Note del corso di Geometria I 2021/22

Sandro Manfredini

Relazioni di Equivalenza

Siano A, B due insiemi.

Definizione:
Il *prodotto cartesiano* di A e B è l'insieme $A \times B$ formato dalle coppie (ordinate) in cui la prima entrata è un elemento di A e la seconda è un elemento di B:

$$A \times B = \{(a, b) |\ a \in A,\ b \in B\}$$

(osserviamo che se $A = \emptyset$ o $B = \emptyset$, allora $A \times B = \emptyset$).
In $A \times A$, la *diagonale* Δ_A è data dalle coppie in cui la prima entrata è uguale alla seconda:

$$\Delta_A = \{(a, b) \in A \times A |\ a = b\} = \{(a, a) |\ a \in A\}.$$

Osservazioni:
▶ L'applicazione che scambia le entrate $\tau : A \times B \to B \times A$, $(a, b) \mapsto (b, a)$, è biunivoca.
▶ La diagonale Δ_A è l'insieme dei punti fissi di $\tau : A \times A \to A \times A$, cioè, dato $(a, b) \in A \times A$, $(a, b) \in \Delta_A \iff \tau((a, b)) = (a, b)$.

Definizione:
Dato un insieme non vuoto X, una *relazione* su X è un sottoinsieme non vuoto $R \subset X \times X$.
Se la coppia $(x, y) \in R$ diciamo che x è in R-relazione con y e scriviamo $x \sim_R y$ (ovvero diciamo semplicemente che x è in relazione con y e scriviamo $x \sim y$, sottintendendo R).

Definizione:
Una relazione R su X si dice:

- *Riflessiva* se $\forall x \in X$, $x \sim_R x$ (ovvero $(x, x) \in R$);
- *Simmetrica* se $\forall x, y \in X$ tali che $x \sim_R y$, allora $y \sim_R x$ (ovvero $(x, y) \in R \Rightarrow (y, x) \in R$);
- *Transitiva* se $\forall x, y, z \in X$ tali che $x \sim_R y$ e $y \sim_R z$, allora $x \sim_R z$ (ovvero $(x, y), (y, z) \in R \Rightarrow (x, z) \in R$).

Osservazioni:
▶ R è riflessiva se e solo se $\Delta_X \subset R$;
▶ R è simmetrica se e solo se R è τ-*invariante*, cioè, $\tau(R) \subset R$.

Esempio:
■ Sia Y un insieme e sia $X = \wp(Y)$ l'*insieme delle parti* (cioè dei sottoinsiemi) di Y. Per $A, B \in X$ poniamo $A \sim B$ se $A \subset B$. Questa è una relazione su X riflessiva e transitiva, ma $A \subset B$ e $B \subset A$ se e solo se $A = B$, quindi non è simmetrica (eccetto il caso $Y = \emptyset$).
Una relazione riflessiva, transitiva, tale che $x \sim y$ e $y \sim x$ se e solo se $x = y$ è detta *relazione d'ordine* o *ordinamento*.
Altro esempio di relazione d'ordine è $\leq$ su $\mathbb{Z}$.

Definizione:
Una relazione R è una *relazione di equivalenza* se è riflessiva, simmetrica e transitiva.
In questo caso, due elementi in R-relazione si dicono R-*equivalenti* (o semplicemente equivalenti, sottintendendo R).

Osserviamo che ogni insieme non vuoto X ammette sempre due relazioni di equivalenza "estreme":

- $R = \Delta_X$, per cui $x \sim_R y$ se e solo se $x = y$, cioè R è la relazione di *uguaglianza*; R è la relazione di equivalenza minima, rispetto all'ordinamento su $\wp(X \times X)$ dato dall'inclusione $\subset$.
- $R = X \times X$, per cui gli elementi di X sono tutti equivalenti tra di loro. R è la relazione di equivalenza massima.

Esempio1:
■ La relazione di equivalenza data da una partizione.

Una *partizione* $\mathcal{P}$ di un insieme non vuoto X è un insieme di sottoinsiemi di X (cioè $\mathcal{P} \subset \wp(X)$) che soddisfa le seguenti proprietà.

1. Ogni $U \in \mathcal{P}$ è non vuoto;
2. $\mathcal{P}$ *ricopre* X, cioè, $\forall x \in X$, $\exists\, U \in \mathcal{P}$ tale che $x \in U$; ovvero, X è l'unione dei sottoinsiemi che appartengono a $\mathcal{P}$, $X = \bigcup_{U \in \mathcal{P}} U$.
3. Se $U, U' \in \mathcal{P}$ e $U \neq U'$, allora $U \cap U' = \emptyset$; ovvero, se $U \cap U' \neq \emptyset$, allora $U = U'$;

Notiamo che da 3 segue che l'U in 2 è unico: $\forall x \in X$, $\exists!\, U \in \mathcal{P}$ tale che $x \in U$.

Sia $\mathcal{P}$ una partizione di X.
La relazione su X data da $x \sim_{\mathcal{P}} y$ se esiste $U \in \mathcal{P}$ tale che $x, y \in U$ è una relazione di equivalenza.
Infatti, la relazione è evidentemente simmetrica; per ogni $x \in X$, $x \sim_{\mathcal{P}} x$ perché $\mathcal{P}$ ricopre X; se poi $x, y, z \in X$ sono tali che $x \sim_{\mathcal{P}} y$ e $y \sim_{\mathcal{P}} z$, allora esistono $U, U' \in \mathcal{P}$ tali che $x, y \in U$ e $y, z \in U'$, e quindi $y \in U \cap U'$. Ne segue che $U = U'$, e $x, z \in U$, cioè $x \sim_{\mathcal{P}} z$.

Esempio2:
■ La relazione di equivalenza data da una applicazione.

Sia $f : X \to Y$ una applicazione.
La relazione su X data da $x \sim_f y$ se e solo se $f(x) = f(y)$ è una relazione di equivalenza.
Infatti, l'uguaglianza (in Y) è una relazione di equivalenza.

Vediamo che tutte le relazioni di equivalenza su X provengono da una partizione e da un'applicazione.

Definizione:
Data una relazione di equivalenza R su X e dato $x \in X$, il sottoinsieme di X dato dagli elementi equivalenti a x di dice la *classe di R-equivalenza* di x,

$$[x]_R = \{y \in X |\ y \sim_R x\} \subset X$$

(sottintendendo R, scriviamo semplicemente $[x]$ e diciamo la classe di equivalenza di x.)

Osservazioni:
▶ $x \in [x]_R$ (per la riflessività);
▶ $x_1 \sim_R x_2$ se e solo se $[x_1]_R = [x_2]_R$ (per la transitività, $[x_1]_R \subset [x_2]_R$, ma vale anche l'inclusione opposta poiché $x_2 \sim_R x_1$ per la simmetricità).
▶ Nell'Esempio1, le classi di equivalenza sono gli elementi di $\mathcal{P}$. In particolare, dato $x \in X$, $[x]_{\mathcal{P}}$ è l'unico $U \in \mathcal{P}$ tale che $x \in U$.
▶ Nell'Esempio2, le classi di equivalenza sono le controimmagini (non vuote) degli elementi di Y tramite f. In generale, se $B \subset Y$, la *controimmagine di B tramite f* è il sottoisieme di X dato da $f^{-1}(B) = \{x \in X |\ f(x) \in B\}$. Le classi di equivalenza sono quindi $f^{-1}(\{y\}) = \{x \in X |\ f(x) = y\}$, al variare di $y \in \operatorname{Im} f$. In particolare, dato $x \in X$, $[x]_f = f^{-1}(\{f(x)\})$.

