se \exists una base \mathcal{B} di V ortonormale per ϕ e di autovettori per f (detta anche **base spettrale**).

Osservazione: Se f è ortogonalmente diagonalizzabile, allora f è autoaggiunto.

Infatti se $\exists \mathcal{B}$ base spettrale $\Rightarrow \mathfrak{M}_{\mathcal{B}}(f) = D$ diagonale. Ma poiché f è autoaggiunto $\Leftrightarrow \exists \mathcal{B}$ base di V ϕ -ortonormale | $\mathfrak{M}_{\mathcal{B}}(f)$ è simmetrica e D lo è, allora f è autoaggiunto.

LEMMA 4.6.1: $f \in \text{End}(V)$ autoaggiunto, W sottospazio di V .

Allora $f(W) \subseteq W \Rightarrow f(W^\perp) \subseteq W^\perp$.

Dimostrazione:

Dobbiamo mostrare che $\forall x \in W^\perp, f(x) \in W^\perp$.

$\forall w \in W, 0 = \phi(x, f(w)) = \phi(f(x), w)$, dunque segue la tesi.

LEMMA 4.6.2: $A \in \mathcal{S}(n, \mathbb{R})$. Allora $p_A(t)$ è completamente fattorizzabile in $\mathbb{R}[t]$.

Dimostrazione:

Consideriamo A come matrice complessa, $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$; sia λ autovalore per A .

Se $X \in \mathbb{C}^n \setminus \{0\}$ autovettore per A relativo a λ , allora $AX = \lambda X$, dunque $A\bar{X} = \bar{\lambda}\bar{X}$.

Allora:

$${}^t X A \bar{X} = \begin{cases} {}^t X (\bar{\lambda} \bar{X}) = \bar{\lambda} {}^t X \bar{X} \\ {}^t (AX) \bar{X} = {}^t (\lambda X) \bar{X} = \lambda {}^t X \bar{X} \end{cases}$$

Dunque $(\bar{\lambda} - \lambda) {}^t X \bar{X} = 0$.

Ma ${}^t X \bar{X} = x_1 \bar{x}_1 + \dots + x_n \bar{x}_n = |x_1|^2 + \dots + |x_n|^2 \in \mathbb{R}^+$, poiché $X \neq 0$.

Dunque $\bar{\lambda} = \lambda \Rightarrow \lambda \in \mathbb{R}$.

TEOREMA SPETTRALE REALE: (V, ϕ) spazio euclideo.

$f \in \text{End}(V)$ è ortogonalmente diagonalizzabile $\Leftrightarrow f$ è autoaggiunto.

Dimostrazione 1:

\Rightarrow) Già vista.

\Leftarrow) Per induzione su $n = \dim(V)$:

Passo base): $n = 1$: ovvio.

Passo induttivo): Per il lemma $\exists \lambda \in \mathbb{R}$ autovalore per f .

Sia $v_1 \in V_\lambda$ di norma 1.

Allora $V = \text{Span}(v_1) \oplus \text{Span}(v_1)^\perp$.

Per il primo lemma, $f(\text{Span}(v_1)^\perp) \subseteq \text{Span}(v_1)^\perp$; inoltre $\dim(\text{Span}(v_1)^\perp) = n - 1$ e $f|_{\text{Span}(v_1)^\perp}$ è autoaggiunto.

Dunque per ipotesi induttiva, $\exists \{v_2, \dots, v_n\}$ base ortonormale di $\text{Span}(v_1)^\perp$ di autovettori per $f|_{\text{Span}(v_1)^\perp}$ (e quindi per f).

Allora $\{v_1, \dots, v_n\}$ è base spettrale per f , da cui la tesi.

Dimostrazione 2:

\Rightarrow) Già vista.

\Leftarrow) Sia \mathcal{S} base ϕ -ortonormale, $A = \mathfrak{M}_{\mathcal{S}}(f)$. Allora ${}^t A = A$.

Per il lemma A è triangolabile, in particolare $\exists P \in O(n) | P^{-1}AP = {}^t PAP = T$ triangolare.

A simmetrica $\Rightarrow T$ simmetrica (poiché congruente ad A) $\Rightarrow T$ è diagonale.

Sia \mathcal{B} base di $V | \mathfrak{M}_{\mathcal{B}, \mathcal{S}}(id) = P$.

$P \in O(n) \Rightarrow \mathcal{B}$ ortonormale. $\mathfrak{M}_{\mathcal{B}}(f) = T$ diagonale $\Rightarrow \mathcal{B}$ di autovettori.

COROLLARIO 4.6.3: $A \in \mathcal{S}(n, \mathbb{R})$. Allora $\exists P \in O(n) | P^{-1}AP = {}^t PAP = D$ diagonale.

Osservazione: Sia $A \in \mathcal{S}(n, \mathbb{R})$. In particolare abbiamo dimostrato che:

- A è simile a D tramite una matrice ortogonale, cioè $A \mathcal{R} D$ in $\mathcal{M}(n, \mathbb{R}) / \text{similit. tramite el. di } O(n)$, che corrisponde a $\text{End}(V) / \text{coniugio tramite el. di } O(V, \phi)$;

- A è congruente a D tramite matrice ortogonale, ossia $A \mathcal{R} D$ in $\mathcal{M}(n, \mathbb{R})$ /congruenza tramite el.di $O(n)$, che corrisponde a $PS(V)$ /isometria tramite el.di $O(V, \phi)$.
Dunque, se $P^{-1}AP = {}^tPAP = D$ diagonale, allora gli elementi sulla diagonale di D sono sia gli autovalori di A , sia permettono di calcolare $\sigma(A)$.

In particolare:

$$i_+(A) = \#\text{autovalori positivi di } A;$$

$$i_-(A) = \#\text{autovalori negativi di } A;$$

$$i_0(A) = \#\text{autovalori nulli di } A.$$

COROLLARIO 4.6.4: $A \in \mathcal{S}(n, \mathbb{R})$ è definita positiva \Leftrightarrow tutti gli autovalori di A sono positivi.

PROPOSIZIONE 4.6.5: (V, ϕ) euclideo, $f \in \text{End}(V)$, $f^* = f$, $\lambda \neq \mu$ autovalori per f .

Allora $V_\lambda \perp V_\mu$ (cioè $\phi(v, w) = 0 \forall v \in V_\lambda, w \in V_\mu$).

Dimostrazione:

$$\lambda\phi(v, w) = \phi(\lambda v, w) = \phi(f(v), w) = \phi(v, f(w)) = \phi(v, \mu w) = \mu\phi(v, w) \forall v \in V_\lambda, w \in V_\mu.$$

Dunque $(\lambda - \mu)\phi(v, w) = 0 \forall v \in V_\lambda, w \in V_\mu$, ma $\lambda \neq \mu$, da cui la tesi.

TEOREMA DI ORTOGONALIZZAZIONE SIMULTANEA: V \mathbb{R} -spazio vettoriale.

$\phi, \psi \in PS(V)$, ϕ definito positivo. Allora $\exists \mathcal{B}$ base di V ortonormale per ϕ e ortogonale per ψ .

Dimostrazione:

Sia \mathcal{S} base ortonormale per ϕ .

Sia $A = \mathfrak{M}_{\mathcal{S}}(\psi) \Rightarrow A$ è simmetrica.

Sia $g \in \text{End}(V) | \mathfrak{M}_{\mathcal{S}}(g) = A \Rightarrow g$ è autoaggiunto.

Per il teorema spettrale $\exists \mathcal{B}$ base ortonormale per ϕ e di autovettori per g , ossia $\mathfrak{M}_{\mathcal{B}}(g) = D$ diagonale.

Sia $M = \mathfrak{M}_{\mathcal{B}, \mathcal{S}}(id)$; poiché \mathcal{B} e \mathcal{S} sono ortonormali, allora $M \in O(n)$.

Inoltre $M^{-1}AM = D$, perciò $M^{-1}AM = {}^tMAM = D$.

Ma $\mathfrak{M}_{\mathcal{B}}(\psi) = {}^tMAM = D$, per cui la base \mathcal{B} è ortogonale per ψ .

PROPOSIZIONE 4.6.6: $A \in \mathcal{M}(n, \mathbb{R})$. A è simmetrica $\Leftrightarrow \begin{cases} A^tA = {}^tAA \\ A \text{ è triangolabile} \end{cases}$

Dimostrazione:

\Rightarrow) Ovvvia.

\Leftarrow) Come conseguenza di Gram-Schmidt $\exists P \in O(n) | P^{-1}AP = {}^tPAP = T$ triangolare superiore.

$$T^tT = {}^tPAP {}^tP^tAP = {}^tPA {}^tAP = {}^tP^tAAP = {}^tP^tAP {}^tPAP = {}^tTT.$$

Ma una matrice T triangolare superiore tale che ${}^tTT = T^tT$ è diagonale, infatti:

$$[T^tT]_{11} = T_1 \cdot T_1 = [T]_{11}^2 + \dots + [T]_{1n}^2$$

$$[{}^tTT]_{11} = [T]_{11}^2,$$

dunque $[T]_{1j} = 0 \forall 2 \leq j \leq n$.

Iterando, ottengo che T è diagonale.

Allora A è simmetrica perché congruente a T tramite $P \in O(n)$.

COROLLARIO 4.6.7: $A \in O(n)$. Se A ha tutti gli autovalori reali, allora è simmetrica (e dunque diagonalizzabile).

PROPOSIZIONE 4.6.8: $A \in \mathcal{S}(n, \mathbb{R})$. A è definita positiva $\Leftrightarrow \exists! S \in \mathcal{S}(n, \mathbb{R})$ definita positiva tale che $A = S^2$ (si dice che S è la **radice quadrata** di A).

Dimostrazione:

$\Leftarrow \forall X \in \mathbb{R}^n \setminus \{0\}$, ${}^tXAX = {}^tXS^2X = {}^tX^tSSX = {}^t(SX)SX > 0$, poiché $S \in GL(n, \mathbb{R})$, tesi.

\Rightarrow Esistenza: Per ipotesi gli autovalori di A sono reali positivi, quindi $\exists M \in O(n)$ tale che

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = D^2 \Rightarrow D = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}.$$

Allora $A = MD^2M^{-1} = (MD^tM)(MD^tM)$.

Basta prendere $S = MD^tM$ (che è simmetrica e definita positiva perché congruente a D simmetrica e definita positiva).

Unicità: Sia S una qualsiasi matrice simmetrica definita positiva tale che $A = S^2$.

Allora $SA = S \cdot S^2 = S^2 \cdot S = AS$.

Se $\mathbb{R}^n = V_{\lambda_1}(A) \oplus \dots \oplus V_{\lambda_k}(A)$, $SA = AS \Rightarrow S(V_{\lambda_j}(A)) \subseteq V_{\lambda_j}(A) \forall j$.

Vediamo ora che $S|_{V_{\lambda_j}(A)}$ è completamente determinato (e quindi S è unica).

Sia μ autovalore per $S|_{V_{\lambda_j}(A)}$ e v un autovettore relativo a μ , cioè $Sv = \mu v$.

$\lambda_j v = Av = S^2 v = \mu^2 v \Rightarrow \mu^2 = \lambda_j \Rightarrow \mu = \sqrt{\lambda_j} \Rightarrow S|_{V_{\lambda_j}(A)} = \sqrt{\lambda_j} \cdot id|_{V_{\lambda_j}(A)}$.

Dunque S è determinato univocamente, cioè ho la tesi.

Osservazione: Con la stessa dimostrazione si prova che A è semidefinito positivo $\Leftrightarrow \exists! S \in \mathcal{S}(n, \mathbb{R})$ semidefinita positiva tale che $A = S^2$.

PROPOSIZIONE 4.6.9: $A \in GL(n, \mathbb{R})$. Allora $\exists! S \in \mathcal{S}(n, \mathbb{R})$ definita positiva e $P \in O(n)$ | $A = SP$ (questa decomposizione prende il nome di **decomposizione polare**).

Dimostrazione:

Esistenza: A^tA è evidentemente simmetrica e definita positiva, in quanto ${}^tXA^tAX = {}^t({}^tAX)^tAX > 0$, poiché $A \in GL(n, \mathbb{R}) \Rightarrow {}^tA \in GL(n, \mathbb{R})$ (oppure tAA è definita positiva perché rappresenta l'identità in una base differente, cioè è una matrice congruente all'identità, che è definita positiva $\Rightarrow {}^tAA$ definita positiva $\Rightarrow A^tA$ è definita positiva).

Dunque per la proposizione precedente $\exists S$ simmetrica definita positiva | $A^tA = S^2$.

Quindi $S^{-1}A^tAS^{-1} = I \Rightarrow S^{-1}A^t(S^{-1}A) = I$, cioè $P = S^{-1}A \in O(n)$.

Inoltre $A = S(S^{-1}A) = SP$, da cui la tesi.

Unicità: Se $A = SP$, allora $A^tA = SP^tPS = S^2$, ma allora S è unica per la proposizione precedente. Poiché $P = S^{-1}A$, anche P è unica.

PROPOSIZIONE 4.6.10: Siano f, g endomorfismi autoaggiunti di uno spazio euclideo (V, ϕ) | $f \circ g = g \circ f$. Allora \exists base ortonormale di V fatta da autovettori sia per f che per g .

Dimostrazione:

f, g diagonalizzabili per il teorema spettrale. Inoltre $\forall \lambda$ autovalore per f , $V_\lambda(f)$ è g -invariante. $g|_{V_\lambda(f)}$ è autoaggiunto nello spazio euclideo $(V_\lambda(f), \phi|_{V_\lambda(f)})$, dunque $g|_{V_\lambda(f)}$ ammette una base \mathcal{B}_λ di autovettori ortogonale rispetto a $\phi|_{V_\lambda(f)}$. Al variare di $\lambda_i \in Sp(f)$, pongo $\mathcal{B} = \cup_i \mathcal{B}_{\lambda_i}$.

Ogni elemento di \mathcal{B} è di autovettori per f e per g ; ma f è autoaggiunto, per cui

$V = V_{\lambda_1}(f) \oplus^\perp \dots \oplus^\perp V_{\lambda_k}(f)$, quindi \mathcal{B} è ortonormale; da questo segue la tesi.

5 SPAZI AFFINI

5.1 ISOMETRIE AFFINI

Riprendiamo la notazione $S(X) = \{f: X \rightarrow X \text{ biunivoche}\}$, $X \neq \emptyset$.

DEFINIZIONE 5.1.1: Ogni sottogruppo di $S(X)$ si chiama **gruppo di trasformazioni** di X .

Esempi: 1) Se V è uno spazio vettoriale, con l'algebra lineare abbiamo studiato $GL(V)$ come gruppo di trasformazioni di V .

2) Allo stesso modo abbiamo studiato $O(V, \phi)$ se $\phi \in PS(V)$.

Osservazione: Non tutte le trasformazioni sono lineari, infatti:

DEFINIZIONE 5.1.2: V spazio vettoriale. La trasformazione $\tau_v: V \rightarrow V \mid \tau_v(w) = w + v$ è detta **traslazione** del vettore v fissato, $v \in V$.

Osservazione: τ_v è lineare $\Leftrightarrow v = 0$.

Notazione: Indicheremo con $T(V) = \{\text{traslazioni di } V\}$.

PROPOSIZIONE 5.1.1: $(T(V), \circ)$ è un gruppo abeliano di trasformazioni di V e $(T(V), \circ) \cong (V, +)$.

Dimostrazione:

$\tau_0 = id$, vale evidentemente la proprietà associativa e $(\tau_v)^{-1} = \tau_{-v}$.

Inoltre $\tau_v \circ \tau_w = \tau_{v+w} = \tau_{w+v} = \tau_w \circ \tau_v$.

Quindi $(T(V), \circ)$ è un gruppo abeliano.

Poiché $F: (V, +) \rightarrow (T(V), \circ) \mid F(v) = \tau_v$ è sicuramente un isomorfismo, segue anche l'altra tesi.

DEFINIZIONE 5.1.3: Sia G un gruppo. Si chiama **azione** di G su un insieme X ogni omomorfismo $\psi: G \rightarrow S(X)$.

L'azione di un gruppo si dice **transitiva** se $\forall x, y \in X, \exists g \in G \mid \psi(g)(x) = y$.

Osservazione: $T(V)$ agisce su V in modo transitivo, infatti $\forall v, w \in V \exists \tau \in T(V) \mid \tau(v) = w$ (in particolare $\tau = \tau_{w-v}$ e $\psi(w-v) = \tau_{w-v}$).

DEFINIZIONE 5.1.4: Sia (V, ϕ) uno spazio euclideo e sia d la distanza indotta da ϕ su V . Allora definiamo $Isom(V, d) = \{f: V \rightarrow V \mid d(P, Q) = d(f(P), f(Q)) \forall P, Q \in V\}$ come l'insieme delle **isometrie affini** di V (spesso in seguito le chiameremo semplicemente **isometrie**).

Osservazione: Sicuramente $O(V, \phi) \subseteq Isom(V, d)$ e $T(V) \subseteq Isom(V, d)$, quindi $\forall v \in V, \forall f \in O(V, \phi), \tau_v \circ f \in Isom(V, d)$.

LEMMA 5.1.2: $\forall v \in V, \forall f \in O(V, \phi)$ (in realtà basterebbe $f \in GL(V)$),
 $f \circ \tau_v = \tau_{f(v)} \circ f$ (dunque in generale non commutano).

Dimostrazione:

$$(f \circ \tau_v)(x) = f(x + v) = f(x) + f(v) = (\tau_{f(v)} \circ f)(x).$$

PROPOSIZIONE 5.1.3: $\{\tau_v \circ f | v \in V, f \in O(V, \phi)\}$ è un gruppo rispetto al prodotto di composizioni.