Proposizione
Sia R una relazione di equivalenza su X. Allora $\mathcal{P} = \{[x]_R |\ x \in X\}$ è una partizione di X e R è la relazione di equivalenza data da $\mathcal{P}$.

Dimostrazione
Poiché per ogni $x \in X$ $x \in [x]_R$, ogni classe di equivalenza è non vuota e le classi di equivalenza ricoprono X.
Dati $x, y \in X$ tali che $[x]_R \cap [y]_R \neq \emptyset$, segliamo z in tale intersezione. Se $u \sim x$, allora $u \sim z \sim y$ (per le proprietà simmetrica e transitiva), da cui $[x]_R \subset [y]_R$. Scambiando i ruoli di x e y otteniamo anche l'inclusione opposta, quindi $[x]_R = [y]_R$. ❏

Definizione:
L'insieme $\mathcal{P}$ delle classi di equivalenza si dice *insieme quoziente* di X rispetto a R e si indica con

$$X/_{\sim_R} = \{[x]_R |\ x \in X\}.$$

Osservazioni:
▶ Le classi di equivalenza hanno un doppio ruolo: sono sottoinsiemi di X ed elementi dell'insieme quoziente: $[x]_R \subset X$ e $[x]_R \in X/_{\sim_R}$.
▶ Nell'Esempio2, l'applicazione $\hat{f} : X/_{\sim_f} \to \operatorname{Im} f$, $[x]_f \mapsto f(x)$, è *ben definita* ed è biunivoca.
Enfatizziamo il fatto che ci sia bisogno di mostrare che $\hat{f}$ è ben definita.
Infatti, data una classe di equivalenza $c \in X/_{\sim_f}$, per definire $\hat{f}(c)$, scegliamo un *rappresentante* di c, cioè un $x \in c$, e definiamo $\hat{f}(c) = f(x)$. Perché questo abbia senso, la definizione di $\hat{f}(c)$ non deve dipendere dalla scelta del rappresentante: potrebbe a priori succedere che scegliendo un altro rappresentante y, $f(y) \neq f(x)$. Ma, se scegliamo un altro rappresentante $y \in c$, $x \sim_f y$, ovvero $f(x) = f(y)$.

Definizione:
L'applicazione $\pi : X \to X/_{\sim_R}$, $x \mapsto [x]_R$, che ad ogni elemento associa la sua classe di equivalenza, è detta *proiezione al quoziente*.

Osservazioni:
▶ π è surgettiva per definizione.
▶ Per ogni $x \in X$, $\pi([x]_R) = [x]_R$ (notare di nuovo il doppio ruolo delle classi di equivalenza).

Proposizione
Sia R una relazione di equivalenza su X e π la proiezione al quoziente.
R è la relazione di equivalenza data da π.

Dimostrazione
Dati $x, y \in X$, $x \sim_R y$ se e solo se $[x]_R = [y]_R$ se e solo se $\pi(x) = \pi(y)$. ❑

Esempio3:
■ La relazione di eqivalenza data un gruppo di trasformazioni.

Sia X un insieme non vuoto e poniamo $\mathrm{S}(X) = \{\sigma : X \to X |\ \sigma \text{ è biunivoca}\}$.
Ricordiamo che l'applicazione identità $id_X : X \to X$, $x \mapsto x$, è biunivoca, che la composizione di applicazioni biunivoche è biunivoca e che ogni applicazione biunivoca ammette una inversa insiemistica che è biunivoca.

Un sottoinsieme $G \subset \mathrm{S}(X)$ tale che

- $id_X \in G$;
- se $\sigma \in G$ allora l'inversa insiemistica di σ, $\sigma^{-1} \in G$,
- se $\sigma, \tau \in G$ allora la composizione $\sigma \circ \tau \in G$;

è detto *gruppo di trasformazioni* di X.

Dato G un gruppo di trasformazioni di X, la relazione su X data da $x \sim_G y$ se esiste $\sigma \in G$ tale che $y = \sigma(x)$ è una relazione di equivalenza.
Infatti, per ogni $x \in X$ $x = id_X(x)$; se $y = \sigma(x)$ con $\sigma \in G$, allora $x = \sigma^{-1}(y)$; se $y = \sigma(x)$, e $z = \tau(y)$ con $\sigma, \tau \in G$, allora $z = (\tau \circ \sigma)(x)$.
Per $x \in X$, la classe di equivalenza $[x]_G = Gx = \{\sigma(x) |\ \sigma \in G\}$ è anche detta la *G-orbita* di x, mentre $G_x = \{g \in G |\ g(x) = x\}$ è detto *stabilizzatore* di x (ed è un sottogruppo di G).
Osserviamo che fissato $\bar{g} \in G$, la composizione a sinistra per $\bar{g}$ dà l'applicazione $s_{\bar{g}} : G \to G$, $g \mapsto \bar{g} \circ g$, e $s_{\bar{g}} \in \mathrm{S}(G)$. Infatti, è surgettiva poiché data $h \in G$, $h = s_{\bar{g}}(\bar{g}^{-1} \circ h)$; è iniettiva poiché se $g_1, g_2 \in G$ sono tali che $s_{\bar{g}}(g_1) = s_{\bar{g}}(g_2)$, allora $\bar{g} \circ g_1 = \bar{g} \circ g_2$, da cui componendo a sinistra per $\bar{g}^{-1}$, $g_1 = g_2$. Ovvero, l'inversa insiemistica di $s_{\bar{g}}$ è $s_{\bar{g}^{-1}}$: per ogni $g \in G$,

$$(s_{\bar{g}} \circ s_{\bar{g}^{-1}})(g) = \bar{g} \circ (\bar{g}^{-1} \circ g) = (\bar{g} \circ \bar{g}^{-1})g = g,$$

$$(s_{\bar{g}^{-1}} \circ s_{\bar{g}})(g) = \bar{g}^{-1} \circ (\bar{g} \circ g) = (\bar{g}^{-1} \circ \bar{g})g = g.$$

Fissiamo $y \in Gx$ un elemento della G-orbita di x, e consideriamo il sottoinsieme (non vuoto, per definizione di G-orbita) di G, $C_y = \{g \in G |\ g(x) = y\}$. Fissato $\bar{g} \in C_y$, $s_{\bar{g}}(G_x) = C_y$. Infatti, se $h \in G_x$, $(\bar{g} \circ h)(x) = \bar{g}(h(x)) = \bar{g}(x) = y$, per cui $\bar{g} \circ h \in C_y$; viceversa, data $h \in C_y$, $(s_{\bar{g}})^{-1}(h) = \bar{g}^{-1} \circ h \in G_x$ poiché $\bar{g}^{-1}(h(x)) = \bar{g}^{-1}(y) = x$. Quindi la restrizione di $s_{\bar{g}}$ a G_x dà una bigezione $s_{\bar{g}}|_{G_x} : G_x \to C_y$.
Osserviamo infine che al variare di $y \in Gx$, i C_y danno una partizione di G (e quindi una relazione di equivalenza su G): abbiamo già osservato che i C_y non sono vuoti e ricoprono G poiché dato $g \in G$, $g \in C_{g(x)}$; se poi $C_y \cap C_z \neq \emptyset$, scelto g in tale intersezione, $y = g(x) = z$.
Se quindi G è finito l'orbita di x contiene $\frac{|G|}{|G_x|}$ elementi.

Strutture Algebriche

Definizione:
Dato un insiseme non vuoto X, una *operazione* su X è un'applicazione $* : X \times X \to X$.
Il risultato dell'operazione tra $x \in X$ e $y \in X$, cioè l'immagine tramite $*$ di $(x, y) \in X \times X$, si indica con $x * y$.

Esempio:
■ Se X è un insieme non vuoto, $\mathrm{S}(X)$ è munito dell'operazione data dalla *composizione*

$$\circ : \mathrm{S}(X) \times \mathrm{S}(X) \to \mathrm{S}(X),\ (\sigma, \tau) \mapsto \sigma \circ \tau,$$

dove $\sigma \circ \tau : X \to X$ è definita da $(\sigma \circ \tau)(x) = \sigma(\tau(x))$.

Osserviamo che se $\rho, \sigma, \tau \in \mathrm{S}(X)$, allora $(\rho \circ \sigma) \circ \tau = \rho \circ (\sigma \circ \tau)$.
Infatti, per ogni $x \in X$, $((\rho \circ \sigma) \circ \tau)(x) = (\rho \circ \sigma)(\tau(x)) = \rho(\sigma(\tau(x))) = \rho((\sigma \circ \tau)(x)) = (\rho \circ (\sigma \circ \tau))(x)$.
Inoltre, l'applicazione identità id_X appartiene a $\mathrm{S}(X)$ ed ha la proprietà $id_X \circ \sigma = \sigma \circ id_X = \sigma$ per ogni $\sigma \in \mathrm{S}(X)$.
Infine, se $\sigma \in \mathrm{S}(X)$, l'inversa insiemistica di σ, σ^{-1}, appartiene a $\mathrm{S}(X)$ ed ha la proprietà $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = id_X$.

Formalizziamo questa struttura nella seguente definizione.

Definizione:
Un insieme munito di una operazione $(X, *)$ si dice *gruppo* se:

- l'operazione $*$ è *associativa*, cioè $\forall x, y, z \in X$, $(x * y) * z = x * (y * z)$;
- l'operazione $*$ ammette *elemento neutro*, cioè $\exists\ u \in X$ tale che, $\forall x \in X$, $u * x = x * u = x$;
- ogni elemento di X ammette un *inverso*, cioè $\forall x \in X$, $\exists\ y \in X$ tale che $x * y = y * x = u$.

Se l'operazione $*$ è anche *commutativa*, cioè se $\forall x, y \in X$, $x * y = y * x$, allora X si dice gruppo *commutativo* o *abeliano*.

Ad esempio, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ e $(\mathbb{R}, +)$ sono gruppi abeliani, $(\mathrm{S}(X), \circ)$ è un gruppo (non abeliano se X contiene almeno tre elementi) detto *gruppo simmetrico su* X (equivalentemente gruppo delle *trasformazioni* di X). Un gruppo di trasformazioni di X è un *sottogruppo* di $\mathrm{S}(X)$.

Osservazioni:
▶ In un gruppo, l'elemento neutro è unico.
Infatti, se $u, u' \in X$ sono elementi neutri, $u' = u' * u = u$.
▶ In un gruppo, l'inverso di un elemento è unico.
Infatti, dato $x \in X$, se $y, y' \in X$ sono inversi di x, $y' = y' * u = y' * (x * y) =$

$(y' * x) * y = u * y = y$.

Di solito, si usa la notazione moltiplicativa e si scrive xy al posto di $x * y$, l'elemento neutro si indica con 1, l'inverso di x si indica con x^{-1}.
La notazione additiva è riservata alle operazioni commutative e si scrive $x + y$ al posto di $x * y$, l'elemento neutro si indica con 0, l'inverso di x si indica con $-x$ e si dice *opposto* di x. Con un piccolo abuso di notazione, si scrive $x - y$ al posto di $x + (-y)$.
Quindi:

- $1x = x1 = x$, $x^{-1}x = xx^{-1} = 1$;
- $0 + x = x + 0 = x$, $x + (-x) = (-x) + x = 0$ (quest'ultima si scrive $x - x = -x + x = 0$).

Notiamo che $\mathbb{Z}$, oltre all'operazione di somma che lo rende un gruppo abeliano, ha anche l'operazione prodotto, che è associativa, ammette elemento neutro ed è distributiva sulla somma.
Formalizziamo questa struttura nella seguente definizione.

Definizione:
Un insieme munito di due operazioni $(\mathcal{A}, +, \cdot)$ è un *anello* se:

- $(\mathcal{A}, +)$ è un gruppo abeliano;
- l'operazione $\cdot$ è associativa ed ammette elemento neutro $1 \neq 0$;
- l'operazione $\cdot$ è *distributiva* su $+$, cioè $\forall x, y, z \in \mathcal{A}$,

$$x(y + z) = xy + xz \quad \text{e} \quad (x + y)z = xz + yz.$$

$\mathcal{A}$ si dice *anello commutativo* se il prodotto è commutativo.

Ad esempio, $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ e $(\mathbb{R}, +, \cdot)$ sono anelli commutativi. Un esempio di anello non commutativo è dato dalle matrici 2×2 a coefficienti in un campo.

Osservazioni:
▶ Per ogni $x \in \mathcal{A}$, $0x = x0 = 0$.
Infatti, $0x = (0+0)x = 0x + 0x$ (dove abbiamo usato la definizione di elemento neutro e la distributività); sommando ad entrambi i membri l'opposto di $0x$, $a = -0x$, otteniamo $0x + a = (0x + 0x) + a = 0x + (0x + a)$ (dove abbiamo usato l'associatività), da cui $0 = 0x + 0 = 0x$ (dove abbiamo usato il fatto che a è l'opposto di $0x$ e la definizione di elemento neutro). In modo analogo si dimostra $x0 = 0$.
▶ Per ogni $x \in \mathcal{A}$, $(-1)x = x(-1) = -x$
(moltiplicare x per l'opposto di 1 dà l'opposto di x).
Infatti, da $0 = 1 + (-1)$, moltiplicando a destra per x otteniamo $0 = 0x = (1 + (-1))x = 1x + (-1)x = x + (-1)x$. Analogamente, moltiplicando a sinistra per x, otteniamo $x + x(-1) = 0$.

▶ Dati a, b in un anello commutativo $\mathcal{A}$, diciamo che a *divide* b o che b è un *multiplo* di a, in simboli $a|b$, se esisite $c \in \mathcal{A}$ tale che $b = ca$. Abbiamo visto che ogni elemento di $\mathcal{A}$ divide 0.

Definizione:
Dato un anello $\mathcal{A}$, poniamo

- $\mathcal{A}^* = \mathcal{A} \setminus \{0\}$;
- $\mathcal{A}' = \{a \in \mathcal{A} | \ \exists\ b \in \mathcal{A} : \ ab = ba = 1\}$ il sottoinsieme degli elementi *invertibili* rispetto al prodotto.