Dimostrazione:

Mostriamo che è chiuso rispetto al prodotto di composizioni:

$$(\tau_v \circ f) \circ (\tau_w \circ g) = \tau_v \circ (f \circ \tau_w) \circ g = \underbrace{\tau_v \circ \tau_{f(w)}}_{\in T(V)} \circ \underbrace{f \circ g}_{\in O(V, \phi)}.$$

Inoltre $id = \tau_0 \circ id$, dunque rimane da far vedere che $(\tau_v \circ f)^{-1} \in \{\tau_v \circ f | v \in V, f \in O(V, \phi)\}$.

$$\tau_v \circ f \circ \tau_w \circ g = \tau_v \circ \tau_{f(w)} \circ f \circ g = id \Leftrightarrow \begin{cases} f \circ g = id \\ f(w) = -v \end{cases} \begin{cases} g = f^{-1} \\ w = -f^{-1}(v) \end{cases}$$

Quindi $(\tau_v \circ f)^{-1} = \tau_{-f^{-1}(v)} \circ f^{-1}$, da cui la tesi.

Osservazione: Con la stessa dimostrazione si prova che $\{\tau_v \circ f | v \in V, f \in GL(V)\}$ è un gruppo di trasformazioni di V .

TEOREMA 5.1.4: $\{\tau_v \circ f | v \in V, f \in O(V, \phi)\} = Isom(V, d)$.

Dimostrazione:

\subseteq) Già vista.

\supseteq) Sia $f \in Isom(V, d)$.

Abbiamo già provato che, se $f(0) = 0$, allora $f \in O(V, \phi)$.

Se invece $f(0) = v \neq 0$, allora $\tau_{-v} \circ f \in Isom(V, d)$ e $(\tau_{-v} \circ f)(0) = 0$, dunque

$\tau_{-v} \circ f = g \in O(V, \phi)$, da cui $f = \tau_v \circ g$ con $g \in O(V, \phi)$.

Osservazione: Ora andremo a studiare le isometrie di $(\mathbb{R}^n, \langle, \rangle)$.

Notiamo che $Isom(\mathbb{R}^n) = \{X \rightarrow AX + B | A \in O(n), B \in \mathbb{R}^n\}$.

Inoltre se $f \in Isom(\mathbb{R}^n)$, allora $Fix(f) = \{X \in \mathbb{R}^n | AX + B = X\} = \{X \in \mathbb{R}^n | (A - I)X = -B\}$,

quindi $Fix(f)$ o è vuoto, oppure è un sottospazio affine di \mathbb{R}^n con giacitura

$$\{X \in \mathbb{R}^n | (A - I)X = 0\} = V_1(A) = Fix(A).$$

DEFINIZIONE 5.1.5: $f \in Isom(\mathbb{R}^n)$ è detta **simmetria** se $f^2 = id$.

PROPOSIZIONE 5.1.5: Sia f una simmetria di \mathbb{R}^n , $f(X) = AX + B$, con $A \in O(n)$. Allora:

- 1) $A^2 = I$ e $f(B) = AB + B = 0$;
- 2) $\frac{B}{2} \in Fix(f)$ (quindi $Fix(f) = \frac{B}{2} + Fix(A)$);
- 3) B è ortogonale a $Fix(A)$.

Dimostrazione:

$$1) \forall X \in \mathbb{R}^n, f^2(X) = X \Rightarrow A(AX + B) + B = A^2X + AB + B = X \Rightarrow A^2 = I \text{ e } AB + B = 0.$$

$$2) f\left(\frac{B}{2}\right) = A \cdot \frac{B}{2} + B \stackrel{\square}{=} -\frac{B}{2} + B = \frac{B}{2} \text{ (il passaggio } \square \text{ segue da 1))}.$$

3) $Fix(A) = V_1(A)$ e $Fix(A)^\perp = V_{-1}(A)$, in quanto $A^2 = I$ e dunque ha come autovalori solo 1 e -1 ; inoltre se $x \in V_{-1}(A)$, $y \in V_1(A)$, si ha $\langle x, y \rangle = {}^t x y = {}^t (-Ax) A y = -{}^t x {}^t A A y = -{}^t x y$,

cioè $\langle x, y \rangle = 0$. Poiché per 1) $AB = -B \Rightarrow B \in V_{-1}(A)$, ho la tesi.

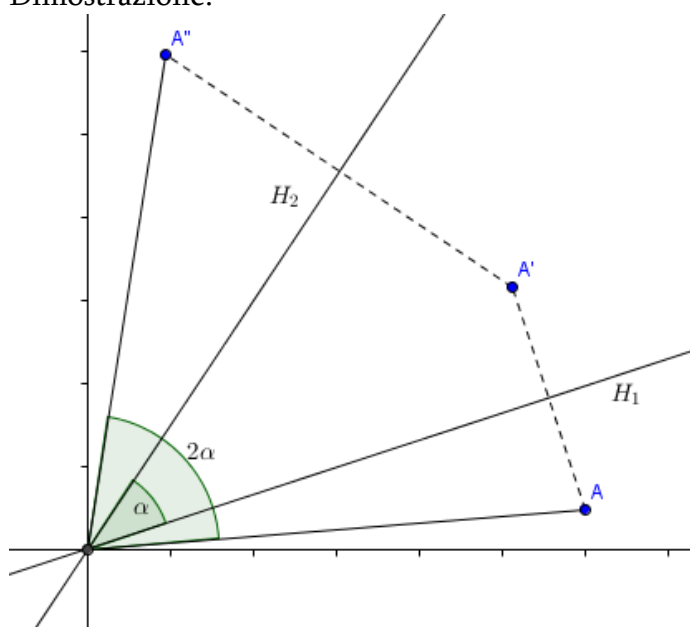
DEFINIZIONE 5.1.6: $f \in Isom(\mathbb{R}^n)$ è detta **riflessione** se $f^2 = id$ e $Fix(f)$ è un iperpiano affine (cioè ha come giacitura un sottospazio di dimensione $n - 1$).

Osservazione: Se $f(X) = AX + B$ è una riflessione, allora $\dim(Fix(A)) = n - 1$, per cui $A \in O(n)$ è una riflessione lineare rispetto alla giacitura di $Fix(f)$.
In particolare $\det(A) = -1$.

DEFINIZIONE 5.1.7: $f \in Isom(\mathbb{R}^n)$ è detta **rotazione** se $Fix(f)$ ha come giacitura un sottospazio di dimensione $n - 2$.

LEMMA 5.1.7: La composizione di due riflessioni ρ_1, ρ_2 di \mathbb{R}^n rispetto a iperpiani incidenti H_1 e H_2 (non coincidenti) è una rotazione r tale che $Fix(r) = H_1 \cap H_2$.

Dimostrazione:



Dobbiamo mostrare che $\dim(Fix(r)) = n - 2$.

Notiamo che, se W_i è la giacitura di H_i per $i = 1, 2$, allora $\dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 + W_2) = n - 1 + n - 1 - n = n - 2$ e $W_1 \cap W_2$ è la giacitura di $H_1 \cap H_2$.

Sicuramente $H_1 \cap H_2 \subseteq Fix(r)$, poiché se $\rho_i(X) = A_i X + B_i$ per $i = 1, 2$, allora:

$$X \in H_1 \cap H_2 \Rightarrow A_i X + B_i = X \text{ per } i = 1, 2 \Rightarrow r(X) = (\rho_1 \circ \rho_2)(X) = A_1(A_2 X + B_2) + B_1 = A_1 X + B_1 = X.$$

Dunque la dimensione della giacitura di r può essere $n - 2$ o $n - 1$ (se fosse $n \Rightarrow r = id$, cioè i due piani sarebbero coincidenti).

Se la dimensione è $n - 1$, allora r è una riflessione, assurdo, perché

$$\det(A_1 A_2) = \det(A_1) \cdot \det(A_2) = 1.$$

Dunque $Fix(r) = H_1 \cap H_2$.

Osservazione: Notiamo anche che la rotazione r è di un angolo doppio rispetto a quello formato dai due iperpiani H_1 e H_2 ; questo è semplicemente dimostrabile in \mathbb{R}^2 e \mathbb{R}^3 , ma non lo dimostriamo in dimensione maggiore per evitare di definire un angolo in \mathbb{R}^n .

PROPOSIZIONE 5.1.8: La composizione di due riflessioni di \mathbb{R}^n può essere una traslazione o una rotazione.

Dimostrazione:

Siano $\rho_1(X) = A_1X + B_1$ e $\rho_2(X) = A_2X + B_2$ le due riflessioni.

Siano $H_1 = \text{Fix}(\rho_1)$ e $H_2 = \text{Fix}(\rho_2)$ i due iperpiani di punti fissi.

Caso 1): H_1 e H_2 sono paralleli (e quindi hanno la stessa giacitura).

Allora $\text{Fix}(A_1) = \text{Fix}(A_2)$, cioè A_1 e A_2 sono riflessioni lineari rispetto allo stesso iperpiano lineare $\Rightarrow A_1 = A_2 = A$. Allora:

$$(\rho_2 \circ \rho_1)(X) = A(AX + B_1) + B_2 = A^2X + AB_1 + B_2 = X + (AB_1 + B_2) = X + (B_2 - B_1) \Rightarrow \rho_2 \circ \rho_1 \text{ è la traslazione } \tau_{\overline{B_1B_2}}.$$

(Osservazione: $B_2 - B_1$ è ortogonale ai due iperpiani e $\|B_2 - B_1\| = 2d(H_1, H_2)$. Pertanto qualsiasi coppia di iperpiani paralleli a H_1, H_2 e aventi la stessa distanza produce $\rho_2 \circ \rho_1$.)

Caso 2): H_1 e H_2 non sono paralleli.

Allora H_1 e H_2 sono incidenti non coincidenti (altrimenti la composizione sarebbe l'identità), dunque la tesi segue per il lemma.

FORMA CANONICA PER UNA SIMMETRIA: Sia $f \in \text{Isom}(\mathbb{R}^n)$ una simmetria, $k = \dim(\text{Fix}(f))$. Allora esiste una base ortonormale $\{v_1, \dots, v_n\}$ di \mathbb{R}^n rispetto alla quale f si scrive:

$$f(X) = \begin{pmatrix} \boxed{I_k} & 0 \\ 0 & \boxed{-I_{n-k}} \end{pmatrix} X + \alpha v_n.$$

TEOREMA 5.1.9: Ogni $f \in \text{Isom}(\mathbb{R}^n)$ è composizione di al più $n + 1$ riflessioni.

Dimostrazione:

Se $f(0) = 0 \Rightarrow f \in O(\mathbb{R}^n) \Rightarrow f$ è composizione di al più n riflessioni lineari.

Se $f(0) = B \neq 0$, consideriamo $H = \{X \in \mathbb{R}^n \mid d(X, 0) = d(X, B)\}$.

H è un iperpiano, poiché $X \in H \Leftrightarrow \|X\|^2 = \|X - B\|^2 \Leftrightarrow \langle X, X \rangle = \langle X - B, X - B \rangle \Leftrightarrow \langle 2\langle B, X \rangle = \langle B, B \rangle$, che è un'unica equazione. Inoltre $\frac{B}{2} \in H$.

In particolare la giacitura di H è $\text{Span}(B)^\perp$.

Sia ρ_H la riflessione rispetto a H . Allora:

$\rho_H \circ f \in \text{Isom}(\mathbb{R}^n)$ e $(\rho_H \circ f)(0) = \rho_H(B) = 0$, quindi $\rho_H \circ f \in O(\mathbb{R}^n)$ è composizione di al più n riflessioni $\Rightarrow f$ è composizione di al più $n + 1$ riflessioni.

DEFINIZIONE 5.1.8: Un'isometria si dice **diretta** se è composizione di un numero pari di riflessioni, **inversa** altrimenti.

Osservazione: $f(X) = AX + B$ è diretta (inversa) $\Leftrightarrow \det(A) = 1$ ($\det(A) = -1$).

DEFINIZIONE 5.1.9: Sia $f \in \text{Isom}(\mathbb{R}^n)$. Diciamo che f è:

- una **glissoriflessione** (o **riflessione traslata**, o **glide**) se è composizione di una riflessione ρ e di una traslazione parallela a $\text{Fix}(\rho)$;

- una **riflessione rotatoria** se è composizione di una riflessione ρ e di una rotazione attorno a una iperretta (sottospazio di dimensione $n - 2$) che contiene una retta ortogonale a $Fix(\rho)$;
- un **avvitamento** (o **twist**) se è composizione di una rotazione r e di una traslazione (non banale) parallela a $Fix(r)$.

Osservazione: Da quello che abbiamo visto, segue che:

- se $\tau \in T(V)$, allora $\det(\tau) = 1$ e $Fix(\tau) = \emptyset$;
- se ρ è una riflessione, $\det(\rho) = -1$ e $\dim(Fix(\rho)) = n - 1$;
- se r è una rotazione, $\det(r) = 1$ e $\dim(Fix(r)) = n - 2$;
- se g è una glissoriflessione, $\det(g) = -1$ e $Fix(g) = \emptyset$;
- se R è una riflessione rotatoria, $\det(R) = -1$ e $\dim(Fix(R)) = n - 3$ (poiché se $x \in Fix(R)$, è chiaro che $x \in Fix(\rho) \cap Fix(r)$, dove $R = \rho \circ r$, e $\dim(Fix(\rho) \cap Fix(r)) = \dim(Fix(\rho)) + \dim(Fix(r)) - \dim(Fix(\rho) + Fix(r)) = n - 1 + n - 2 - n = n - 3$);
- se t è un avvitamento, allora $\det(t) = 1$ e $Fix(t) = \emptyset$.

PROPOSIZIONE 5.1.10: Sia $f \in Isom(\mathbb{R}^n)$. Se f ha almeno un punto fisso, allora f è composizione di al più n riflessioni.

Dimostrazione:

Sia $f(0) = B$. Se $B = 0$, segue subito la tesi.

Sia allora $B \neq 0$.

Sia $Q|f(Q) = Q$; $H = \{X \in \mathbb{R}^n | d(X, 0) = d(X, B)\}$.

Come provato prima, $\rho_H \circ f$ è lineare. Inoltre $Q \in H$, infatti:

$$d(Q, 0) = d(f(Q), f(0)) = d(Q, B).$$

Allora $\rho_H(Q) = Q$ e quindi $(\rho_H \circ f)(Q) = Q \Rightarrow \rho_H \circ f$ è lineare e $\dim Fix(\rho_H \circ f) \geq 1$, quindi $\rho_H \circ f$ è composizione di al più $n - 1$ riflessioni, quindi f è composizione di al più n riflessioni, da cui la tesi.

Osservazione: Se $Fix(f) = \emptyset$, possono effettivamente essere necessarie $n + 1$ riflessioni:

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2 | f((x, y)) = (x + 2, -y); \text{ allora } f: \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

$Fix(f) = \emptyset$; non bastano 2 riflessioni perché f non è né una traslazione, né una riflessione, né una rotazione.

D'altra parte, se $f = \tau_v$, $Fix(f) = \emptyset$ ma bastano 2 riflessioni.

TEOREMA DI CLASSIFICAZIONE DELLE ISOMETRIE PIANE: Ogni $f \in Isom(\mathbb{R}^2)$ è:

- 1) una traslazione;
- 2) una rotazione;
- 3) una riflessione;
- 4) una glissoriflessione.

Dimostrazione:

f è composizione di al più 3 riflessioni.

Se f è composizione di 0 riflessioni, $f = id$.

Se f è composizione di 1 riflessione, allora f è una riflessione.

Se f è composizione di 2 riflessioni, abbiamo visto che f è una traslazione o una rotazione.

Se f è composizione di 3 riflessioni, diciamo $f = \rho_1 \circ \rho_2 \circ \rho_3$, allora:

Caso 1): $\rho_1 \circ \rho_2 = \tau_v$ traslazione.

Scrivo $v = v_1 + v_2$, con $v_1 // v_3$ e $v_2 \perp r_3$. Allora:

$$f = \tau_v \circ \rho_3 = \tau_{v_1} \circ \tau_{v_2} \circ \rho_3.$$

Ma $\tau_{v_2} \circ \rho_3$ è una riflessione con asse $r'_3 // r_3$, dunque f è una glissoriflessione.

(Osservazione: Se $v_1 = 0$ (cioè $v \perp r_3$), f è una riflessione. Quindi:

$f = \tau_v \circ \rho_3$ è una riflessione \Leftrightarrow gli assi di riflessione sono 3 rette parallele.)

Caso 2): $\rho_1 \circ \rho_2$ è una rotazione attorno ad un punto P con angolo α .

a) $P \notin r_3$.

Allora scrivo la rotazione $R = \rho_1 \circ \rho_2$ come composizione di due riflessioni ρ'_1 e ρ'_2 rispetto a rette incidenti di angolo $\frac{\alpha}{2}$ e $r'_2 // r_3$.

$$f = R \circ \rho_3 = \rho'_1 \circ \underbrace{\rho'_2 \circ \rho_3}_{\text{traslazione}} \Rightarrow f \text{ è glissoriflessione.}$$

b) $P \in r_3$ (gli assi di ρ_1, ρ_2, ρ_3 si dicono **concorrenti** in P).

Scrivo $R = \rho_1 \circ \rho_2$ come $\rho'_1 \circ \rho_3$, dove ρ'_1 è una riflessione rispetto ad una opportuna retta passante per $P \Rightarrow f = \rho'_1 \circ \rho_3 \circ \rho_3 = \rho'_1 \Rightarrow f$ è riflessione.

(Osservazione: $\rho_1 \circ \rho_2 \circ \rho_3$ è riflessione \Leftrightarrow i 3 assi sono concorrenti o paralleli).

Osservazione: La composizione di 3 riflessioni di \mathbb{R}^3 rispetto a 3 piani incidenti in un punto P e a due a due ortogonali è una simmetria con $Fix(f) = \{P\}$. Quest'ultima è detta **simmetria centrale** di centro P .

In generale la composizione fra due isometrie "elementari" (riflessioni, traslazioni e rotazioni) non è commutativa, tranne che nei casi speciali della glissoriflessione, della riflessione rotatoria e dell'avvitamento. Ne tralasciamo la dimostrazione.