Ad esempio, $\mathbb{Z}' = \{1, -1\}$, $\mathbb{Q}' = \mathbb{Q}^*$, $\mathbb{R}' = \mathbb{R}^*$.

Osservazioni:
▶ Anche nel caso di un anello, se l'inverso (rispetto al prodotto) di un elemento esiste, esso è unico. L'inverso di $a \in \mathcal{A}$ si indica con a^{-1}.
▶ Poiché il prodotto di due elementi invertibili è invetibile ($(ab)^{-1} = b^{-1}a^{-1}$), e l'inverso di un elemento invertibile è invertibile ($(a^{-1})^{-1} = a$), $\mathcal{A}'$ dotato della restrizione del prodotto è un gruppo.
▶ $\mathcal{A}' \subset \mathcal{A}^*$, in quanto 0 non può essere invertibile. Infatti, per ogni $x \in \mathcal{A}$, $0x = 0$ e ricordando che $0 \neq 1$, segue che 0 non ha un inverso.
▶ Se $\mathcal{A}$ è commutativo, ogni $x \in \mathcal{A}$ invertibile divide tutti gli elementi di A, ovvero $\forall x \in \mathcal{A}', \forall a \in \mathcal{A}$, $x|a$. Infatti, $a = (ax^{-1})x$.

Definizione:
Sia $\mathcal{A}$ un anello. $x \in \mathcal{A}^*$ si dice *divisore di zero* se esiste $y \in A^*$ tale che $xy = 0$ o $yx = 0$.

Osservazioni:
▶ $\mathcal{A}'$ non contiene divisori di zero, infatti, se $xy = 0$ e x è invertibile, moltiplicando a sinistra per x^{-1} e usando quanto visto sopra otteniamo $y = x^{-1}xy = x^{-1}0 = 0$ (idem se $yx = 0$, moltiplicando a destra).
▶ Se x non è un divisore di zero, valgono le leggi di cancellazione per x: $xa = xb \Rightarrow a = b$, $ax = bx \Rightarrow a = b$ (nel primo caso, ad esempio, si ha $x(a - b) = 0$, da cui $a - b = 0$).

Definizione:
Sia $\mathcal{A}$ un anello commutativo. $\mathcal{A}$ si dice:

- *dominio di integrità* (o più semplicemente *dominio*) se è privo di divisori di zero.
- *campo* se $\mathcal{A}' = \mathcal{A}^*$.

Ad esempio, $(\mathbb{Z}, +, \cdot)$ è un dominio, $(\mathbb{Q}, +, \cdot)$ e $(\mathbb{R}, +, \cdot)$ sono campi.

Osservazioni:
▶ In un dominio, $ab = 0$ se e solo se $a = 0$ o $b = 0$, e quindi vale la legge di

cancellazione: se $a \neq 0$, $ab = ac \Rightarrow b = c$.
▶ Un campo è un dominio.

Ricordiamo altre proprietà fondamentali di $\mathbb{Z}$.

• Su $\mathbb{Z}$ è definita la *divisione con resto*, ovvero vale la seguente proprietà:

$$\forall\ a, m \in \mathbb{Z},\ m \neq 0,\ \exists!\ q, r \in \mathbb{Z},\ \text{con } 0 \leq r < |m|,\ \text{tali che } a = qm + r.$$

q è il *quoziente* e r è il *resto* della divisione di a per m. Notiamo che $m|a$ se e solo se $r = 0$.

• Dati $a, b \in \mathbb{Z}$ non nulli, il massimo comun divisore di a e b, indicato con $mcd(a, b)$ (o semplicemente con (a, b)) è caratterizzato dalle seguenti proprietà:

- è positivo ($mcd(a, b) > 0$);
- divide sia a che b (ovvero a e b sono multipli di $mcd(a, b)$; in simboli, $mcd(a, b)|a$, $mcd(a, b)|b$);
- è maggiore di ogni altro divisore positivo comune di a e b ($m \in \mathbb{Z}$, $m|a$, $m|b$ allora $|m| \leq mcd(a, b)$, e in effetti $m|mcd(a, b)$).

Per completezza si pone $mcd(a, 0) = |a|$.
Osserviamo che $mcd(a, b) = mcd(b, a) = mcd(|a|, |b|)$ e che $a|b$ se e solo se $mcd(a, b) = a$.
L'esistenza del massimo comun divisore si dimostra usando l'*algoritmo Euclideo*, basato sulla reiterazione di divisioni con resto: se $a \geq b > 0$, si divide a per b, $a = q_1 b + r_1$ ($r_1 < b$) e si osserva che $m|a, b \iff m|b, r_1$ e quindi $mcd(a, b) = mcd(b, r_1)$; reiterando, si divide b per r_1, $b = q_2 r_1 + r_2$ ($r_2 < r_1$) e di nuovo $mcd(a, b) = mcd(r_1, r_2)$; proseguendo, la procedura si arresta quando si trova $r_{i+1} = 0$, cioè quando $r_i | r_{i-1}$, e allora $mcd(a, b) = mcd(r_i, r_{i-1}) = r_i$.

• Vale l'*identità di Bézout* (ottenuta ad esempio sostituendo reiteratamente l'espressione di $r_i = r_{i-2} - q_i r_{i-1}$ nella formula successiva dell'algoritmo Euclideo):

$$\forall\ a, b \in \mathbb{Z},\ \exists\ c, d \in \mathbb{Z} \text{ tali che } mcd(a, b) = ac + bd.$$

• $p \in \mathbb{Z}$ si dice *primo*, se per ogni $m \in \mathbb{Z}$, $1 < m < |p|$, $m \not| p$. Osserviamo che p è primo se e solo se $mcd(m, p) = 1$ per ogni $m \in \mathbb{Z}$, $1 < m < |p|$. Ogni $m \in \mathbb{Z}$ si fattorizza in modo unico (a meno dell'ordine e del segno) come prodotto di numeri primi.

Vediamo che esistono altri anelli con proprietà simili a quelle di $\mathbb{Z}$.

Definizione:
Sia $\mathbb{K}$ un campo. Un *polinomio nell'indeterminata t a coefficienti in $\mathbb{K}$* è una scrittura formale $p = a_0 t^0 + a_1 t^1 + a_2 t^2 + \cdots + a_k t^k$ per qualche $k \in \mathbb{N}$,

$a_0, \dots, a_k \in \mathbb{K}$ detti *coefficienti* di p. In notazione compatta $p = \sum_{i=0}^{k} a_i t^i$.
Di solito, nella scrittura di un polinomio si omettono i termini con coefficiente nullo, il termine $a_0 t^0$ si scrive semplicemente a_0 e il temine $a_1 t^1$ si scrive semplicemente $a_1 t$; inoltre, $1t^i$ si scrive semplicemente t^i, e $(-a)t^i$ si scrive semplicemente $-at^i$ (ad esempio, $2t^0 + 1t^1 + 0t^2 + (-2)t^3 + 0t^4 = 2 + t - 2t^3$). Se si vuole evidenziare l'indeterminata, si scrive $p = p(t)$. Il coefficiente a_0 si dice *termine noto* di p. I polinomi del tipo $p = a_0 t^0 = a_0$ (con coefficienti nulli dall'indice 1 in poi) si dicono *costanti*.
Essendo scritture formali, due polinomi sono uguali se e solo se hanno tutti i coefficienti ordinatamente uguali.
Dati due polinomi, $p = \sum_{i=0}^{k} a_i t^i$, $q = \sum_{i=0}^{k} b_i t^i$ (notiamo che a patto di aggiungere termini con coefficiente nullo, possiamo supporre che i due polinomi abbiano una scrittura di questo tipo con lo stesso numero di termini), possiamo definire la loro somma e il loro prodotto:

$$p + q = \sum_{i=0}^{k} (a_i + b_i) t^i$$

$$pq = \sum_{i,j=1}^{k} a_i b_j t^{i+j}$$

ovvero, le operazioni sono definite dalle regole

$$a_i t^i + b_i t^i = (a_i + b_i) t^i,$$

$$(a_i t^i)(b_j t^j) = (a_i b_j) t^{i+j}$$

e dall'imporre la proprietà distributiva.
È facile vedere che si ottiene un anello commutativo, dove 0 è il polinomio costante nullo (il polinomio con tutti i coefficienti nulli), 1 è il polinomio costante con termine noto $a_0 = 1$.
Indichiamo con $\mathbb{K}[t]$ l'anello commutativo dei polinomi nell'indeterminata t a coefficienti in $\mathbb{K}$.

Definizione:
Dato $p \in \mathbb{K}[t]$, $p = \sum_{i=0}^{k} a_i t^i$, il massimo d degli indici i per cui $a_i \neq 0$ si dice *grado* di p, $d = \deg p$ (se $p = 0$ è il polinomio nullo, si pone $\deg p = -\infty$). Il coefficiente a_d si dice *coefficiente direttore* di p. Se $a_d = 1$, p si dice *monico*.

Osservazioni:
▶ I polinomi costanti sono i polinomi di grado minore o uguale a 0.
▶ Per ogni $p, q \in \mathbb{K}[t]$, $\deg(pq) = \deg p + \deg q$, $\deg(p + q) \leq \max(\deg p, \deg q)$.
▶ $\mathbb{K}[t]$ è un domino: se infatti $p, q \in \mathbb{K}[t]$ sono tali che $pq = 0$, allora $\deg p + \deg q = \deg 0 = -\infty$, per cui $p = 0$ o $q = 0$.
▶ Se $p, q \in \mathbb{K}[t]$ sono tali che $p|q$ e $q|p$, allora i due polinomi differiscono per una costante moltiplicativa non nulla. Infatti, scriviamo $p = rq$ e $q = sp$ con $r, s \in \mathbb{K}[t]$. Notiamo che $p = 0$ se e solo se $q = 0$, e in tal caso possiamo

prendere $r = s = 1$. Altrimenti, da $p = rsp$, comparando i gradi abbiamo $\deg p = \deg r + \deg s + \deg p$, da cui $\deg r = \deg s = 0$, cioè r e s sono costanti e non nulli (inoltre, esendo $\mathbb{K}[t]$ un dominio, $rs = 1$).
▶ Se $p, q \in \mathbb{K}[t]$ sono tali che $p|q$ e $\deg p = \deg q$, allora i due polinomi differiscono per una costante moltiplicativa non nulla. Infatti, scriviamo $q = rp$ con $r \in \mathbb{K}[t]$. Notiamo che $p = 0$ se e solo se $q = 0$, e in tal caso possiamo prendere $r = 1$. Altrimenti, comparando i gradi, abbiamo $\deg r = 0$, cioè r è costante e non nullo.
▶ $\mathbb{K}$ si include in $\mathbb{K}[t]$ come i polinomi costanti. $\mathbb{K}[t]$ è quindi un anello che estende $\mathbb{K}$.
▶ Gli elementi invertibili di $\mathbb{K}[t]$ sono i polinomi costanti non nulli, ovvero $\mathbb{K}[t]' = \{p \in \mathbb{K}[t] |\ \deg p = 0\}$: se infatti $p, q \in \mathbb{K}[t]$ sono tali che $pq = 1$, allora $\deg p + \deg q = \deg 1 = 0$, per cui $\deg p = \deg q = 0$. Nell'identificazione di $\mathbb{K}$ con i polinomi costanti, $\mathbb{K}[t]' = \mathbb{K}^*$.

Anche su $\mathbb{K}[t]$ possiamo fare la *divisone con resto*, ovvero vale la seguente proprietà:
Proposizione
Per ogni $p_1, p_2 \in \mathbb{K}[t]$, $p_2 \neq 0$, esistono unici $q, r \in \mathbb{K}[t]$ con $\deg r < \deg p_2$ tali che $p_1 = p_2 q + r$ (per cui $p_2|p_1$ se e solo se $r = 0$).

Dimostrazione
Proviamo l'esistenza per induzione sul grado di p_1.
Se $p_1 = 0$, allora scegliamo $q = r = 0$. Se $p_1 = a_0 \neq 0$ è costante, e $\deg p_2 > 1$, allora scegliamo $q = 0, r = a_0$; se invece anche $p_2 = b_0 \neq 0$ è costante, allora scegliamo $q = \frac{a_0}{b_0}, r = 0$.
Supponiamo quindi $d_1 = \deg p_1 > 1$.
Se $d_2 = \deg p_2 > d_1$, allora $q = 0, r = p_1$.
Se $d_2 \leq d_1$, allora siano a_{d_1} e b_{d_2} i coefficienti direttori di p_1 e p_2 e notiamo che $Q(t) = \frac{a_{d_1}}{b_{d_2}} t^{d_1 - d_2} p_2(t)$ ha grado d_1 e coefficiente direttore a_{d_1}. Quindi $p_3 = p_1 - Q$ ha grado strettamente minore del grado di p_1, e quindi per ipotesi induttiva, esistono $q', r' \in \mathbb{K}[t]$ con $\deg r' < \deg p_2$ tali che $p_3 = q' p_2 + r'$. Ma allora $p_1 = p_3 + Q = q' p_2 + r' + \frac{a_{d_1}}{b_{d_2}} t^{d_1 - d_2} p_2 = (q' + \frac{a_{d_1}}{b_{d_2}} t^{d_1 - d_2}) p_2 + r'$, per cui $q = q' + \frac{a_{d_1}}{b_{d_2}} t^{d_1 - d_2}$ e $r = r'$.
Per l'unicità, se $p_1 = q_1 p_2 + r_1 = q_2 p_2 + r_2$ con $\deg r_1, \deg r_2 < \deg p_2$, allora $(q_1 - q_2) p_2 = r_2 - r_1$. Poiché $\deg(r_2 - r_1) < \deg p_2$, questo può succedere se e solo se $q_1 - q_2 = 0$ e quindi $r_2 - r_1 = 0$. ■

Quindi, in perfetta analogia a quanto visto con $\mathbb{Z}$, anche su $\mathbb{K}[t]$ abbiamo l'algoritmo Euclideo, e tramite esso possiamo definire il massimo comun divisore di due polinomi (unico se si richiede monico). Vale inoltre l'identità di Bézout: per ogni $p, q \in \mathbb{K}[t]$ esistono $r, s \in \mathbb{K}[t]$ tali che $mcd(p, q) = pr + qs$.