TEOREMA DI CLASSIFICAZIONE DELLE ISOMETRIE DI \mathbb{R}^3 : Ogni $f \in Isom(\mathbb{R}^3)$ è una:

- traslazione;
- riflessione;
- rotazione;
- glissoriflessione;
- avvitamento;
- riflessione rotatoria.

Dimostrazione:

$f(X) = AX + B$, con $A \in O(n)$, per cui $\exists B$ base ortonormale di \mathbb{R}^3 , $B = \{v_1, v_2, v_3\} | \mathfrak{M}_B(A) =$

a) I ;

b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$;

c) $\left(\begin{array}{cc|c} rot(\theta) & & 0 \\ & & 1 \end{array} \right), \theta \in (0, 2\pi)$;

d) $\left(\begin{array}{cc|c} rot(\theta) & & 0 \\ & & -1 \end{array} \right), \theta \in (0, 2\pi)$.

Studiamo i vari casi:

a) f è una traslazione, $f = \tau_0$.

b) $f(X) = AX$ è una riflessione (poiché $f|_{\text{span}(v_1, v_2)} = id$), quindi il tipo di $f(X) = AX + B$ dipende dalla posizione della traslazione B .

Sia $H = \text{Span}(v_1, v_2)$ e $\forall v \in \mathbb{R}^3$, v_H, v_H^\perp le proiezioni di v su H e $H^\perp = \text{Span}(v_3)$ rispetto a $\mathbb{R}^3 = H \oplus H^\perp$.

$$f(X) = (AX + B_H^\perp) + B_H.$$

Studiamo $\text{Fix}(X \rightarrow AX + B_H^\perp)$:

$$AX + B_H^\perp = X \Leftrightarrow (A - I)X = -B_H^\perp.$$

Ora $(A - I)|_H \equiv 0$ e $(A - I)|_{H^\perp} = -2I$, dunque:

$$\begin{aligned} (A - I)X &= (A - I)(X_H + X_H^\perp) = -2X_H^\perp \Rightarrow AX + B_H^\perp = X \Leftrightarrow -2X_H^\perp = -B_H^\perp \Leftrightarrow X_H^\perp = \frac{B_H^\perp}{2} \\ &\Leftrightarrow X = \frac{B_H^\perp}{2} + h, h \in H \Leftrightarrow X \in \frac{B_H^\perp}{2} + H \end{aligned}$$

Perciò $\text{Fix}(X \rightarrow AX + B_H^\perp)$ è un piano affine e $X \rightarrow AX + B_H^\perp$ è una riflessione ρ .

$f = \tau_{B_H} \circ \rho$, $B_H \in H$ che è la giacitura di $\text{Fix}(S)$.

Dunque f è una riflessione se $B_H = 0$, una glissoriflessione se $B_H \neq 0$.

c) Usando la stessa notazione del punto precedente, vogliamo mostrare che $X \rightarrow AX + B_H$ è una rotazione.

$$AX + B_H = X \Leftrightarrow (A - I)X = -B_H.$$

$(A - I)|_H: H \rightarrow H$ è invertibile, $(A - I)|_{H^\perp} \equiv 0$; per cui:

$\exists! X_0 \in H | (A - I)X_0 = -B_H$; inoltre $\text{Ker}(A - I) = H^\perp$, quindi

X risolve $(A - I)X = -B_H \Leftrightarrow X \in X_0 + H^\perp$.

Quindi $\text{Fix}(X \rightarrow AX + B_H)$ è una retta affine, per cui $X \rightarrow AX + B_H$ è una rotazione.

Dunque se $B_H^\perp = 0$, f è una rotazione, altrimenti f è un avvitamento.

$$d) \underbrace{\begin{pmatrix} \text{rot}(\theta) & | & 0 \\ 0 & | & -1 \end{pmatrix}}_{=A} = \underbrace{\begin{pmatrix} \text{rot}(\theta) & | & 0 \\ 0 & | & 1 \end{pmatrix}}_{=A_H} + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{=A_{H^\perp}} - I$$

Dunque:

$$\begin{aligned} f(X) = AX + B &= (A_H + A_{H^\perp} - I)X + B_H + B_{H^\perp} = \\ &= (A_H X + B_H - X_H) + (A_{H^\perp} X + B_{H^\perp} - X_{H^\perp}) \end{aligned}$$

Sappiamo già per il punto c) che $(A_H X + B_H - X_H): H \rightarrow H$ è una rotazione e per il punto b) che $(A_{H^\perp} X + B_{H^\perp} - X_{H^\perp}): H^\perp \rightarrow H^\perp$ è una riflessione (con piano di riflessione \perp all'asse di rotazione).

Dunque $\forall B$, f è una riflessione rotatoria.

5.2 SPAZI E SOTTOSPAZI AFFINI

DEFINIZIONE 5.2.1: Definiamo $A(V) = \{\tau_v \circ f \mid v \in V, f \in GL(V)\}$.

Osservazione: Abbiamo visto che $A(V)$ è un gruppo di trasformazioni di V (che in particolare coincide con il sottogruppo di $S(V)$ generato da $T(V)$ e $GL(V)$)

DEFINIZIONE 5.2.2: G gruppo di trasformazioni di un insieme X . $\forall x \in X$, definiamo **stabilizzatore** di x l'insieme $St_x(G) = \{g \in G \mid g(x) = x\}$.

PROPOSIZIONE 5.2.1: 1) $St_v(GL(V)) = GL(V) \Leftrightarrow v = 0$
 2) $St_v(A(V)) \cong GL(V) \quad \forall v \in V$.

Dimostrazione:

1) Ovviamente $St_0(GL(V)) = GL(V)$.

D'altra parte se $v \neq 0$, $\exists f \in GL(V) \mid f(v) \neq v$.

2) Vediamo intanto che $\forall v, w \in V$, $St_v(A(V))$ e $St_w(A(V))$ sono isomorfi.

Infatti se $u = v - w$ (così che $\tau_u(w) = v$):

$$\begin{aligned} St_v(A(V)) &\rightarrow St_w(A(V)) \\ f &\rightarrow \tau_{-u} \circ f \circ \tau_u \end{aligned}$$

è un isomorfismo.

Inoltre $\tau_v \circ f \in St_0(A(V)) \Leftrightarrow (\tau_v \circ f)(0) = 0 \Leftrightarrow v = 0$. Cioè $St_0(A(V)) = GL(V)$, tesi.

Osservazione: Questo fa capire che in V c'è un punto privilegiato, cioè l'origine O , mentre rispetto ad $A(V)$ tutti i punti di V sono equivalenti.

Dunque nella seguente trattazione distingueremo gli elementi di V pensati come punti e i vettori di V pensati come traslazioni (poiché sappiamo che $(T(V), \circ) \cong (V, +)$).

DEFINIZIONE 5.2.3: V spazio vettoriale. Un insieme A (anche \emptyset) si dice **spazio affine** su V se $\exists F: A \times A \rightarrow V$ che associa ad ogni coppia di punti $P, Q \in A$ un vettore di V , denotato \overrightarrow{PQ} , in modo che:

1) $\forall P \in A, \forall v \in V, \exists! Q \in A \mid \overrightarrow{PQ} = v$;

2) $\forall P, Q, R \in A, \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ (relazione di Chasles).

DEFINIZIONE 5.2.4: Definiamo **dimensione** di uno spazio affine A , $\dim(A) = \dim(V)$ se $A \neq \emptyset$, altrimenti $\dim(\emptyset) = -1$.

Osservazione: Dalla relazione di Chasles discende:

- $\forall P \in A, \overrightarrow{PP} = 0$ (basta prendere $P = Q = R$);
- $\forall P, Q \in A, \overrightarrow{QP} = -\overrightarrow{PQ}$ (basta prendere $P = R$).

Esempi: Alcuni esempi di spazi affini possono essere:

1) $A = V, F: V \times V \rightarrow V \mid (P, Q) \rightarrow \overrightarrow{PQ} \stackrel{\text{def}}{=} Q - P$ (detto **spazio affine standard** su V)

2) $f: U \rightarrow W$ lineare, $b \in W$.

Sia $A = f^{-1}(b)$, $V = f^{-1}(0) = \text{Ker}(f)$.

Allora A è spazio affine su V tramite $F: A \times A \rightarrow V \mid (a_1, a_2) \rightarrow \overrightarrow{a_1 a_2} \stackrel{\text{def}}{=} a_2 - a_1$.

DEFINIZIONE 5.2.5: Definiamo **traslazione** del vettore v , $\tau_v: A \rightarrow A \mid P \rightarrow Q$, con $\overrightarrow{PQ} = v$.

Notazione: Se $P \in A$ e $v \in V$, denoteremo con $P + v$ l'unico punto Q di A tale che $\overrightarrow{PQ} = v$.

Osservazione: Con questa notazione si ha:

- $\overrightarrow{P(P+v)} = v$;
- $P + \overrightarrow{PQ} = Q$;
- $\tau_v(P) = P + v$.

Si ha così l'applicazione $A \times V \rightarrow A \mid (P, v) \rightarrow P + v$.

LEMMA 5.2.2: $\forall P \in A, \forall v_1, v_2 \in V, (P + v_1) + v_2 = P + (v_1 + v_2)$.

(Osservazione: Questa relazione non è affatto ovvia, in quanto la scrittura $P + v$ è solamente una notazione e non ha niente a che fare con la somma tradizionale.)

Dimostrazione:

Sia $P_1 = P + v_1$, ossia $\overrightarrow{PP_1} = v_1$.

Sia $P_2 = P_1 + v_2$, ossia $\overrightarrow{P_1 P_2} = v_2$.

$P + (v_1 + v_2) = P + (\overrightarrow{PP_1} + \overrightarrow{P_1 P_2}) = P + \overrightarrow{PP_2} = P_2$, da cui la tesi.

DEFINIZIONE 5.2.6: Fissato $P \in A$, definiamo l'applicazione $F_P: A \rightarrow V \mid Q \rightarrow \overrightarrow{PQ}$ (a volte questa applicazione prende il nome di **sollevamento** di A su V).

Osservazione: Dagli assiomi segue che F_P è bigettiva e $F_P(P) = \overrightarrow{PP} = 0$, dunque F_P trasforma P nell'origine di V .

In altre parole, con F_P "identifico" A con V , ossia "sollevo" su A una struttura di spazio vettoriale. Questo si può fare $\forall P \in A$.

Osservazione: $\forall v \in V, F_P^{-1}(v) = P + v$.

Osservazione: Cerchiamo di sollevare la nozione di combinazione lineare di vettori nello spazio affine: siano $P_1, \dots, P_k \in A$. Fissiamo $P \in A$.

F_P trasforma P_i in $\overrightarrow{PP_i} \forall i$.

Inoltre $\forall t_1, \dots, t_k \in \mathbb{K}, \exists t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k} \in V$ e

$F_P^{-1}(t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k}) = P + t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k} \in A$.

Affinché sia una buona definizione, il risultato non deve dipendere da P .

Cioè se $A \ni Q \neq P$, allora deve valere $P + \sum_{i=1}^k t_i \overrightarrow{PP_i} = Q + \sum_{i=1}^k t_i \overrightarrow{QP_i}$.

$$P + \sum_{i=1}^k t_i \overrightarrow{PP_i} = P + \sum_{i=1}^k t_i (\overrightarrow{PQ} + \overrightarrow{QP_i}) = P + \left(\sum_{i=1}^k t_i \right) \overrightarrow{PQ} + \sum_{i=1}^k t_i \overrightarrow{QP_i},$$

che coincide con $Q + \sum_{i=1}^k t_i \overrightarrow{QP_i} \Leftrightarrow P + (\sum_{i=1}^k t_i) \overrightarrow{PQ} = Q \Leftrightarrow \sum_{i=1}^k t_i = 1$.

DEFINIZIONE 5.2.7: Dati $P_1, \dots, P_k \in A$ e $t_1, \dots, t_k \in \mathbb{K} \mid \sum_{i=1}^k t_i = 1$, definiamo combinazione affine dei punti P_i con coefficienti t_1, \dots, t_k il punto $P + t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k}$, dove P è un qualunque punto di A .

Osservazione: Notiamo che $\forall P, P_1, \dots, P_k \in A$ e $\forall t_1, \dots, t_k \in \mathbb{K} \mid \sum_{i=1}^k t_i = 1$, vale:

$$P + t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k} = t_1 P_1 + \dots + t_k P_k.$$

Infatti $F_P(P + t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k}) = t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k}$ e

$$F_P(t_1 P_1 + \dots + t_k P_k) = \overrightarrow{P(t_1 P_1 + \dots + t_k P_k)} = t_1 \overrightarrow{PP_1} + \dots + t_k \overrightarrow{PP_k},$$

e poiché F_P è bigettiva, ho la tesi.

Esempi: 1) Se $P, Q \in A$, $\frac{1}{2}P + \frac{1}{2}Q$ è detto punto medio fra P e Q ($\text{char}(\mathbb{K}) \neq 2$)

2) Se $P_1, \dots, P_k \in A$, $\frac{1}{k}P_1 + \dots + \frac{1}{k}P_k$ è detto baricentro dei P_i ($\text{char}(\mathbb{K}) \nmid k$).

DEFINIZIONE 5.2.8: Definiamo **insieme delle combinazioni affini** dei punti $P_1, \dots, P_k \in A$ $\text{Comb}_a(P_1, \dots, P_k) = \{\text{combinazioni affini di } P_1, \dots, P_k\}$.

Esempio: $A = \mathbb{K}^n$, $P_1, P_2 \in A$ punti distinti.

$$\begin{aligned} \text{Comb}_a(P_1, P_2) &= \{t_1 P_1 + t_2 P_2 \mid t_1 + t_2 = 1\} = \{P_1 + (1-t)\overrightarrow{P_1 P_1} + t\overrightarrow{P_1, P_2} \mid t \in \mathbb{K}\} = \\ &= \{P_1 + t(P_2 - P_1) \mid t \in \mathbb{K}\}, \end{aligned}$$

cioè è la retta passante per P_1 e P_2 .

Analogamente se vede che $\text{Comb}_a(P_1, P_2, P_3)$ è il piano contenente P_1, P_2, P_3 (se i tre punti non sono allineati).

Similmente al caso lineare, definisco:

DEFINIZIONE 5.2.9: $H \subseteq A$ (anche $H = \emptyset$) è detto sottospazio affine di A se è chiuso per combinazioni affini.

Osservazione: Come nel caso lineare, l'intersezione di una famiglia arbitraria di sottospazi affini è un sottospazio affine.

Notazione: Se $P_0 \in A$ e W è sottospazio di V , indicheremo con $P_0 + W = \{P_0 + w \mid w \in W\}$.

Osservazione: $P_0 + W = \{\tau_w(P_0) \mid w \in W\}$ e $F_P^{-1}(W) = P + W$.

PROPOSIZIONE 5.2.3: $\forall P_0 \in A, \forall W$ sottospazio di V , $P_0 + W$ è sottospazio affine di A .

Dimostrazione:

Verifico che $P_0 + W$ è chiuso per combinazioni affini.

Siano $P_0 + w_1, \dots, P_0 + w_k \in P_0 + W$; $t_1 + \dots + t_k = 1$.

$$\begin{aligned} t_1(P_0 + w_1) + \dots + t_k(P_0 + w_k) &= P_0 + t_1 \overrightarrow{P_0(P_0 + w_1)} + \dots + t_k \overrightarrow{P_0(P_0 + w_k)} = \\ &= P_0 + t_1 w_1 + \dots + t_k w_k \in P_0 + W, \text{ da cui la tesi.} \end{aligned}$$

PROPOSIZIONE 5.2.4: Sia $H \neq \emptyset$ un sottospazio affine di A . Allora esiste un unico sottospazio W_H di V (detto la **giacitura** di H), tale che $\forall P \in H, H = P + W_H$.

Dimostrazione:

Sia $P_0 \in H$. Cerco W_0 sottospazio di $V \mid H = P_0 + W_0$.

Poiché $P_0 + W_0 = F_{P_0}^{-1}(W_0)$, deve essere $W_0 = F_{P_0}(H) = \{w \in V \mid P_0 + w \in H\}$.

Verifico che l'insieme W_0 è sottospazio di V .

Siano $w_1, w_2 \in W_0, \alpha_1, \alpha_2 \in \mathbb{K}$.

Allora $P_0 + w_1 \in H, P_0 + w_2 \in H$.

Devo mostrare che $\alpha_1 w_1 + \alpha_2 w_2 \in W_0$, cioè che $P_0 + \alpha_1 w_1 + \alpha_2 w_2 \in H$.

$$\begin{aligned} P_0 + \alpha_1 w_1 + \alpha_2 w_2 &= P_0 + \alpha_1 \overrightarrow{P_0 P_0 + w_1} + \alpha_2 \overrightarrow{P_0 P_0 + w_2} = \\ &= P_0 + (1 - \alpha_1 - \alpha_2) \overrightarrow{P_0 P_0} + \alpha_1 \overrightarrow{P_0 P_0 + w_1} + \alpha_2 \overrightarrow{P_0 P_0 + w_2}, \end{aligned}$$

che appartiene a H perché combinazione affine dei punti $P_0, P_0 + w_1, P_0 + w_2$ che stanno in H .

Resta da controllare che W_0 non dipende dalla scelta di P_0 , ossia che $\forall P_1, P_2 \in H$,

$$F_{P_1}(H) = F_{P_2}(H).$$

$$F_{P_1}(H) = \{\overrightarrow{P_1 Q} \mid Q \in H\}.$$

Sia $\overrightarrow{P_1 Q} \in F_{P_1}(H)$.

$$\text{Allora } \overrightarrow{P_1 Q} = \overrightarrow{P_1 P_2} + \overrightarrow{P_2 Q} = \underbrace{-\overrightarrow{P_2 P_1}}_{\in F_{P_2}(H)} + \underbrace{\overrightarrow{P_2 Q}}_{\in F_{P_2}(H)} \in F_{P_2}(H),$$

cioè $F_{P_1}(H) \subseteq F_{P_2}(H)$.

Analogamente si prova l'inclusione opposta \Rightarrow tesi.

DEFINIZIONE 5.2.10: Se H è sottospazio affine di A , si pone $\dim(H) = \dim(W_H)$.

- H è detto **retta (affine)** se $\dim(H) = 1$;
- H è detto **piano (affine)** se $\dim(H) = 2$;
- H è detto **iperpiano (affine)** se $\dim(H) = \dim(A) - 1$.

Osservazione: H, L sottospazi affini di $A, H \cap L \neq \emptyset$.

Allora $\forall P \in H \cap L, H = P + W_H, L = P + W_L$.

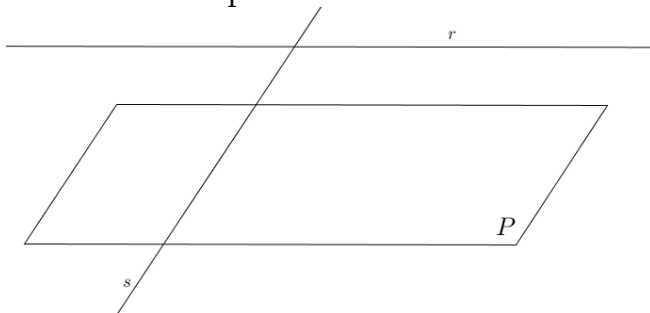
$$H \cap L = (P + W_H) \cap (P + W_L) = P + (W_H \cap W_L) \Rightarrow W_{H \cap L} = W_H \cap W_L.$$

Dunque $\dim(H \cap L) = \dim(W_H \cap W_L)$.

DEFINIZIONE 5.2.11: Due sottospazi affini H, L si dicono:

- incidenti se $H \cap L \neq \emptyset$;
- paralleli se $W_H \subseteq W_L \vee W_L \subseteq W_H$;
- sghembi se $H \cap L = \emptyset \wedge W_H \cap W_L = \{0\}$.

Osservazione: Il parallelismo non è una relazione di equivalenza:



infatti dalla figura si vede che $W_r \subseteq W_p$ e $W_s \subseteq W_p$, ma $W_s \not\subseteq W_r \wedge W_r \not\subseteq W_s$.

PROPOSIZIONE 5.2.5: $P_0, \dots, P_k \in A$. Allora:

$$Comb_a(P_0, \dots, P_k) = P_0 + Span(\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k}).$$

Quindi $Comb_a(P_0, \dots, P_k)$ è un sottospazio affine di A , detto il sottospazio affine generato da P_0, \dots, P_k . È il più piccolo sottospazio affine di A contenente P_0, \dots, P_k .

Dimostrazione:

Per definizione di combinazione affine vale \subseteq .

Mostriamo ora l'inclusione opposta:

$\forall t_1, \dots, t_k \in \mathbb{K}, P_0 + t_1\overrightarrow{P_0P_1} + \dots + t_k\overrightarrow{P_0P_k} = P_0 + (1 - \sum_{i=1}^k t_i)\overrightarrow{P_0P_0} + t_1\overrightarrow{P_0P_1} + \dots + t_k\overrightarrow{P_0P_k}$, che appartiene a $Comb_a(P_0, \dots, P_k) \Rightarrow$ tesi.

Osservazione: Dunque $\dim(Comb_a(P_0, \dots, P_k)) = \dim(Span(\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k}))$.

Notazione: In generale $\forall X \subseteq A, X \neq \emptyset$, denotiamo con $Comb_a(X)$ l'insieme delle combinazioni affini dei punti di X .

PROPOSIZIONE 5.2.6: $X \subseteq A. \forall R \in X, Comb_a(X) = R + Span(\{\overrightarrow{RQ} | Q \in X\})$.

Inoltre se $H \subseteq A$ è un sottospazio affine di A che contiene X , allora $Comb_a(X) \subseteq H$.

Dimostrazione:

È analoga a quella della proposizione precedente.

DEFINIZIONE 5.2.12: Se H, L sono sottospazi affini di A , poniamo $H + L = Comb_a(H \cup L)$.

Osservazione: $H \subseteq L \Rightarrow W_H \subseteq W_L$.

Dimostrazione:

$$\forall P \in H, H = P + W_H, L = P + W_L.$$

$$\forall w \in W_H, P + w \in H \subseteq L \Rightarrow w \in W_L.$$

PROPOSIZIONE 5.2.7: H, L sottospazi affini di A . Allora $\forall P \in H, \forall Q \in L$:

$$W_{H+L} = W_H + W_L + Span(\overrightarrow{PQ}).$$

Dimostrazione:

$$\supseteq) H \subseteq H + L, L \subseteq H + L \Rightarrow W_H + W_L \subseteq W_{H+L}.$$

$$\text{Poiché } Comb_a(P, Q) \subseteq H + L \Rightarrow Span(\overrightarrow{PQ}) \subseteq W_{H+L}.$$

$$\subseteq) \text{ Considero il sottospazio affine } S = P + (W_H + W_L + Span(\overrightarrow{PQ})).$$

Allora $S \supseteq P + W_H = H$; inoltre $S \supseteq L$, infatti:

$$Q = P + \overrightarrow{PQ} \in S \Rightarrow S = Q + (W_H + W_L + Span(\overrightarrow{PQ})) \Rightarrow S \supseteq Q + W_L = L.$$

$$\text{Per minimalità, } S \supseteq H + L \Rightarrow W_S = W_H + W_L + Span(\overrightarrow{PQ}) \supseteq W_{H+L}.$$

LEMMA 5.2.8: H, L sottospazi affini di A . Allora $H \cap L = \emptyset \Leftrightarrow \forall P \in H, \forall Q \in L, \overrightarrow{PQ} \notin W_H + W_L$.

Dimostrazione:

\Rightarrow) Per assurdo supponiamo che $\exists P \in H, Q \in L | \overrightarrow{PQ} = w_1 + w_2$, con $w_1 \in W_H, w_2 \in W_L$.

$$\underbrace{P + w_1}_{\in H} = P + (\overrightarrow{PQ} - w_2) = (P + \overrightarrow{PQ}) - w_2 = \underbrace{Q - w_2}_{\in L}, \text{ assurdo.}$$

⇐) Per assurdo supponiamo $\exists R \in H \cap L$.

Siano $P \in H, Q \in L$, per cui $H = P + W_H, L = Q + W_L$.

Allora $\overrightarrow{PQ} = \underbrace{\overrightarrow{PR}}_{\in W_H} + \underbrace{\overrightarrow{RQ}}_{\in W_L} \in W_H + W_L$, assurdo.

FORMULA DI GRASSMANN AFFINE: Siano H, L sottospazi di A .

1) Se $H \cap L \neq \emptyset \Rightarrow \dim(H + L) = \dim(H) + \dim(L) - \dim(H \cap L)$.

2) Se $H \cap L = \emptyset \Rightarrow \dim(H + L) = \dim(H) + \dim(L) - \dim(W_H \cap W_L) + 1$.

Dimostrazione:

$\dim(H + L) = \dim(W_{H+L})$ e $W_{H+L} = W_H + W_L + \text{Span}(\overrightarrow{PQ})$, con $P \in H$ e $Q \in L$.

1) Per il lemma, se $H \cap L \neq \emptyset, \overrightarrow{PQ} \in W_H + W_L \Rightarrow W_{H+L} = W_H + W_L$.

La tesi segue dalla formula di Grassmann vettoriale:

$$\begin{aligned} \dim(W_{H+L}) &= \dim(W_H + W_L) = \dim(W_H) + \dim(W_L) - \dim(W_H \cap W_L) = \\ &= \dim(H) + \dim(L) - \dim(H \cap L). \end{aligned}$$

2) Se $H \cap L = \emptyset$, per il lemma $\overrightarrow{PQ} \notin W_H + W_L$.

Quindi $\dim(W_{H+L}) = \dim(W_H + W_L + \text{Span}(\overrightarrow{PQ})) = \dim(W_H + W_L) + 1$ e si conclude come prima.

Osservazione: Se H e L sono sghembi, allora $\dim(H + L) = \dim(H) + \dim(L) + 1$ (infatti tutte le combinazioni affini di due rette sghembe in \mathbb{R}^3 generano tutto lo spazio \mathbb{R}^3).

DEFINIZIONE 5.2.13: P_0, \dots, P_k si dicono **affinemente indipendenti** se

$\dim(\text{Comb}_a(P_0, \dots, P_k)) = k$ (ossia se $\dim(\text{Span}(\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k})) = k$, cioè se $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k}$ sono linearmente indipendenti).

DEFINIZIONE 5.2.14: Sia $\dim(A) = n$. Si chiama **riferimento affine** di A ogni insieme ordinato $R = \{P_0, \dots, P_n\}$ di $n + 1$ punti di A affinemente indipendenti.

Osservazione: Essendo R ordinato, scelgo P_0 come “punto base”.

Inoltre $\{\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}\}$ è una base di V .

Osservazioni: 1) Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di $V, \forall P_0 \in A, \{P_0, P_0 + v_1, \dots, P_0 + v_n\}$ è riferimento affine di A .

2) Se $A = V$ e $\mathcal{B} = \{v_1, \dots, v_n\}$ è base di V allora $\{0, v_1, \dots, v_n\}$ è riferimento affine.

Ad esempio in $\mathbb{K}^n, \{0, e_1, \dots, e_n\}$ è detto **riferimento affine standard**.

PROPOSIZIONE 5.2.9: Se $R = \{P_0, \dots, P_n\}$ è un riferimento affine di A , ogni punto $P \in A$ si scrive in modo unico come $P = a_0P_0 + \dots + a_nP_n$, con $a_0 + \dots + a_n = 1$.

((a_1, \dots, a_n) sono quindi dette le **coordinate affini** di P rispetto a R).

Dimostrazione:

Per ipotesi $A = \text{Comb}_a(P_0, \dots, P_n)$, dunque gli a_i esistono.

Mostriamo che sono unici:

se $P = a_0P_0 + \dots + a_nP_n = a'_0P_0 + \dots + a'_nP_n$, con $\sum_{i=0}^n a_i = \sum_{i=0}^n a'_i = 1$, allora:

$$a_0P_0 + \dots + a_nP_n = P_0 + a_1\overrightarrow{P_0P_1} + \dots + a_n\overrightarrow{P_0P_n};$$

$$a'_0P_0 + \dots + a'_nP_n = P_0 + a'_1\overrightarrow{P_0P_1} + \dots + a'_n\overrightarrow{P_0P_n},$$

dunque $a_1\overrightarrow{P_0P_1} + \dots + a_n\overrightarrow{P_0P_n} = a'_1\overrightarrow{P_0P_1} + \dots + a'_n\overrightarrow{P_0P_n}$, ma i $\overrightarrow{P_0P_i}$ sono linearmente indipendenti, perciò $a_i = a'_i \forall i \Rightarrow a_0 = a'_0$, da cui la tesi.

DEFINIZIONE 5.2.15: A spazio affine su V , B spazio affine su W , V, W \mathbb{K} -spazi vettoriali.
 $f: A \rightarrow B$ si dice **trasformazione affine** se conserva le combinazioni affini.

DEFINIZIONE 5.2.16: $f: A \rightarrow B$ trasformazione affine si dice **isomorfismo affine** se è biunivoca.

DEFINIZIONE 5.2.17: Un isomorfismo affine $f: A \rightarrow A$ si dice **affinità**.
 Poniamo $Aff(A) = \{\text{affinità } A \rightarrow A\}$.

Esempi: 1) Le traslazioni sono affinità.

2) Se R è riferimento affine e $\dim(A) = n$,

$$[\]_R: \begin{matrix} A \rightarrow \mathbb{K}^n \\ P \rightarrow [P]_R \end{matrix}, \text{ dove } [P]_R \text{ sono le coordinate affini di } P \text{ rispetto a } R,$$

è un isomorfismo affine (quindi $A \cong \mathbb{K}^n$).

Osservazione: $[\]_R$ trasforma i punti di R in $0, e_1, \dots, e_n$.

PROPOSIZIONE 5.2.10: $A = V$. Se $f: V \rightarrow V$ è un'affinità e $f(0) = 0$, allora f è lineare.

Dimostrazione:

$$\forall v_1, v_2 \in V, \forall t_1, t_2 \in \mathbb{K}:$$

$$\begin{aligned} f(t_1v_1 + t_2v_2) &= f((1 - t_1 - t_2) \cdot 0 + t_1v_1 + t_2v_2) = (1 - t_1 - t_2)f(0) + t_1f(v_1) + t_2f(v_2) = \\ &= t_1f(v_1) + t_2f(v_2), \end{aligned}$$

da cui la tesi.

PROPOSIZIONE 5.2.11: $f \in Aff(V)$. Allora esistono unici $v \in V, g \in GL(V) \mid f = \tau_v \circ g$.

Dimostrazione:

Sia $v = f(0)$. Allora $g = \tau_{-v} \circ f \in Aff(V)$ e $g(0) = 0 \Rightarrow g \in GL(V)$.

Poiché $f(0)$ è unico, allora v è unico e quindi anche g è univocamente determinato, tesi.

Osservazione: Dunque $Aff(V) = \{\tau_v \circ g \mid v \in V, g \in GL(V)\}$.

Di conseguenza $Aff(V)$ coincide con il gruppo di trasformazioni che avevamo precedentemente denotato con $A(V)$, generato dalle traslazioni $T(V)$ di V e da $GL(V)$.

In particolare:

$$Aff(\mathbb{K}^n) = \{X \rightarrow AX + B \mid A \in GL(n, \mathbb{K}), B \in \mathbb{K}^n\}.$$

Osservazione: Prendiamo $f: A \rightarrow B$ biunivoca, con A, B spazi affini su V, W . Allora:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ F_{P_0} \downarrow & & \downarrow F_{f(P_0)} \\ V & & W \end{array}$$

Sia $\varphi_{P_0}: V \rightarrow W$ l'unica applicazione che rende commutativo il diagramma:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ F_{P_0} \downarrow & & \downarrow F_{f(P_0)} \\ V & \xrightarrow{\varphi_{P_0}} & W \end{array}$$

Dunque φ_{P_0} è univocamente determinata da $\varphi_{P_0} = F_{f(P_0)} \circ f \circ F_{P_0}^{-1}$.

Osservazioni: Sia φ_{P_0} l'applicazione trovata nel punto precedente. Allora:

- 1) $A \ni P \xrightarrow{F_{P_0}} \overline{P_0 P} \xrightarrow{\varphi_{P_0}} \varphi_{P_0}(\overline{P_0 P}) \xrightarrow{F_{f(P_0)}^{-1}} f(P_0) + \varphi_{P_0}(\overline{P_0 P})$.
Dunque $f(P) = f(P_0) + \varphi_{P_0}(\overline{P_0 P})$ e quindi $f(P_0 + v) = f(P_0) + \varphi_{P_0}(v)$.
- 2) $V \ni v \xrightarrow{F_{P_0}^{-1}} P_0 + v \xrightarrow{f} f(P_0 + v) \xrightarrow{F_{f(P_0)}} \overline{f(P_0) f(P_0 + v)}$
Dunque $\varphi_{P_0}(v) = \overline{f(P_0) f(P_0 + v)}$.

PROPOSIZIONE 5.2.12: f è affine $\Leftrightarrow \varphi_{P_0}$ è lineare.

Dimostrazione:

- \Rightarrow $f(P_0) + \varphi_{P_0}(\sum_{i=1}^k t_i v_i) \stackrel{\square}{=} f(P_0 + \sum_{i=1}^k t_i v_i) = f((1 - \sum_{i=1}^k t_i)P_0 + \sum_{i=1}^k t_i(P_0 + v_i)) =$
 $= (1 - \sum_{i=1}^k t_i)f(P_0) + \sum_{i=1}^k t_i f(P_0 + v_i) \stackrel{\square}{=} (1 - \sum_{i=1}^k t_i)f(P_0) + \sum_{i=1}^k t_i (f(P_0) + \varphi_{P_0}(v_i)) =$
 $= f(P_0) + \sum_{i=1}^k t_i \varphi_{P_0}(v_i)$, da cui:
 $\varphi_{P_0}(\sum_{i=1}^k t_i v_i) = \sum_{i=1}^k t_i \varphi_{P_0}(v_i)$, cioè φ_{P_0} è lineare (i passaggi con $\stackrel{\square}{=}$ seguono dalle osservazioni precedenti).
- \Leftarrow Siano $P_0, \dots, P_k \in A$, $t_1 + \dots + t_k = 1$.
 $f(t_0 P_0 + \dots + t_k P_k) = f(P_0 + t_1 \overline{P_0 P_1} + \dots + t_k \overline{P_0 P_k}) = f(P_0) + \varphi_{P_0}(\sum_{i=1}^k t_i \overline{P_0 P_i}) =$
 $= f(P_0) + \sum_{i=1}^k t_i \varphi_{P_0}(\overline{P_0 P_i}) = f(P_0) + \sum_{i=1}^k t_i \overline{f(P_0) f(P_i)} = t_0 f(P_0) + \dots + t_k f(P_k)$, cioè f è affine.

PROPOSIZIONE 5.2.13: Sia $f: A \rightarrow B$ trasformazione affine. $\forall P_0, P_1 \in A$, $\varphi_{P_0} = \varphi_{P_1}$.