L'analogo dei numeri primi interi sono i polinomi irriducibili:

Definizione:

$p \in \mathbb{K}[t]$ non costante si dice *irriducibile* se per ogni polinomio q con $0 < \deg q < \deg p$, $q \not| p$, ovvero gli unici polinomi non nulli che dividono p sono i polinomi invertibili e quelli che differiscono da p per una costante moltiplicativa non nulla.

Osserviamo che se $p \in \mathbb{K}[t]$ è irriducibile, allora dati $q_1, q_2 \in \mathbb{K}[t]$, se $p|q_1q_2$ allora $p|q_1$ o $p|p_2$ (e per induzione, se $p|q_1q_2\cdots q_n$ allora esiste $1 \leq i \leq n$ tale che $p|q_i$). Infatti, consideriamo $d = mcd(p, q_1)$. d divide p ed è monico, quindi ci sono solo due possibilità: $d = 1$ o $d = \frac{1}{a}p$, dove a è il coefficiente direttore di p. Nel secondo caso, $d|q_1$ ed anche $p = ad|q_1$. Nel primo caso, usiamo l'identità di Bézout e scriviamo $1 = rp + sq_1$ con $r, s \in \mathbb{K}[t]$. Allora $q_2 = q_2rp + sq_1q_2$ e poiché $p|sq_1q_2$, allora $p|q_2$.

Proposizione
Ogni polinomio si scrive in modo unico (a meno dell'ordine e di invertibili) come prodotto di polinomi irriducibili.

Dimostrazione
Dato $p \in \mathbb{K}[t]$, se p è irriducibile allora p è prodotto di polinomi irrducibili. Altrimenti, $p = q_1q_2$ con $\deg q_1, \deg q_2 < \deg p$. Se q_1 e q_2 sono entrambi irriducibili allora p è prodotto di polinomi irrducibili. Altrimenti si reitera il procedimento con q_1 o q_2 (ad esempio se q_1 non è irriducibile, scriviamo $q_1 = r_1r_2$ con $\deg r_1, \deg r_2 < \deg q_1$ e $p = r_1r_2q_2$). Notando che ad ogni passo il grado di almeno uno dei fattori cala, dopo un numero finito di passi otteniamo solo polinomi irriducibili e p ne è il prodotto.
Se poi $p = p_1p_2\cdots p_m = q_1q_2\cdots q_n$ con i p_i e i q_j irriducibili, allora $p_1|q_1q_2\cdots q_n$ e quindi p_1 divide uno dei q_j che, a meno di riordinare, possiamo supporre sia q_1. q_1 è però irriducibile, e quindi da $p_1|q_1$ segue $p_1 = \alpha q_1$ con $\alpha \in \mathbb{K}^*$. Adesso, a meno di invertibili, abbiamo $p_2p_3\cdots p_m = q_2q_3\cdots q_n$ e si può reiterare per quanto possibile trovando ad ogni passo un p_i e un q_j che differiscono per una costante moltiplicativa non nulla. Se fosse $m < n$, allora troveremmo $1 = \beta q_{m+1}\cdots q_n$, con $\beta \in \mathbb{K}^*$, ma allora $q_{m+1}, \ldots, q_n$ sarebbero costanti contro il fatto che sono irriducibili $\lightning$. Analogamente, se fosse $n < m$, allora troveremmo $1 = \beta p_{n+1}\cdots p_m$, con $\beta \in \mathbb{K}^*$, ma allora $p_{n+1}, \ldots, p_m$ sarebbero costanti contro il fatto che sono irriducibili $\lightning$. Quindi $m = n$ e i p_i coincidono con i q_j a meno dell'ordine e di una costante moltiplicativa non nulla. ❑

Definizione:
Dato $p \in \mathbb{K}[t]$, $p = \sum_{i=0}^{k} a_it^i$, e $a \in \mathbb{K}$, la *valutazione di p su a* è l'elemento di $\mathbb{K}$ dato da $p(a) = a_0 + a_1a + \cdots + a_ka^k$; $a \in \mathbb{K}$ si dice *radice* di p se $p(a) = 0$.

Osserviamo che, per come sono definite le operazioni tra polinomi, per ogni $p, q \in \mathbb{K}[t]$, $(p + q)(a) = p(a) + q(a)$, $(pq)(a) = p(a)q(a)$.
Ogni polinomio $p \in \mathbb{K}[t]$ definisce quindi una *funzione polinomiale* (che indichiamo ancora con p) $p : \mathbb{K} \to \mathbb{K}$, $a \mapsto p(a)$. Notiamo che due polinomi diversi possono dare la stessa funzione polinomiale (ma questo non succede se il campo è infinito).

Come corollario immediato della divisione con resto abbiamo il:

Teorema di Ruffini:
$a \in \mathbb{K}$ è una radice di $p \in \mathbb{K}[t]$ se e solo se $(t-a)|p$.

Dimostrazione
Se infatti $p(t) = (t-a)q(t)$, allora valutando in a, $p(a) = (a-a)q(a) = 0q(a) = 0$; vicevera, dividiamo p per $t-a$, $p = (t-a)q + r$ con $\deg r < 1$ (ovvero, r è costante); valutando in a, $0 = p(a) = 0q(a) + r$, da cui $r = 0$. ∎

Segue dal teorema di Ruffini che un polinomio di grado $d > 0$ non può avere più di d radici.
Quindi, se $\mathbb{K}$ è infinito, il polinomio nullo è l'unico ad avere infinite radici, ovvero l'unico polinomio che dà la funzione polinomiale nulla è il polinomio nullo. Applicato alla differenza, se $\mathbb{K}$ è infinito, due polinomi sono uguali se e solo se danno la stessa funzione polinomiale.

Se $a \in \mathbb{K}$ è una radice di $p \in \mathbb{K}[t]$, reiterando il teorema di Ruffini, possiamo scrivere $p(t) = (t-a)^k q(t)$ con $q(a) \neq 0$. k si dice *molteplicità algebrica* della radice a.
Un polinomio non costante $p \in \mathbb{K}[t]$ si dice *completamente fattorizzabile* in $\mathbb{K}[t]$ se la somma delle molteplicità algebriche delle sue radici è uguale al suo grado, ovvero, se $t_1, \dots, t_h \in \mathbb{K}$ sono le radici di p, con molteplicità algebriche $m_1, \dots, m_h$, allora $p(t) = \alpha(t-t_1)^{m_1} \cdots (t-t_h)^{m_h}$, con $\alpha \in \mathbb{K}^*$.
Osserviamo che p è completamente fattorizzabile in $\mathbb{K}[t]$ se e solo se la sua fattorizzazione in irriducibili contiene solo polinomi irriducibili di grado 1, ovvero p non è completamente fattorizzabile in $\mathbb{K}[t]$ se esiste un polinomio irriducibile $q \in \mathbb{K}[t]$ con $\deg q > 1$ tale che $q|p$.

Il campo $\mathbb{K}$ si dice *algebricamente chiuso* se ogni polinomio non costante in $\mathbb{K}[t]$ ha almeno una radice. Notiamo che per il teorema di Ruffini, questo è equivalente a dire che tutti i polinomi non costanti in $\mathbb{K}[t]$ sono completamente fattorizzabili.
Il campo $\mathbb{C}$ dei numeri complessi è un esempio di campo algebricamente chiuso.