Dimostrazione:

Valgono le relazioni:

$$\begin{aligned} f(P_0 + v) &= f(P_0) + \varphi_{P_0}(v); \\ f(P_1 + v) &= f(P_1) + \varphi_{P_1}(v), \forall v \in V. \end{aligned}$$

Queste relazioni possono essere riscritte nella forma:

$$\begin{aligned} \varphi_{P_0}(v) &= f(P_0 + v) - f(P_0); \\ \varphi_{P_1}(v) &= f(P_1 + v) - f(P_1). \end{aligned}$$

Sia $P = P_0 + v$. Allora $P_1 + v = P_1 + P - P_0$. Dunque:

$$\begin{aligned} \varphi_{P_1}(v) &= \underbrace{f(P_1 + P - P_0) - f(P_1)}_{\text{comb. affine}} = f(P_1) + f(P) - f(P_0) - f(P_1) = f(P) - f(P_0) = \\ &= f(P_0 + v) - f(P_0) = \varphi_{P_0}(v), \end{aligned}$$

da cui la tesi.

COROLLARIO 5.2.14: $f: A \rightarrow B$ è affine $\Leftrightarrow \exists \varphi: V \rightarrow W$ applicazione lineare tale che $\forall P_0 \in A$, $f(P) = f(P_0) + \varphi(\overline{P_0 P})$.

DEFINIZIONE 5.2.18: Questa $\varphi: V \rightarrow W$ prende il nome di **applicazione lineare associata** a f .

Osservazione: Sia $f: A \rightarrow B$ applicazione affine con applicazione lineare associata $\varphi: V \rightarrow W$.

Sia $g: B \rightarrow C$ applicazione affine con applicazione lineare associata $\psi: W \rightarrow Z$.

Sia $P_0 \in A, Q_0 = f(P_0)$.

Allora:

$$f(P) = f(P_0) + \varphi(\overrightarrow{P_0P}) \quad \forall P \in A,$$

$$g(Q) = g(Q_0) + \psi(\overrightarrow{Q_0Q}) \quad \forall Q \in B.$$

$$(g \circ f)(P) = g(f(P)) = g(Q_0) + \psi(\underbrace{\overrightarrow{f(P_0)f(P)}}_{\varphi(\overrightarrow{P_0P})}) = g(f(P_0)) + (\psi \circ \varphi)(\overrightarrow{P_0P}).$$

PROPOSIZIONE 5.2.15: Se $f: A \rightarrow B$ è affine e invertibile, allora $f^{-1}: B \rightarrow A$ è affine.

Dimostrazione:

Sia $\varphi: V \rightarrow W$ l'applicazione lineare associata. φ è invertibile e quindi isomorfismo.

Sia $P_0 \in A, Q_0 = f(P_0)$.

Sia $g: B \rightarrow A$ definita da $g(Q) = f^{-1}(Q_0) + \varphi^{-1}(\overrightarrow{Q_0Q})$.

Allora $(g \circ f)(P_0) = g(Q_0) = f^{-1}(Q_0) + \varphi^{-1}(\overrightarrow{Q_0Q_0}) = P_0$.

Per la formula precedente:

$$(g \circ f)(P) = P_0 + (\varphi^{-1} \circ \varphi)(\overrightarrow{P_0P}) = P_0 + \overrightarrow{P_0P} = P \quad \forall P.$$

Dunque $f^{-1}(Q) = f^{-1}(Q_0) + \varphi^{-1}(\overrightarrow{Q_0Q})$, da cui la tesi.

Osservazione: Perciò $Aff(A)$ è un gruppo.

PROPOSIZIONE 5.2.16: A spazio affine su V . $\{P_0, \dots, P_n\}, \{Q_0, \dots, Q_n\}$ riferimenti affini di A .

Allora esiste una sola affinità f di A tale che $f(P_i) = Q_i \quad \forall 0 \leq i \leq n$.

Dimostrazione:

$\forall P \in A, P$ si scrive in modo unico come $P = \sum_{i=0}^n t_i P_i$, con $\sum_{i=0}^n t_i = 1$.

Necessariamente dobbiamo definire $f(P) = \sum_{i=0}^n t_i Q_i$.

Verifico che f è un'affinità (che sarà dunque unica per quanto appena visto):

sia $\varphi \in GL(V)$ l'unico isomorfismo che trasforma $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}$ in $\overrightarrow{Q_0Q_1}, \dots, \overrightarrow{Q_0Q_n}$.

$$f(P) = \sum_{i=0}^n t_i Q_i = Q_0 + t_1 \overrightarrow{Q_0Q_1} + \dots + t_n \overrightarrow{Q_0Q_n} = f(P_0) + \sum_{i=1}^n t_i \varphi(\overrightarrow{P_0P_i}) = f(P_0) + \varphi\left(\sum_{i=1}^n t_i \overrightarrow{P_0P_i}\right)$$

Ma $P = \sum_{i=0}^n t_i P_i = P_0 + \sum_{i=1}^n t_i \overrightarrow{P_0P_i}$; inoltre $P = P_0 + \overrightarrow{P_0P} \Rightarrow \overrightarrow{P_0P} = \sum_{i=1}^n t_i \overrightarrow{P_0P_i}$.

Dunque $f(P) = f(P_0) + \varphi(\overrightarrow{P_0P}) \quad \forall P \in A$, da cui la tesi.

5.3 GEOMETRIA AFFINE EUCLIDEA

Notazione: Indicheremo con $X \cdot Y$ il prodotto scalare standard fra X e Y in \mathbb{R}^n .

DEFINIZIONE 5.3.1: $v \in \mathbb{R}^n$ si dice **ortogonale** al sottospazio affine S se $v \in W_S^\perp$.

Esempio: Se H è l'iperpiano $B \cdot X + d = 0$, allora B è ortogonale a H .

DEFINIZIONE 5.3.2: Due sottospazi affini S e S' si dicono **ortogonali** (e lo denoteremo $S \perp S'$) se $W_S \subseteq W_{S'}^\perp$ ($\Leftrightarrow W_{S'} \subseteq W_S^\perp$).

Esempi: 1) Siano r la retta $X = At + C$ e r' la retta $X = A't + C'$. Allora $r \perp r' \Leftrightarrow W_r \subseteq W_{r'}^\perp \Leftrightarrow A \in (\text{Span}(A'))^\perp \Leftrightarrow A \cdot A' = 0$.

2) r retta di equazione $X = At + C$ e H iperpiano di equazione $B \cdot X + d = 0$. Allora $W_H = \{X | B \cdot X = 0\}$ e $W_H^\perp = \text{Span}(B)$, da cui $r \perp H \Leftrightarrow A \parallel B$.

PROPOSIZIONE 5.3.1: Sia S un sottospazio affine di dimensione k . Allora:

- 1) $S' \perp S \Rightarrow \dim(S') \leq n - k$.
- 2) $\forall 0 \leq d \leq n - k$ esiste un sottospazio affine S' di dimensione d tale che $S' \perp S$.
- 3) Tutti i sottospazi affini S' ortogonali a S di dimensione massima ($\dim(S') = n - k$) sono paralleli e ciascuno di essi interseca S in uno e un solo punto.
- 4) $\forall P \in \mathbb{R}^n \exists!$ sottospazio affine S' ortogonale a S di dimensione massima passante per P .

Dimostrazione:

1) $W_S \subseteq W_{S'}^\perp$. Poiché $\dim(W_S) = k$, allora $\dim(W_{S'}^\perp) \geq k$ e dunque $\dim(W_{S'}) \leq n - k$.

2) Basta prendere S' tale che $\dim(S') = \dim(W_{S'}^\perp) = d$ e $W_{S'} \subseteq W_S^\perp$.

3) Sia $S = R + W_S$ e sia S' come nelle ipotesi; $S' = Q + W_{S'}$ con $W_{S'} \subseteq W_S^\perp$.

Poiché $\dim(W_{S'}) = \dim(W_S^\perp) = n - k$, si ha $W_{S'} = W_S^\perp$, e dunque ogni S' ha giacitura W_S^\perp .

Consideriamo $S' = Q + W_{S'}$. Visto che $\mathbb{R}^n = W_S \oplus W_S^\perp$, si ha $R - Q = v + w$, con $v \in W_S$ e $w \in W_S^\perp$. Dunque $R - v = Q + w \in S \cap S'$, poiché $R - v \in R + W_S = S$ e $Q + w \in Q + W_S^\perp = S'$. Per ragioni di dimensione, il punto è unico.

4) Basta prendere $S' = P + W_S^\perp$.

Esempio: Sia r una retta in \mathbb{R}^3 e $P \in \mathbb{R}^3$. Esiste un unico piano passante per P e ortogonale a r ; tale piano interseca r in uno e un solo punto.

In particolare, se r ha equazione $X = At + C$, allora $W_r = \text{Span}(A)$; visto che $W_r \subseteq W_H^\perp$, per ragioni di dimensione $W_r = W_H^\perp$, cioè $W_H^\perp = \text{Span}(A)$. Dunque H ha equazione $A \cdot X = A \cdot P$.

Osservazione: Siano H e H' due iperpiani in \mathbb{R}^n , di equazione rispettivamente $B \cdot X + d = 0$ e $B' \cdot X + d' = 0$; H e H' si dicono ortogonali $\Leftrightarrow B \perp B' \Leftrightarrow B \cdot B' = 0$.

DEFINIZIONE 5.3.3: Sia $P \in \mathbb{R}^n$ e S un sottospazio affine di \mathbb{R}^n . Definiamo **distanza** di P da S $d(P, S) = \inf_{X \in S} d(P, X)$.

PROPOSIZIONE 5.3.2: $\exists P_0 \in S$ tale che $d(P, S) = \|P - P_0\|$ (e quindi l'inf è un minimo).

Dimostrazione:

Sia $\dim(S) = k$. Per la proposizione precedente $\exists!$ S' sottospazio affine ortogonale a S , di dimensione massima e passante per P ; sia P_0 tale che $S \cap S' = \{P_0\}$.

Osserviamo che $(P - P_0) \perp S$.

Vogliamo mostrare che $d(P, S) = \|P - P_0\|$, cioè che $\forall X \in S, X \neq P_0, d(P, X) > \|P - P_0\|$.

Abbiamo:

$$\begin{aligned} d(P, X)^2 &= \|P - X\|^2 = \|(P - P_0) + (P_0 - X)\|^2 = \\ &= ((P - P_0) + (P_0 - X)) \cdot ((P - P_0) + (P_0 - X)) = \\ &= d(P, P_0)^2 + d(P_0, X)^2 + 2 \underbrace{(P_0 - X) \cdot (P - P_0)}_{=0} = d(P, P_0)^2 + \underbrace{d(P_0, X)^2}_{>0} > \\ &> \|P - P_0\|^2, \end{aligned}$$

da cui la tesi.

COROLLARIO 5.3.3: H iperpiano di \mathbb{R}^n di equazione $B \cdot X + d = 0, P \in \mathbb{R}^n$. Allora $d(P, H) = \frac{|B \cdot P + d|}{\|B\|}$.

Dimostrazione:

$d(P, H) = \|P - P_0\|$, dove $P_0 = H \cap r$ con r la retta per P ortogonale a H . r ha equazione $X = Bt + P$.

Per determinare $P_0 = H \cap r$, cerco t tale che $B \cdot (Bt + P) + d = 0$, cioè $t = \frac{-d - B \cdot P}{\|B\|^2}$, ossia

$$P_0 = \frac{-d - B \cdot P}{\|B\|^2} B + P.$$

$$\text{Finalmente abbiamo } d(P, H) = \|P - P_0\| = \left\| \frac{d + B \cdot P}{\|B\|^2} B \right\| = \frac{|d + B \cdot P|}{\|B\|^2} \|B\| = \frac{|d + B \cdot P|}{\|B\|}.$$

Osservazione: Il lettore può trovare in modo analogo la formula per la distanza di un punto da una retta in \mathbb{R}^3 .

DEFINIZIONE 5.3.4: S, S' sottospazi affini di \mathbb{R}^n . Definiamo **distanza** fra S e S' $d(S, S') = \inf_{X \in S, Y \in S'} d(X, Y)$.

DISTANZA FRA DUE PIANI IN \mathbb{R}^3 :

- Se $H_1 \cap H_2 \neq \emptyset, d(H_1, H_2) = 0$.
- Se $H_1 \parallel H_2, d(H_1, H_2) = d(P, H_2) \forall P \in H_1$, che si può calcolare con la formula precedente.

DISTANZA RETTA-PIANO IN \mathbb{R}^3 :

- Se $r \cap H \neq \emptyset, d(r, H) = 0$.
- Se $r \parallel H, d(r, H) = d(P, H) \forall P \in r$, che si calcola ancora con la formula precedente.

DISTANZA FRA DUE RETTE IN \mathbb{R}^3 :

- Se $r_1 \cap r_2 \neq \emptyset, d(r_1, r_2) = 0$.
- Se $r_1 \parallel r_2, d(r_1, r_2) = d(P, r_2) \forall P \in r_1$.
- Se le rette sono sghembe, diciamo che r_1 e r_2 hanno equazione rispettivamente $X = A_1 t + C_1$ e $X = A_2 t + C_2$; proviamo che $\exists!$ l retta ortogonale a r_1 e r_2 che le interseca entrambe. Se P_1 e P_2 sono i punti di intersezione, evidentemente si ha che $d(r_1, r_2) = \|P_1 - P_2\|$.

Dimostrazione:

Il generico punto di r_1 è $P(t) = A_1 t + C_1$, mentre il generico punto di r_2 è $Q(\theta) = A_2 \theta + C_2$. La retta $l(t, \theta)$ congiungente $P(t)$ e $Q(\theta)$ è ovviamente incidente sia a r_1 che a r_2 ; provo che $\exists! (t, \theta)$ tale che $l(t, \theta)$ è ortogonale a entrambe le rette.

Poiché l è parallela al vettore $P(t) - Q(\theta) = A_1 t + C_1 - A_2 \theta - C_2$, basta imporre:

$$\begin{cases} (A_1 t + C_1 - A_2 \theta - C_2) \cdot A_1 = 0 \\ (A_1 t + C_1 - A_2 \theta - C_2) \cdot A_2 = 0 \end{cases}$$

La matrice dei coefficienti di questo sistema 2×2 è:

$$M = \begin{pmatrix} A_1 \cdot A_1 & -A_1 \cdot A_2 \\ A_1 \cdot A_2 & -A_2 \cdot A_2 \end{pmatrix};$$

Se $A_1 = (\alpha_1 \ \beta_1 \ \gamma_1)$ e $A_2 = (\alpha_2 \ \beta_2 \ \gamma_2)$, si ha $\det(M) = -(A_1 \cdot A_1)(A_2 \cdot A_2) + (A_1 \cdot A_2)^2 = -(\alpha_1 \beta_2 - \alpha_2 \beta_1)^2 - (\alpha_1 \gamma_2 - \alpha_2 \gamma_1)^2 - (\beta_1 \gamma_2 - \beta_2 \gamma_1)^2$, dunque $\det(M) = 0 \Leftrightarrow$ la matrice:

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{pmatrix}$$

ha rango 1, cioè A_1 e A_2 sono linearmente dipendenti. Ma le rette sono sghembe, dunque A_1 e A_2 sono linearmente indipendenti e dunque $\det(M) \neq 0$.

Da questo segue l'unicità (e l'esistenza) della soluzione (t_0, θ_0) del sistema; i punti $P_1 = P(t_0)$ e $P_2 = Q(\theta_0)$ sono quelli cercati.

5.4 AFFINITÀ DI \mathbb{K}^n

Osservazione: $Aff(\mathbb{K}^n) = \{X \rightarrow MX + N \mid M \in GL(n, \mathbb{K}), N \in \mathbb{K}^n\}$.

Possiamo vedere $Aff(\mathbb{K}^n)$ come un sottogruppo di $GL(n+1, \mathbb{K})$, infatti:

sia $H = \left\{ X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} \in \mathbb{K}^{n+1} \mid x_{n+1} = 1 \right\}$. H è sottospazio affine di \mathbb{K}^{n+1} con giacitura

$$W_H = \{X \in \mathbb{K}^{n+1} \mid x_{n+1} = 0\}.$$

Sia $f: \mathbb{K}^n \rightarrow H \mid X \rightarrow \begin{pmatrix} X \\ 1 \end{pmatrix}$ (se $X \in \mathbb{K}^n$, la notazione $\begin{pmatrix} X \\ 1 \end{pmatrix}$ indica il vettore X' di \mathbb{K}^{n+1} tale che

$$x'_i = x_i \ \forall i \leq n \text{ e } x'_{n+1} = 1).$$

Poiché $\begin{pmatrix} X \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{=e_{n+1}} + \begin{pmatrix} X \\ 0 \end{pmatrix}$, allora $f = \tau_{e_{n+1}} \circ \varphi$, dove $\varphi: \mathbb{K}^n \rightarrow W_H \mid X \rightarrow \begin{pmatrix} X \\ 0 \end{pmatrix}$ è un isomorfismo

lineare.

Allora f è un isomorfismo affine e $\mathbb{K}^n \cong_{aff} H$.

Sia $G(H) = \{g \in GL(n+1, \mathbb{K}) \mid g(H) = H\}$ e sia $g \in G(H)$.

Allora:

$$g(Y) = \begin{pmatrix} M & N \\ tP & q \end{pmatrix} \cdot Y$$

per certi $M \in \mathcal{M}(n, \mathbb{K})$, $N, P \in \mathbb{K}^n$, $q \in \mathbb{K} \ \forall Y \in \mathbb{K}^{n+1}$.

Se $Y = \begin{pmatrix} X \\ 1 \end{pmatrix} \in H$, allora:

$$g\left(\begin{pmatrix} X \\ 1 \end{pmatrix}\right) = \begin{pmatrix} M & N \\ tP & q \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix} = \begin{pmatrix} * \\ tPX + q \end{pmatrix}$$

e dunque ${}^tPX + q = 1 \quad \forall X \in \mathbb{K}^n$, cioè $P = 0, q = 1$.

Si ha perciò:

$$G(H) = \left\{ \widetilde{M}_N = \begin{pmatrix} M & | & N \\ 0 & | & 1 \end{pmatrix} \mid M \in GL(n, \mathbb{K}), N \in \mathbb{K}^n \right\}.$$

Osservazione: $\begin{pmatrix} M_1 & | & N_1 \\ 0 & | & 1 \end{pmatrix} \begin{pmatrix} M_2 & | & N_2 \\ 0 & | & 1 \end{pmatrix} = \begin{pmatrix} M_1M_2 & | & M_1N_2 + N_1 \\ 0 & | & 1 \end{pmatrix} \Rightarrow (G(H), \circ)$ è un sottogruppo di $GL(n+1, \mathbb{K})$.

PROPOSIZIONE 5.4.1: $L: (Aff(\mathbb{K}^n), \circ) \rightarrow (G(H), \circ) \mid (X \rightarrow MX + N) \rightarrow \begin{pmatrix} M & | & N \\ 0 & | & 1 \end{pmatrix}$ è un isomorfismo di gruppi (e dunque $Aff(\mathbb{K}^n)$ è isomorfo a un sottogruppo di $GL(n+1, \mathbb{K})$).

Dimostrazione:

Siano $f_1, f_2 \in Aff(\mathbb{K}^n)$; $f_1(X) = M_1X + N_1, f_2(X) = M_2X + N_2$.

Vediamo che $(f_1 \circ f_2)(X) = f_1(M_2X + N_2) = M_1(M_2X + N_2) + N_1 = M_1M_2X + M_1N_2 + N_1$.

Poiché prima abbiamo visto che:

$$L(f_1)L(f_2) = \begin{pmatrix} M_1 & | & N_1 \\ 0 & | & 1 \end{pmatrix} \begin{pmatrix} M_2 & | & N_2 \\ 0 & | & 1 \end{pmatrix} = \begin{pmatrix} M_1M_2 & | & M_1N_2 + N_1 \\ 0 & | & 1 \end{pmatrix},$$

allora effettivamente $L(f_1 \circ f_2) = L(f_1)L(f_2)$, cioè L è lineare.

Poiché è evidentemente bigettiva, ho la tesi.

DEFINIZIONE 5.4.1: Sia G un gruppo di trasformazioni di \mathbb{K}^n .

Due sottoinsiemi F_1, F_2 di \mathbb{K}^n sono detti **G-equivalenti** se $\exists g \in G \mid g(F_1) = F_2$.

DEFINIZIONE 5.4.2: $F_1, F_2 \subseteq \mathbb{K}^n$ sono detti **affinamente equivalenti** (rispettivamente, **metricamente equivalenti**) se $\exists g \in Aff(\mathbb{K}^n)$ (rispettivamente, $g \in Isom(\mathbb{K}^n)$) tale che $g(F_1) = F_2$.

Notazione: Se F_1 e F_2 sono affinamente equivalenti, lo indicheremo con $F_1 \sim_{aff} F_2$.

Esempi: $A = \mathbb{K}^n$.

1) Siano $F_1 = \{P_0, \dots, P_k\}, F_2 = \{Q_0, \dots, Q_k\}$ $(k+1)$ -uple di punti di \mathbb{K}^n affinamente indipendenti. Abbiamo visto che $\exists g \in Aff(\mathbb{K}^n) \mid g(P_i) = Q_i \quad \forall i$ e dunque F_1 e F_2 sono affinamente equivalenti.

2) H_1, H_2 iperpiani affini di \mathbb{K}^n .

Allora $\exists g \in Aff(\mathbb{K}^n) \mid g(H_1) = H_2$, infatti, se $H_1 = Comb_a(P_0, \dots, P_{n-1})$ e

$H_2 = Comb_a(Q_0, \dots, Q_{n-1})$, allora scegliendo P_n, Q_n in modo che sia $i P_i$ che $i Q_i$ siano un riferimento affine, so che $\exists g \in Aff(\mathbb{K}^n) \mid g(P_i) = Q_i \quad \forall i$, e dunque $g(H_1) = H_2$.

Dunque $\{iperpiani\text{ affini di } \mathbb{K}^n\} / \sim_{aff}$ ha una sola classe di equivalenza.

DEFINIZIONE 5.4.3: $\forall g \in \mathbb{K}[x_1, \dots, x_n]$ (dove $\mathbb{K}[x_1, \dots, x_n]$ rappresenta l'anello dei polinomi in x_1, \dots, x_n), definiamo **luogo di zeri** di g l'insieme $V(g) = \{X \in \mathbb{K}^n \mid g(X) = 0\}$.

Osservazione: L'applicazione $V: \mathbb{K}[x_1, \dots, x_n] \rightarrow \mathbb{K}^n \mid g \rightarrow V(g)$ non è iniettiva, infatti:

- $V(\alpha g) = V(g) \quad \forall \alpha \in \mathbb{K} \setminus \{0\}$;

- $V(g^m) = V(g) \quad \forall m \in \mathbb{N}^+$;
- Se $\mathbb{K} = \mathbb{C}$ e g_1, g_2 non contengono fattori multipli, allora:
 $V(g_1) = V(g_2) \Leftrightarrow \exists \alpha \in \mathbb{C} \setminus \{0\} \mid g_1 = \alpha g_2$.
- Se $\mathbb{K} = \mathbb{R}$ la proprietà precedente non vale, infatti $\forall c > 0, V(x^2 + y^2 + c) = \emptyset$.

DEFINIZIONE 5.4.4: $g_1, g_2 \in \mathbb{K}[x_1, \dots, x_n]$. Diciamo che g_1, g_2 sono **proporzionali** $g_1 \sim g_2 \Leftrightarrow \exists \alpha \in \mathbb{K} \setminus \{0\} \mid g_1 = \alpha g_2$.

Osservazione: La relazione di proporzionalità fra polinomi è di equivalenza in $\mathbb{K}[x_1, \dots, x_n]$ (la verifica è immediata).

DEFINIZIONE 5.4.5: Definiamo **ipersuperficie affine** (o semplicemente **ipersuperficie**) ogni classe di proporzionalità di polinomi di $\mathbb{K}[x_1, \dots, x_n]$ di grado positivo.

DEFINIZIONE 5.4.6: Se $I = [g]$ è ipersuperficie, $g(X) = 0$ è detta **equazione** di I e $V(g) \subseteq \mathbb{K}^n$ è detto **supporto** di I (indicato anche come $Supp(I)$).

DEFINIZIONE 5.4.7: Se $I = [g]$ è ipersuperficie, definiamo grado di I , $\deg([g]) = \deg(g)$.

Osservazione: È una buona definizione, poiché se $g' \sim g \Rightarrow \deg(g_1) = \deg(g_2)$.

- Se $n = 2$, I è detta **curva affine**;
- se $n = 3$, I è detta **superficie affine**;
- le ipersuperfici di grado 2 sono dette **quadriche** (**coniche** se $n = 2$).

Osservazione: Come visto, l'ipersuperficie determina il suo supporto, non il viceversa.

È dunque improprio parlare di equivalenza affine solo per i supporti, in quanto due ipersuperfici possono avere lo stesso luogo di zeri.

Introduciamo quindi il concetto di equivalenza affine anche per le ipersuperfici:

DEFINIZIONE 5.4.8: $I = [g]$ ipersuperficie. $\psi(X) = MX + N \in Aff(\mathbb{K}^n)$.

Definiamo **ipersuperficie "rimontata"** di I tramite ψ l'ipersuperficie $\psi^{-1}(I)$ di equazione $g(\psi(X)) = 0$.

Osservazione:

$$\begin{array}{ccc} \mathbb{K}^n \supseteq V(g \circ \psi) & \xrightarrow{\psi} & V(g) \subseteq \mathbb{K}^n \\ & \searrow \swarrow & \\ & \mathbb{K} & \end{array}$$

La definizione precedente è coerente con il fatto che ψ trasforma il supporto di $\psi^{-1}(I)$ nel supporto di I , infatti:

$$x_0 \in Supp(\psi^{-1}(I)) \Leftrightarrow g(\psi(x_0)) = 0 \Leftrightarrow \psi(x_0) \in V(g) = Supp(I).$$

DEFINIZIONE 5.4.9: Due ipersuperfici affini I e J si dicono affinemente equivalenti se $\exists \psi \in Aff(\mathbb{K}^n) \mid I = \psi^{-1}(J)$.

In altre parole, $I = [f]$ e $J = [g]$ sono affinemente equivalenti se $\exists \psi \in Aff(\mathbb{K}^n) \mid f = g \circ \psi$.

Osservazione:

$$\mathbb{K}[x_1, \dots, x_n] / \sim_{aff} \ni I \xleftarrow{\psi^{-1}} J \in \mathbb{K}[x_1, \dots, x_n] / \sim_{aff}$$

$$\begin{array}{ccc} V \downarrow & & \downarrow V \\ \mathbb{K}^n \ni \text{Supp}(I) & \xrightarrow{\psi} & \text{Supp}(J) \ni \mathbb{K}^n \end{array}$$

$I \sim_{aff} J \Rightarrow \text{Supp}(I) \sim_{aff} \text{Supp}(J)$. Il viceversa è falso.

Osservazione: \sim_{aff} classifica dunque i polinomi (e non i supporti) a meno di coordinate affini.

Osservazione: $I = [g]$ ipersuperficie | $V(g)$ è iperpiano. Allora $\deg(g) = 1$.

Sia $g(X) = {}^tAX + b$, $A \in \mathbb{K}^n \setminus \{0\}$, $b \in \mathbb{K}$.

Sia $\psi(X) = MX + N$ affinità di \mathbb{K}^n .

Allora $(g \circ \psi)(X) = {}^tA(MX + N) + b = ({}^tAM)X + {}^tAN + b$.

Dunque se $J = [f]$ è un altro iperpiano, $f(X) = {}^tA'X + b'$.

$I \sim_{aff} J \Leftrightarrow \exists \alpha \neq 0, \exists M \in GL(n, \mathbb{K}), \exists N \in \mathbb{K}^n |$

$$\begin{cases} {}^tA' = \alpha {}^tAM \\ b' = \alpha({}^tAN + b) \end{cases}$$

Poiché il sistema ha sempre soluzione, troviamo che due qualsiasi iperpiani sono affinemente equivalenti anche come ipersuperfici, non solo come supporti.

5.5 QUADRICHE

DEFINIZIONE 5.5.1: Il supporto di un'ipersuperficie si dice **cono** se vale la seguente proprietà: se contiene un punto P , allora contiene tutti i punti $tP \ \forall t \in \mathbb{K}$.

Osservazione: Prendiamo una quadrica $I = [g]$, dunque $\deg(g) = 2$.

L'equazione generica della quadrica è:

$g(X) = {}^tXAX + 2{}^tBX + c$, con $A \in \mathcal{S}(n, \mathbb{K})$, $A \neq 0$, $B \in \mathbb{K}^n$, $c \in \mathbb{K}$.

Esempio: $g \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 2x_1x_2 + 3x_2^2 + 4x_2 + 5$.

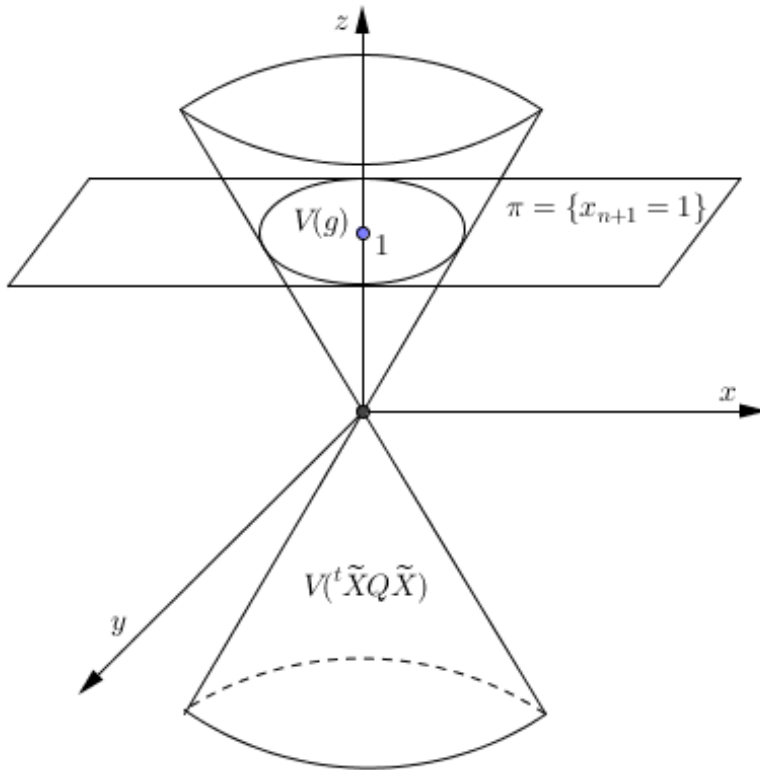
$$A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}; B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}; c = 5.$$

Osservazione: Se denoto l'equazione con $Q = \begin{pmatrix} \boxed{A} & B \\ {}^tB & c \end{pmatrix}$, allora, ponendo $\tilde{X} = \begin{pmatrix} X \\ 1 \end{pmatrix}$:

$$\begin{aligned} {}^t\tilde{X}Q\tilde{X} &= ({}^tX \mid 1) \begin{pmatrix} \boxed{A} & B \\ {}^tB & c \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix} = ({}^tXA + {}^tB \mid {}^tXB + c) \begin{pmatrix} X \\ 1 \end{pmatrix} = \\ &= {}^tXAX + \underbrace{{}^tBX + {}^tXB}_{\text{sono numeri}} + c = {}^tXAX + 2{}^tBX + c = g(X). \end{aligned}$$

Dunque $V(g) = V({}^t\tilde{X}Q\tilde{X})$.

Osservazione: L'equazione ${}^t\tilde{X}Q\tilde{X}$ è omogenea di secondo grado, quindi $V({}^t\tilde{X}Q\tilde{X})$ è un cono. Dunque vedo $V(g)$ come un cono in \mathbb{K}^{n+1} intersecato con l'iperpiano $x_{n+1} = 1$:



Esempio: La figura precedente mostra l'equivalenza fra $x_1^2 + x_2^2 - 1 = 0$ e $\begin{cases} x_1^2 + x_2^2 - x_3^2 = 0 \\ x_3 = 1 \end{cases}$.

Osservazione: Sia I la quadrica di equazione ${}^t\tilde{X}Q\tilde{X} = 0$.

Sia $\psi(X) = MX + N$ affinità; calcoliamo l'equazione della quadrica $\psi^{-1}(I)$.

Poniamo $\widetilde{\psi(X)} = \widetilde{M}_N \tilde{X} = \begin{pmatrix} M & | & N \\ 0 & | & 1 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix} = \begin{pmatrix} MX + N \\ 1 \end{pmatrix}$.

Dunque $\psi^{-1}(I)$ ha equazione:

$${}^t(\widetilde{\psi(X)})Q(\widetilde{\psi(X)}) = {}^t\tilde{X}{}^t\widetilde{M}_N Q \widetilde{M}_N \tilde{X} = 0.$$

Dunque la matrice associata alla quadrica $\psi^{-1}(I)$ è ${}^t\widetilde{M}_N Q \widetilde{M}_N$.

Perciò studiare $\{\text{quadriche di } \mathbb{K}^n\} / \sim_{aff}$ corrisponde a studiare

$$\left\{ Q = \left(\begin{array}{c|c} \boxed{A} & B \\ \hline {}^tB & c \end{array} \right) \middle| A \in \mathcal{S}(n, \mathbb{K}) \setminus \{0\} \right\} / \sim_{aff},$$

dove $Q \sim_{aff} Q' \Leftrightarrow \exists \alpha \in \mathbb{K} \setminus \{0\}, \exists \widetilde{M}_N \in Aff(\mathbb{K}^n) \mid Q' = \alpha {}^t\widetilde{M}_N Q \widetilde{M}_N$.

Vogliamo trovare dei “**rappresentanti canonici**” per equivalenza affine, cioè una famiglia $\{F_1, \dots, F_k\}$ di quadriche di \mathbb{K}^n tali che:

- $\forall J$ quadrica di \mathbb{K}^n , $\exists i | J \sim_{aff} F_i$;
- $F_i \not\sim_{aff} F_j \quad \forall i, j$.

Questi rappresentanti prendono il nome di **forme canoniche**.

Restringiamoci al caso $\mathbb{K} = \mathbb{R} \vee \mathbb{K} = \mathbb{C}$.

Cominciamo dal caso $n = 2$, cioè dalle coniche di \mathbb{K}^2 .

CLASSIFICAZIONE AFFINE DELLE CONICHE:

Sia C la conica di equazione $g(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33}$, $a_{ij} \in \mathbb{K} = \mathbb{R} \vee \mathbb{C}$, a_{11}, a_{12}, a_{22} non tutti nulli.

Se $\tilde{X} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$, $c = a_{33}$, $Q = \begin{pmatrix} \boxed{A} & B \\ {}^tB & c \end{pmatrix}$, allora

C ha equazione ${}^t\tilde{X}Q\tilde{X} = 0$.

Se $\psi(X') = MX' + N$ è affinità, allora $\psi^{-1}(C)$ ha equazione ${}^t\tilde{X}'{}^t\tilde{M}_N Q \tilde{M}_N \tilde{X}' = 0$.

$$Q' = {}^t\tilde{M}_N Q \tilde{M}_N = \begin{pmatrix} {}^tM & | & 0 \\ {}^tN & | & 1 \end{pmatrix} \begin{pmatrix} \boxed{A} & B \\ {}^tB & c \end{pmatrix} \begin{pmatrix} M & | & N \\ 0 & | & 1 \end{pmatrix} = \begin{pmatrix} {}^tMAM & | & {}^tMAN + {}^tMB \\ {}^tN{}^tAM + {}^tBM & | & {}^tNAN + 2{}^tNB + c \end{pmatrix}$$

Quindi $A' = {}^tMAM$, $B' = {}^tM(AN + B)$.

Osservazione: M invertibile e $A \neq 0 \Rightarrow A' \neq 0$, quindi l'affinità trasforma la conica in una conica.