Definizione:
Sia $\mathcal{A}$ un anello commutativo. Un *ideale* di $\mathcal{A}$ è un sottoinsieme $\mathcal{I} \subset \mathcal{A}$ tale che

- $0 \in \mathcal{I}$;
- è chiuso per somma: $a, b \in \mathcal{I} \Rightarrow a + b \in \mathcal{I}$;
- è chiuso per prodotto *esterno*: $a \in \mathcal{A}$, $b \in I \Rightarrow ab \in \mathcal{I}$ (questa proprietà viene detta anche di *assorbimento del prodotto*).

Ad esempio, $\{0\}$ e $\mathcal{A}$ sono ideali di $\mathcal{A}$. Se $x \in \mathcal{A}$, l'insieme dei multipli di x, $\mathcal{I} = x\mathcal{A} = \{xa \in \mathcal{A} |\ a \in \mathcal{A}\}$ è un ideale detto l'*ideale principale generato da* x (infatti, $0 = x0$; $a_1x + a_2x = (a_1 + a_2)x$); $b(xa) = x(ba)$).

Notiamo che $x\mathcal{A} = \{0\} \iff x = 0$; $x\mathcal{A} = \mathcal{A} \iff x \in \mathcal{A}'$.
Sia $\mathcal{A}$ un anello commutativo e $\mathcal{I}$ un ideale di $\mathcal{A}$. La relazione su $\mathcal{A}$ definita da $a \sim_{\mathcal{I}} b$ se $a - b \in I$ è una relazione di equivalenza.
Infatti, $a - a = 0 \in I$; $a - b \in I \Rightarrow b - a = (-1)(a - b) \in I$; se $a - b \in I$ e $b - c \in I$, allora $a - c = (a - b) + (b - c) \in I$.
Osserviamo che per $x \in \mathcal{A}$, $[x]_{\mathcal{I}} = x + I = \{x + a \in \mathcal{A} |\ a \in \mathcal{I}\}$. Sull'insieme quoziente $\mathcal{A}/\mathcal{I}$, definiamo le operazioni di somma e prodotto:

$$[a]_{\mathcal{I}} + [b]_{\mathcal{I}} = [a + b]_{\mathcal{I}}, \quad [a]_{\mathcal{I}}[b]_{\mathcal{I}} = [ab]_{\mathcal{I}}.$$

Sono ben definite, in quanto (tralasciando il pedice $\mathcal{I}$) se $[a'] = [a]$ e $[b'] = [b]$, $[a' + b'] = [a' - a + a + b' - b + b] = [a + b + (a' - a) + (b' - b)] = [a + b]$ in quanto $(a' - a) + (b' - b) \in I$; $[a'b'] = [(a' + a - a)(b' + b - b)] = [ab + a(b' - b) + (a' - a)b'] = [ab]$ in quanto $a(b' - b) + (a' - a)b' \in I$.
Osserviamo che, poiché definite tramite le operazioni su $\mathcal{A}$, la somma e il prodotto su $\mathcal{A}/\mathcal{I}$ sono associative e commutative e vale la proprietà distributiva; la somma ha per elemento neutro $[0] = \mathcal{I}$, il prodotto ha come elemento neutro $[1] = 1 + \mathcal{I}$.
Se $\mathcal{I} \neq \mathcal{A}$, allora $[0] \neq [1]$ e quindi $\mathcal{A}/\mathcal{I}$ è un anello commutativo.

Ogni ideale $\mathcal{I}$ di $\mathbb{K}[t]$ è principale. Infatti, se $\mathcal{I} \neq \{0\}$, allora esiste un unico $\mu \in K[t]$ monico tale che $\mathcal{I}$ è l'ideale principale generato da μ: $\mathcal{I} = \mu\mathbb{K}[t] = \{p \in \mathbb{K}[t] |\ \mu | p\}$, i multipli polinomiali di μ.
Infatti, sia $p \neq 0$ un polinomio di grado minimo in $\mathcal{I}$, cioè $p \in \mathcal{I}$ e $\deg p \leq \deg q$ per ogni $q \in \mathcal{I}$ (un tale p esiste poiché $\mathcal{I} \neq \{0\}$). Dato $p_1 \in \mathcal{I}$, dividiamo p_1 per p: $p_1 = pq + r$ con $q, r \in \mathbb{K}[t]$, $\deg r < \deg p$. Ma $p_1 \in \mathcal{I}$, $pq \in \mathcal{I}$ implica $r = p_1 - pq \in \mathcal{I}$ e per la minimalità del grado di p, $r = 0$, per cui $p | p_1$. Viceversa, se $p | p_1$, $p_1 = pq \in \mathcal{I}$. Quindi $\mathcal{I} = p\mathbb{K}[t]$ e p è un generatore di $\mathcal{I}$. Se p' è un altro polinomio di grado minimo in $\mathcal{I}$, allora $p | p'$ e $p' | p$, e quindi differiscono per una costante moltiplicativa non nulla. Esiste quindi un unico polinomio monico che genera $\mathcal{I}$, ottenuto dividendo per il coefficiente direttore un qualsiasi generatore.

Se $\mathcal{I} \subset \mathbb{K}[t]$ è l'ideale principale generato da p di grado $d \geq 1$, consideriamo la funzione $R_p : \mathbb{K}[t] \to \mathbb{K}[t]$, che associa ad ogni polinomio q il resto della divisione di q per p. Allora, la relazione di equivalenza $\sim_{\mathcal{I}}$ data da $\mathcal{I}$ e la relazione di equivanlenza $\sim_{R_p}$ data da R_p sono uguali. Infatti, se $p_1 \sim_{\mathcal{I}} p_2$, allora $p_2 = p_1 + sp$ per un $s \in \mathbb{K}[t]$. Dividendo p_1 per p otteniamo $p_1 = qp + r$, con $q, r \in \mathbb{K}[t]$, $\deg r < \deg p$, per cui $p_2 = (q + s)p + r$ e per l'unicità della divisione con resto, $R_p(p_2) = r = R_p(p_1)$. Viceversa, se $R_p(p_1) = R_p(p_2) = r$, scriviamo $p_1 = q_1 p + r$, $p_2 = q_2 p + r$ con $q_1, q_2 \in \mathbb{K}[t]$, e quindi $p_2 - p_1 = (q_2 - q_1)p \in \mathcal{I}$.
Otteniamo quindi che $\mathbb{K}[t]/\mathcal{I}$ è in bigezione con i possibili resti della divisione per p, ovvero con i polinomi di grado minore di d, che indichiamo con $\mathbb{K}_{d-1}[t] = \{q \in \mathbb{K}[t] |\ \deg q \leq d - 1\}$. Ovvero, in ogni classe di equivalenza esiste un unico polinomio di grado minore di p.
Osserviamo che se $d = 1$, allora $\mathbb{K}[t]/\mathcal{I} = \mathbb{K}$ (i polinomi di grado minore o uguale