Osservazione: Q e Q' sono congruenti e A e A' sono congruenti $\Rightarrow rk(A)$ e $rk(Q)$ sono invarianti per equivalenza affine (non cambiano se Q è moltiplicata per $\alpha \neq 0$).

DEFINIZIONE 5.5.2: C è detta **degenere** se $\det(Q) = 0$. Più precisamente, si dice:

- **semplicemente degenere** se $rk(Q) = 2$;
- **doppiamente degenere** se $rk(Q) = 1$.

DEFINIZIONE 5.5.3: $C = [g]$ è detta conica a centro se $\exists N \in \mathbb{K}^2 | g(X) = g(\sigma_N(X)) \quad \forall X$, dove σ_N è la simmetria centrale di centro N .

Osservazione: Se $N = (0,0)$, $\sigma_N(x, y) = (-x, -y)$, dunque $(0,0)$ è centro per $C = [g] \Leftrightarrow g(x, y)$ non contiene monomi di primo grado (ossia $B = 0$).

In generale, se $N \in \mathbb{K}^2$ è un centro per C e considero la traslazione $\tau(X) = X + N$, allora $\tau^{-1}(C)$ ha centro in $(0,0)$, e quindi $B' = 0$.

Poiché $B' = {}^tM(AN + B) = AN + B$, in quanto in una traslazione $M = {}^tM = I$, allora N è centro per $C \Leftrightarrow AN = -B$.

In altre parole, i centri della conica C sono le soluzioni del sistema $AY = -B$.

Per la trattazione delle coniche, distinguiamo il caso delle coniche a centro da quelle non a centro:

CONICHE NON A CENTRO: Sia C la conica non a centro tale che

$$Q = \begin{pmatrix} \boxed{A} & B \\ {}^tB & c \end{pmatrix}.$$

Siamo nel caso in cui il sistema $AY = -B$ non ha soluzione, dunque $rk(A) = 1$, poiché $0 < rk(A) < 2$.

Allora $\exists M \in GL(2, \mathbb{K}) \mid {}^tMAM = \begin{pmatrix} \pm 1 & 0 \\ 0 & 0 \end{pmatrix}$, quindi con la trasformazione lineare $X \rightarrow MX$ (ossia \widetilde{M}_0), ed eventualmente cambiando di segno all'equazione, Q diventa:

$$Q_1 = \begin{pmatrix} \boxed{1} & 0 & b_1 \\ 0 & 0 & b_2 \\ b_1 & b_2 & d \end{pmatrix}.$$

Poiché C non ha centri, $b_2 \neq 0$ (altrimenti il sistema $AY = -B$ avrebbe infinite soluzioni) e quindi $rk(Q) = 3$.

Vediamo che $\exists N = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \mid$ la traslazione $X \rightarrow X + N$ trasforma Q_1 in:

$$Q_2 = \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 0 & c_2 \\ 0 & c_2 & 0 \end{pmatrix}, \text{ con } c_2 \neq 0.$$

Infatti impongo che:

$$\begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ * \end{pmatrix} \\ \begin{pmatrix} \alpha & \beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + 2 \begin{pmatrix} \alpha & \beta \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + d = 0 \end{cases} \begin{cases} \alpha + b_1 = 0 \\ \alpha^2 + 2\alpha b_1 + 2\beta b_2 + d = 0 \end{cases} \begin{cases} \alpha = -b_1 \\ 2\beta b_2 = b_1^2 - d' \end{cases}$$

che ha soluzione perché $b_2 \neq 0$.

Infine con la trasformazione $\begin{cases} x = x' \\ y = \frac{y'}{2c_2} \end{cases}$, cioè con l'affinità:

$$M' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2c_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ si ottiene l'equazione } x'^2 - y' = 0, \text{ ossia:}$$

$$Q_3 = \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}, \text{ chiamato tipo } \mathbf{C}_1 \text{ o parabola.}$$

CONICHE A CENTRO: Sia C una conica a centro.

Passo 1): Eliminazione dei termini di primo grado con una traslazione.

In particolare, se N è il centro di C , con la traslazione $\tau: X \rightarrow X + N$ si ottiene la nuova conica $\tau^{-1}(C)$ che ha per matrice associata:

$$Q_1 = \begin{pmatrix} \boxed{A} & 0 \\ 0 & d \end{pmatrix}.$$

Se $d \neq 0$ posso dividere l'equazione per d , ossia posso supporre $d = 0$ o $d = 1$.

A si modifica per congruenza, quindi la forma canonica dipende dal campo.

Caso $\mathbb{K} = \mathbb{C}$

Il rango di A è un invariante completo per congruenza, quindi:

- se $rk(A) = 2$, A è congruente a $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;
- se $rk(A) = 1$, A è congruente a $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Vediamo dunque come si semplifica l'equazione di C a seconda della coppia $(rk(A), rk(Q))$:

- a) $\begin{cases} rk(A) = 2 \\ rk(Q) = 3 \end{cases}$. Quindi $d = 1$. Allora $C \sim_{aff} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, che ha equazione $x^2 + y^2 + 1 = 0$;
- b) $\begin{cases} rk(A) = 2 \\ rk(Q) = 2 \end{cases}$. Quindi $d = 0$. Allora $C \sim_{aff} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, che ha equazione $x^2 + y^2 = 0$
(dunque il supporto è l'unione delle rette incidenti $x - iy = 0$ e $x + iy = 0$);
- c) $\begin{cases} rk(A) = 1 \\ rk(Q) = 2 \end{cases}$. Quindi $d = 1$. Allora $C \sim_{aff} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, che ha equazione $x^2 + 1 = 0$
(ossia il supporto è l'unione delle rette parallele $x - i = 0$ e $x + i = 0$);
- d) $\begin{cases} rk(A) = 1 \\ rk(Q) = 1 \end{cases}$. Quindi $d = 0$. Allora $C \sim_{aff} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, che ha equazione $x^2 = 0$, detta **retta doppia** perché è l'unione di due rette coincidenti.

Osservazione: Ricordiamo che nel caso della conica non a centro si aveva $rk(A) = 1$, $rk(Q) = 3$.

Abbiamo così provato:

TEOREMA DI CLASSIFICAZIONE AFFINE DELLE CONICHE DI \mathbb{C}^2 : Ogni conica di \mathbb{C}^2 è affinementemente equivalente ad una e una sola delle seguenti:

- 1) $x^2 - y = 0$;
- 2) $x^2 + y^2 + 1 = 0$;
- 3) $x^2 + y^2 = 0$;
- 4) $x^2 + 1 = 0$;
- 5) $x^2 = 0$.

La coppia $(rk(A), rk(Q))$ è un sistema completo di invarianti per equivalenza affine in \mathbb{C}^2 .

Caso $\mathbb{K} = \mathbb{R}$

Su \mathbb{R} il rango non è un invariante completo per congruenza.

La segnatura lo è, ma non è invariante per \sim_{aff} (in quanto se moltiplico un'equazione per $\alpha < 0$ cambia la segnatura da $(i_+(Q), i_-(Q), i_0(Q))$ a $(i_-(Q), i_+(Q), i_0(Q))$).

Possiamo però usare l'indice di Witt che è insensibile alla moltiplicazione per scalare $\neq 0$ (infatti $\min(i_+(Q), i_-(Q))$ rimane lo stesso).

Passo 2 – Caso reale): Semplifichiamo A .

Infatti dopo il passo 1 ci siamo ridotti alla forma:

$$Q = \begin{pmatrix} \boxed{A} & 0 \\ 0 & d \end{pmatrix}, \text{ con } d = 0 \vee d = 1.$$

Distinguiamo i casi a seconda dei valori di $(rk(A), rk(Q), w(A), w(Q))$.

1) $\begin{cases} rk(A) = 2 \\ rk(Q) = 3 \end{cases}$, quindi C è non degenere, cioè $d = 1$.

Si possono avere i seguenti sottocasi:

$$\begin{cases} w(A) = 0 \Rightarrow w(Q) = < \begin{matrix} 0 \\ 1 \end{matrix} \\ w(A) = 1 \Rightarrow w(Q) = 1 \end{cases}$$

• $rk(A) = 2, rk(Q) = 3, w(A) = 0, w(Q) = 0$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ ossia } x^2 + y^2 + 1 = 0, \text{ tipo } C_2, \text{ detta "ellisse immaginaria"} (V(f) = \emptyset).$$

• $rk(A) = 2, rk(Q) = 3, w(A) = 0, w(Q) = 1$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \text{ ossia } x^2 + y^2 - 1 = 0, \text{ tipo } C_3, \text{ detta "ellisse reale"}$$

$(V(f) = \text{circonferenza})$.

• $rk(A) = 2, rk(Q) = 3, w(A) = 1, w(Q) = 1$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ ossia } x^2 - y^2 + 1 = 0, \text{ tipo } C_4, \text{ detta "iperbole"} (V(f) = \text{iperbole}).$$

2) $\begin{cases} rk(A) = 2 \\ rk(Q) = 2 \end{cases}$, quindi C è semplicemente degenere, cioè $d = 0$.

Si possono avere i seguenti sottocasi:

$$\begin{cases} w(A) = 0 \Rightarrow w(Q) = 1 \\ w(A) = 1 \Rightarrow w(Q) = 2 \end{cases}$$

• $rk(A) = 2, rk(Q) = 2, w(A) = 0, w(Q) = 1$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ ossia } x^2 + y^2 = 0, \text{ tipo } C_5, \text{ detta "rette complesse incidenti"}$$

$(V(f) = \{(0,0)\})$.

• $rk(A) = 2, rk(Q) = 2, w(A) = 1, w(Q) = 2$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ ossia } x^2 - y^2 = 0, \text{ tipo } C_6, \text{ detta "rette incidenti"}$$

$(V(f) = \text{rette incidenti})$

3) $\begin{cases} rk(A) = 1 \\ rk(Q) = 2 \end{cases}$, cioè $d = 1$.

Si possono avere i seguenti sottocasi:

$$\begin{cases} w(A) = 1 \Rightarrow w(Q) = < \begin{matrix} 1 \\ 2 \end{matrix} \end{cases}$$

• $rk(A) = 1, rk(Q) = 2, w(A) = 1, w(Q) = 1$:

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ ossia } x^2 + 1 = 0, \text{ tipo } \mathbf{C}_7, \text{ detta "rette complesse parallele"} (V(f) = \emptyset).$$

- $rk(A) = 1, rk(Q) = 2, w(A) = 1, w(Q) = 2:$

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \text{ ossia } x^2 - 1 = 0, \text{ tipo } \mathbf{C}_8, \text{ detta "rette parallele"}$$

$(V(f) = \text{rette parallele}).$

- 4) $\begin{cases} rk(A) = 1 \\ rk(Q) = 1 \end{cases}$, cioè $d = 0$.

Si possono avere i seguenti sottocasi:

$$\{w(A) = 1 \Rightarrow w(Q) = 2$$

- $rk(A) = 1, rk(Q) = 1, w(A) = 1, w(Q) = 2:$

$$Q \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ ossia } x^2 = 0, \text{ tipo } \mathbf{C}_9, \text{ detta "retta doppia"} (V(f) = \{x = 0\}).$$

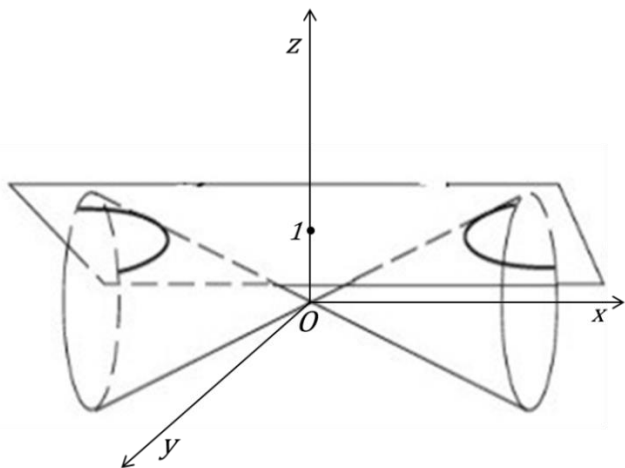
Abbiamo dunque provato:

TEOREMA DI CLASSIFICAZIONE AFFINE DELLE CONICHE DI \mathbb{R}^2 : Ogni conica di \mathbb{R}^2 è affinementemente equivalente a una e una sola delle seguenti:

- 1) $x^2 - y = 0;$
- 2) $x^2 + y^2 + 1 = 0;$
- 3) $x^2 + y^2 - 1 = 0;$
- 4) $x^2 - y^2 + 1 = 0;$
- 5) $x^2 + y^2 = 0;$
- 6) $x^2 - y^2 = 0;$
- 7) $x^2 + 1 = 0;$
- 8) $x^2 - 1 = 0;$
- 9) $x^2 = 0.$

Inoltre la quaterna $(rk(A), rk(Q), w(A), w(Q))$ è un sistema completo di invarianti per equivalenza affine in \mathbb{R}^2 .

Osservazione: La conica C di equazione $x^2 - y^2 + 1 = 0$ (l'iperbole) può essere vista come: $C = S \cap \{z = 1\}$, dove $S = \{(x, y, z) \in \mathbb{R}^3 | x^2 - y^2 + z^2 = 0\}$.



Esempio: Sia C la conica di \mathbb{R}^2 di equazione $x^2 + y^2 + 4xy - 2x - 10y + 1 = 0$.

Determinare il modello canonico affine \tilde{C} di C e determinare un'affinità

$\psi \in \text{Aff}(\mathbb{R}^2) \mid \psi(\tilde{C}) = C$.

Svolgimento:

$$Q = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -5 \\ -1 & -5 & 1 \end{pmatrix}, A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} -1 \\ -5 \end{pmatrix}.$$

$$\det(A) = -3 \Rightarrow \sigma(A) = (1, 1, 0) \Rightarrow w(A) = 1.$$

$$\det(Q) = -3 + 5(-3) - 1(-9) = -9 \Rightarrow \sigma(Q) = (2, 1, 0) \Rightarrow w(Q) = 1.$$

Dunque $\tilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $x^2 - y^2 + 1 = 0$, cioè è un'iperbole.

Sia $f(x, y) = x^2 + y^2 + 4xy - 2x - 10y + 1$.

Sia $\tilde{f}(x, y) = x^2 - y^2 + 1$.

1) Ricerchiamo il centro di C .

Dobbiamo risolvere il sistema $AX = -B$, cioè:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \Rightarrow \begin{cases} \alpha = 3 \\ \beta = -1 \end{cases} \Rightarrow N = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Sia $\psi_1(X) = X + N = X + \begin{pmatrix} 3 \\ -1 \end{pmatrix}$.

Allora $\tilde{M}_1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ rappresenta l'affinità ψ_1 .

La conica $\psi_1^{-1}(C)$ ha come matrice associata $Q_1 = {}^t\tilde{M}_1 Q \tilde{M}_1$.

Per quanto visto nella teoria:

$$Q_1 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

2) Trasformo A nella forma di Sylvester.

$$v_1 = e_1, v_2 = e_2 - \frac{2}{1}e_1 = e_2 - 2e_1.$$

$$\langle v_2, v_2 \rangle = \langle e_2 - 2e_1, e_2 - 2e_1 \rangle = 1 + 4 - 8 = -3. \text{ Dunque:}$$

$$\mathfrak{M}_{\{v_1, v_2\}, \{e_1, e_2\}}(id) = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Sia $\psi_2(X) = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} X$, che è rappresentata da $\tilde{M}_2 = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

La matrice relativa alla conica $C_2 = \psi_2^{-1}(C_1)$ è:

$$Q_2 = {}^t\tilde{M}_2 Q_1 \tilde{M}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Dobbiamo trasformare il coefficiente di x^2 da 1 a 3, poiché in questo modo otterremo l'ipersuperficie cercata (poiché nella classe di equivalenza ci sta il polinomio di \tilde{C} moltiplicato per una qualsiasi $\alpha \neq 0$, in questo caso $\alpha = 3$):

$$\tilde{M}_3 = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'affinità cercata è $\tilde{M}_1 \cdot \tilde{M}_2 \cdot \tilde{M}_3 = \begin{pmatrix} \sqrt{3} & -2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, cioè:

$$\psi(X) = \begin{pmatrix} \sqrt{3} & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Osservazione: Si possono classificare metricamente le coniche di \mathbb{R}^2 in modo simile al caso affine. Si diagonalizza A non con Sylvester ma con il teorema spettrale.

Se $M \in O(2)$, $\det(\widetilde{M}_N) = \det \begin{pmatrix} \boxed{M} & N \\ 0 & 1 \end{pmatrix} = \det(M) = \pm 1$, quindi:

- $tr({}^tMAM) = tr(A)$;
- $\det({}^tMAM) = \det(A)$;
- $\det({}^t\widetilde{M}_N Q \widetilde{M}_N) = \det(Q)$ (anche se $\widetilde{M}_N \notin O(3)$).

Se moltiplico l'equazione della conica per $\alpha \neq 0$, la matrice della conica diventa:

$$\alpha Q = \begin{pmatrix} \boxed{\alpha A} & \alpha B \\ \alpha {}^t B & \alpha c \end{pmatrix}, \text{ dunque:}$$

- $\det(\alpha Q) = \alpha^3 \det(Q)$;
- $\det(\alpha A) = \alpha^2 \det(A)$;
- $tr(\alpha A) = \alpha tr(A)$.

Quindi (se $tr(A) \neq 0$) i numeri $\frac{\det(A)}{tr(A)^2}$ e $\frac{\det(Q)}{tr(A)^3}$ sono invarianti metrici.

Se C e C' sono metricamente equivalenti, allora $\exists \alpha \neq 0$

- $\det(Q') = \alpha^3 \det(Q)$;
- $\det(A') = \alpha^2 \det(A)$;
- $tr(A') = \alpha tr(A)$.

Non vale però il viceversa, cioè questi invarianti non sono sufficienti per decidere se due coniche sono metricamente equivalenti. Infatti:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{pmatrix}$$

non sono metricamente equivalenti $\forall c \neq 1$, ma hanno gli stessi invarianti precedenti.

DEFINIZIONE 5.5.4: Sia $f(x_1, \dots, x_n) \in \mathbb{K}[x_1, \dots, x_n]$ un polinomio di grado 2.

Allora $f(X) = {}^tXAX + 2{}^tBX + c$, con $A \in \mathcal{S}(n, \mathbb{K})$.

Definiamo **matrice della quadrica** $[f]$:

$$Q = \begin{pmatrix} \boxed{A} & B \\ {}^t B & c \end{pmatrix} \in \mathcal{M}(n+1, \mathbb{K}).$$

La quadrica C si dice **degenere** se $\det(Q) = 0$.

DEFINIZIONE 5.5.5: Una quadrica C si dice **cono di vertice** $P_0 \in C$ se $\forall P \in C, P \neq P_0$, tutta la retta congiungente P_0 e P è contenuta in C .

DEFINIZIONE 5.5.6: Una quadrica C è detta **cilindro** se $\exists r$ retta di \mathbb{K}^n tale che $\forall P \in C$ la retta passante per P e parallela a r è contenuta in C .

- Esempi: 1) Se f è un polinomio omogeneo di secondo grado, $[f]$ è un cono di vertice l'origine.
 2) Se f è un polinomio di secondo grado in x_1, \dots, x_n in cui non compare una variabile x_j , allora $[f]$ è un cilindro parallelo all'asse x_j .
 3) $x^2 - y^2 = 0$ in \mathbb{R}^3 (che è l'unione di due piani incidenti), è sia un cono di vertice $(0,0,0)$, sia un cilindro parallelo all'asse z .

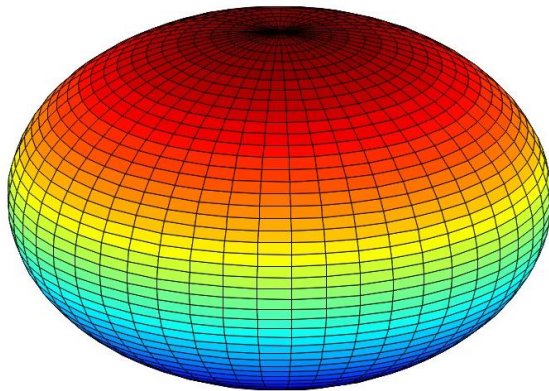
In modo simile al caso delle coniche si prova:

TEOREMA DI CLASSIFICAZIONE AFFINE DELLE QUADRICHE: Ogni quadrica di \mathbb{K}^n è affinementemente equivalente ad una e una sola delle seguenti:

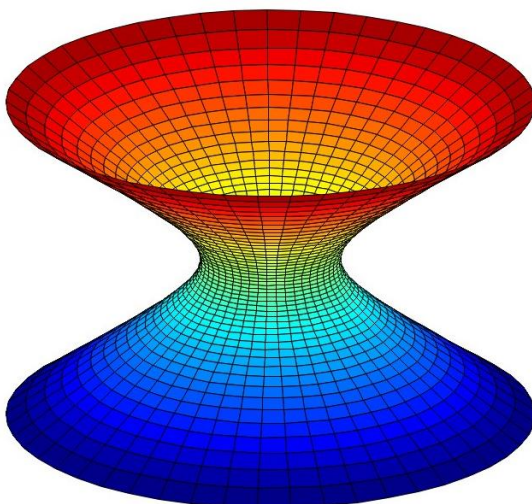
- 1) Se $\mathbb{K} = \mathbb{C}$:
 - $x_1^2 + \dots + x_r^2 + d = 0$, con $d = 0 \vee d = 1$ (conica a centro);
 - $x_1^2 + \dots + x_r^2 - x_n = 0$, detto “**paraboloide**”.
- 2) Se $\mathbb{K} = \mathbb{R}$:
 - $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2 + d = 0$, con $d = 0 \vee d = 1$ (conica a centro);
 - $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2 - x_n = 0$, detti “**paraboloidi**”.

LISTA DEI MODELLI AFFINI PER LE QUADRICHE DI \mathbb{R}^3 :

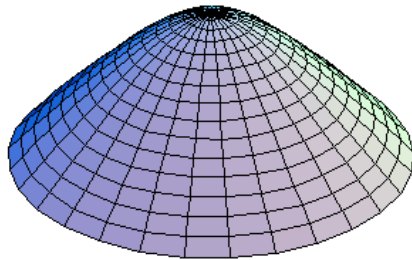
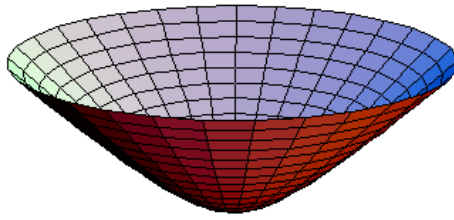
- a) A centro:
 - A centro non degeneri (cioè $c \neq 0$):
 - 1) $x^2 + y^2 + z^2 + 1 = 0$, detto “**ellissoide immaginario**” ($V(f) = \emptyset$);
 - 2) $x^2 + y^2 + z^2 - 1 = 0$, detto “**ellissoide**”;



- 3) $x^2 + y^2 - z^2 - 1 = 0$, detto “**iperboloide a una falda**”;



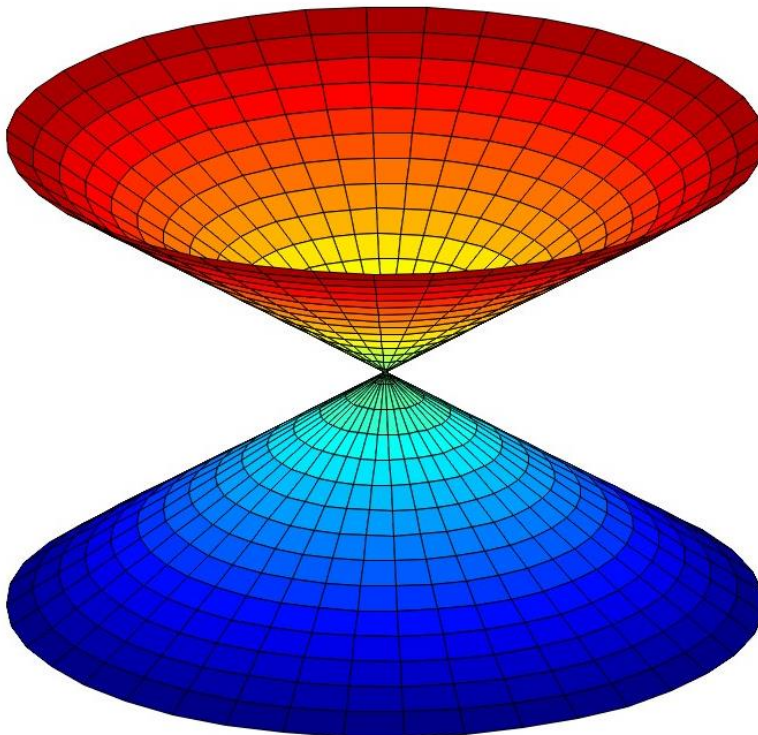
4) $x^2 + y^2 - z^2 + 1 = 0$, detto “iperboloide a due falde”



- A centro degeneri con $c = 0$ (dunque coni):

5) $x^2 + y^2 + z^2 = 0$, detto “punto”, o “cono immaginario” ($V(f) = \{(0,0,0)\}$);

6) $x^2 + y^2 - z^2 = 0$, detto “cono reale”;



7) $x^2 + y^2 = 0$, detto “piani complessi incidenti” ($V(f) = \text{asse } z$);

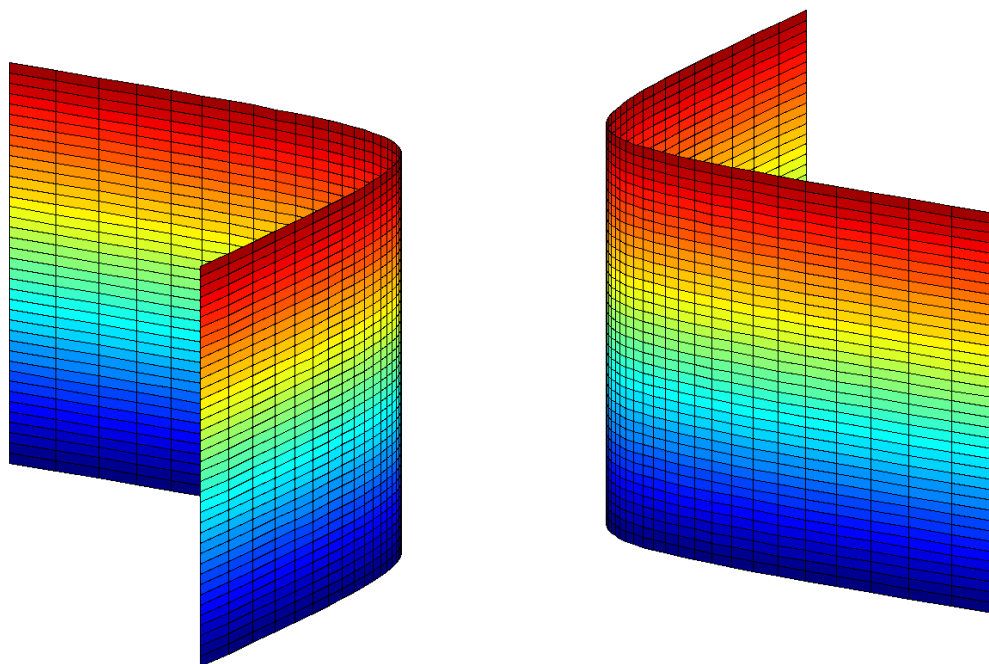
8) $x^2 - y^2 = 0$, detto “piani incidenti”;

9) $x^2 = 0$, detto “piano doppio”;

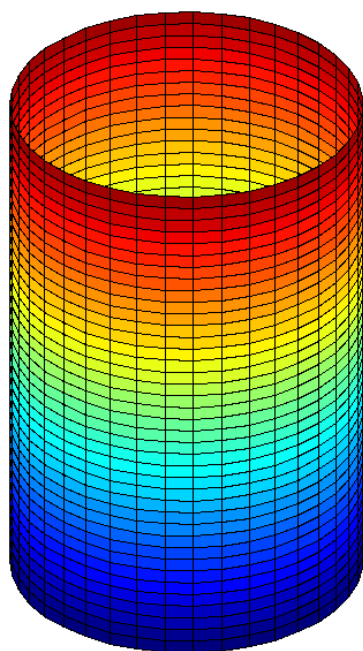
- A centro degeneri con $c \neq 0$:

10) $x^2 + y^2 + 1 = 0$, detto “cilindro immaginario” ($V(f) = \emptyset$);

11) $x^2 - y^2 + 1 = 0$, detto “cilindro iperbolico”;



12) $x^2 + y^2 - 1 = 0$, detto “cilindro ellittico”;



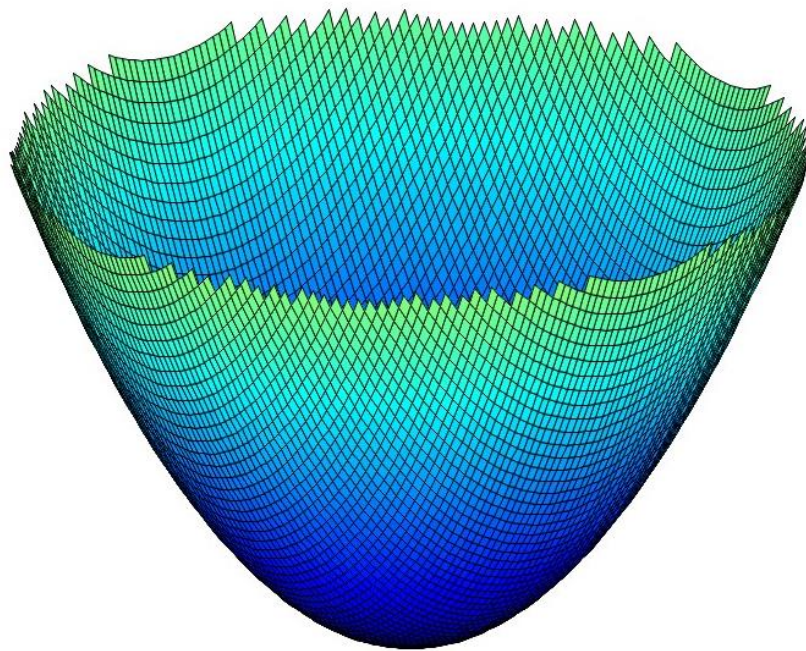
13) $x^2 - 1 = 0$, detto “piani paralleli”;

14) $x^2 + 1 = 0$, detto “piani complessi paralleli” ($V(f) = \emptyset$);

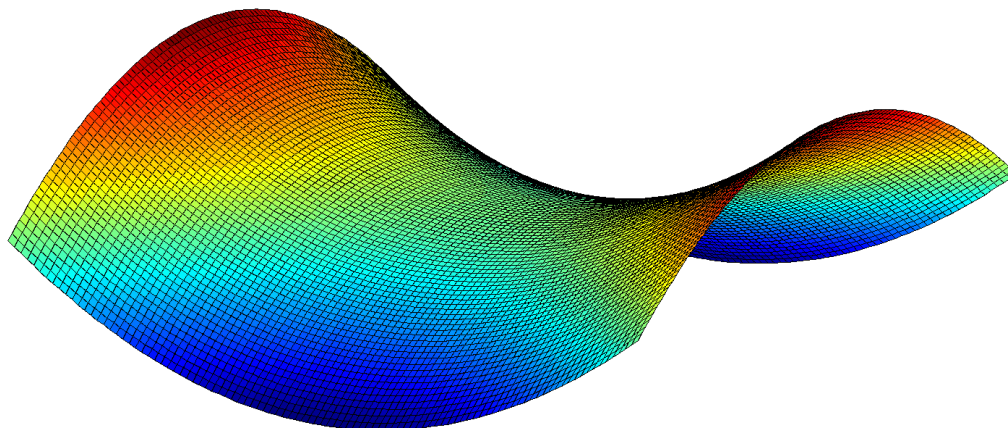
b) Non a centro (paraboloidi):

- Non a centro non degeneri:

15) $x^2 + y^2 - z = 0$, detto “paraboloide ellittico”;

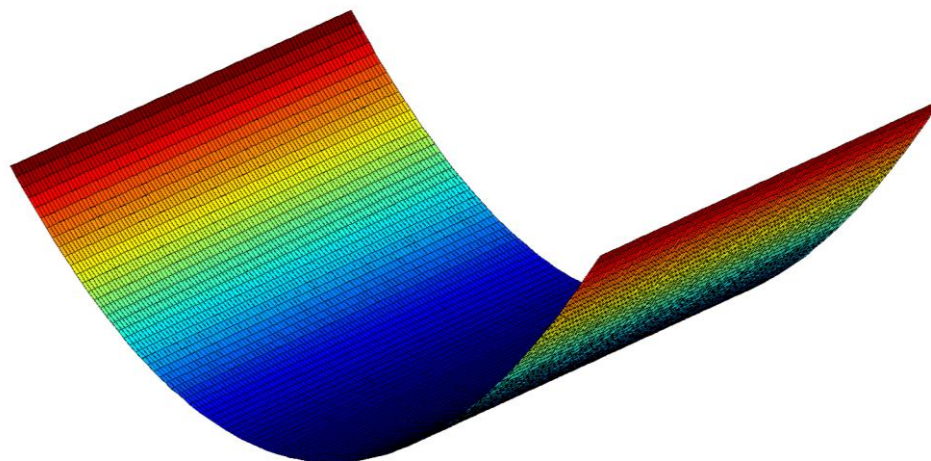


16) $x^2 - y^2 - z = 0$, detto “paraboloide iperbolico” o “sella”;



• Non a centro degeneri:

17) $x^2 - z = 0$, detto “cilindro parabolico”;



Osservazione: L'iperboloide a una falda è spesso detto "rigato" perché per ogni suo punto passano due rette incidenti completamente giacenti sull'iperboloide. Infatti:

$$x^2 + y^2 - z^2 - 1 = 0 \Rightarrow x^2 - z^2 = 1 - y^2 \Rightarrow (x + z)(x - z) = (1 + y)(1 - y).$$

Dunque ci sono due tipi di queste rette:

$$1^\circ \text{ tipo: } \begin{cases} \lambda(x - z) = \mu(1 - y) \\ \mu(x + z) = \lambda(1 + y) \end{cases}$$

$$2^\circ \text{ tipo: } \begin{cases} \lambda(x - z) = \mu(1 + y) \\ \mu(x + z) = \lambda(1 - y) \end{cases}$$

PROPOSIZIONE 5.5.1: Sia Q una quadrica di \mathbb{R}^n e R un centro di Q . Allora, se $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è un'affinità, $f(R)$ è un centro di $f(Q)$ (cioè le affinità mantengono i centri).

Dimostrazione:

Utilizzando la notazione usuale:

$$\begin{aligned} R \text{ è centro per } Q &\Leftrightarrow {}^tMAM \cdot R = -{}^tM(AN + B) \quad \boxed{\Leftrightarrow} \quad AMR = -AN - B \Leftrightarrow \\ &\Leftrightarrow AMR + AN = -B \Leftrightarrow A(MR + N) = -B \Leftrightarrow A(f(R)) = -B \Leftrightarrow f(R) \text{ è centro per } f(Q), \end{aligned}$$

dove il passaggio $\boxed{\Leftrightarrow}$ segue dal fatto che M è invertibile, poiché f è affinità.

SITOGRAFIA:

- <http://upload.wikimedia.org/wikipedia/commons/9/91/HyperboloidOfTwoSheets.png>