a 0), e che in generale, la proiezione al quoziente ristretta a $\mathbb{K}$ immerge $\mathbb{K}$ in $\mathbb{K}[t]/\mathcal{I}$ come sottoanello.
$\mathbb{K}$ si identifica quindi con le classi dei polinomi di grado minore o uguale a 0. Per $a \in \mathbb{K}$, scriviamo a al posto di $[a]_{\mathcal{I}}$.
Notiamo che se p non è irriducibile, allora $\mathbb{K}[t]/\mathcal{I}$ non è un domino, in quanto contiene divisori di zero. Scrivendo $p = q_1 q_2$, con $q_1, q_2 \in \mathbb{K}[t]$ non costanti, $\deg q_1, \deg q_2 < \deg p$, abbiamo $[q_1], [q_2] \neq [0]$, ma $[q_1][q_2] = [p] = [0]$.
Se invece p è irriducibile, allora data una classe di equivalenza non nulla $c \in \mathbb{K}[t]/\mathcal{I}$, rappresentiamola con l'unico polinomio $q \in c$ con $\deg q < \deg p$ ($q \neq 0$). Essendo p e q coprimi, per l'identità di Bézout, $1 = mcd(p, q) = rp + sq$, con $r, s \in \mathbb{K}[t]$. Passando alle classi di equivalenza, $[1] = [rp + sq] = [sq] = [s][q]$ e quindi c è invertibile (e $c^{-1} = [s]$). Quindi, dato $p \in \mathbb{K}[t]$, $\mathbb{K}[t]/p\mathbb{K}[t]$ è un campo se e solo se p è irriducibile. In tal caso è un campo che estende $\mathbb{K}$ (e diverso da $\mathbb{K}$ se $\deg p > 1$).
Esiste un'altra procedura che coinvolge $\mathbb{K}[t]$ e produce un campo che estende $\mathbb{K}$ (e che estende $\mathbb{K}[t]$ come anello).
Sia $\mathcal{D}$ un dominio e su $\mathcal{D} \times \mathcal{D}^*$ definiamo la relazione $(x, y) \sim (a, b)$ se $ya = xb$.
È facile verificare che è una relazione di equivalenza: è chiramente riflessiva e simmetrica, se poi $ya = xb$ e $az = wb$, allora $abyw = abxz$ da cui, se $a \neq 0$, $yw = xz$ (poiché $\mathcal{D}$ essendo un dominio ha la proprietà di cancellazione per il prodotto); se $a = 0$, allora $x = w = 0$ e continua a valere $yw = xz$.
Chiamiamo $Frac(\mathcal{D})$ l'insieme quoziente e scriviamo $\frac{a}{b}$ per la classe di equivalenza di (a, b).

Definiamo le operazioni di somma e prodotto su $Frac(\mathcal{D})$:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

È facile verificare che le due operazioni sono ben definite e rendono $Frac(\mathcal{D})$ un anello commutativo, dove $0 = \frac{0}{1} = \frac{0}{m}$, $1 = \frac{1}{1} = \frac{m}{m}$ per ogni $m \in D^*$.
Infatti, se $ya = xb$, $zc = wd$, allora $yz(ad + bc) = xbzd + wdyb = bd(xz + wy)$, per cui $\frac{x}{y} + \frac{w}{z} = \frac{xz+yw}{yz} = \frac{ad+bc}{db} = \frac{a}{b} + \frac{c}{d}$. Inoltre, $yzac = xwbd$, per cui $\frac{x}{y} \cdot \frac{w}{z} = \frac{xw}{yz} = \frac{ac}{bd} = \frac{a}{b} \cdot \frac{c}{d}$.

$\mathcal{D}$ si include in $Frac(\mathcal{D})$ tramite $j : \mathcal{D} \to Frac(\mathcal{D})$, $m \mapsto \frac{m}{1}$ (infatti $\frac{m}{1} = \frac{n}{1}$ se e solo se $m = n$), così che $(Frac(\mathcal{D}), +, \cdot)$ estende $(\mathcal{D}, +, \cdot)$ come anello.
Se $\frac{a}{b} \neq 0$, allora $a \neq 0$ e

$$\frac{a}{b}\frac{b}{a} = \frac{ab}{ab} = 1$$

e quindi $Frac(\mathcal{D})$ è un campo detto *campo delle frazioni di* $\mathcal{D}$.

Per $\mathcal{D} = \mathbb{K}[t]$, otteniamo $\mathbb{K}(t) = Frac(\mathbb{K}[t])$, il *campo delle funzioni razionali*

nell'indeterminata t a coefficenti in $\mathbb{K}$, i cui elementi sono (classi di equivalenza di) rapporti di polinomi con denominatore non nullo ($\frac{p}{q}$, con $p, q \in \mathbb{K}[t]$, $q \neq 0$ e $\frac{p}{q} = \frac{pr}{qr}$ per ogni $r \in \mathbb{K}[t]^*$.
$\mathbb{K} \subset \mathbb{K}[t] \subset \mathbb{K}(t)$, per cui $\mathbb{K}(t)$ è un campo che estende il campo $\mathbb{K}$.

Esempio:
■ Il campo $\mathbb{C}$ dei numeri complessi.

Il polinomio $p = t^2 + 1 \in \mathbb{R}[t]$ è irriducibile (infatti non ha radici reali) e quindi $\mathbb{C} = \mathbb{R}[t]/(p)$ è un campo che estende $\mathbb{R}$.

Poiché $\deg p = 2$, ogni classe di equivalenza in $\mathbb{C}$ è del tipo $[a + bt]$ con $a, b \in \mathbb{R}$ e in ogni classe di equivalenza c'è un unico rappresentante di questo tipo.
Per come sono definite le operazioni sul quoziente, $[a+bt] = [a]+[b][t] = a+b[t]$ (usando l'abuso notazionale di cui sopra, per cui se $x \in \mathbb{R}$, scriviamo x al posto di $[x]$). È comune scrivere $i = [t]$ (detta *unità immaginaria*), per cui ogni $z \in \mathbb{C}$ si scrive in modo unico come $z = a + bi$ con $a, b \in \mathbb{R}$.

Osservazioni: $[t^2 + 1] = 0$, per cui $i^2 = -1$. Considerando il polinomio p come polinomio a coefficienti in $\mathbb{C}$, $p \in \mathbb{C}[t]$, non è più irriducibile, avendo i e $-i$ come radici.

Con questa notazione le operazioni su $\mathbb{C}$ diventano:

$$(a + bi) + (c + di) = (a + b) + (b + d)i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

$0 = 0 + 0i$, $1 = 1 + 0i$, se $z = a + ib$, $-z = -a + (-b)i$ (che scriviamo $-a - bi$).
a si dice la *parte reale* di z, b la *parte immaginaria*, $a = \Re(z)$, $b = \Im(z)$.

$(\mathbb{C}, +, \cdot)$ estende $(\mathbb{R}, +, \cdot)$ e $\mathbb{R}$ si include in $\mathbb{C}$ come gli elementi nella forma $a+0b$ (che scriveremo semplicemente a) con parte immaginaria nulla.
Gli elementi del tipo $0 + bi$ (che scriveremo semplicemente bi) sono detti *immaginari puri* (e sono un'ulteriore copia di $\mathbb{R}$ dentro $\mathbb{C}$, in cui la somma funziona come su $\mathbb{R}$, ma non è chiusa rispetto al prodotto). Quindi dato $z \in \mathbb{C}$, z è reale se e solo se $z = \Re(z)$, z è immaginario puro se e solo se $z = \Im(z)i$.

Definiamo l'applicazione *coniugio*, $- : \mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$, dove se $z = a + bi$, $\overline{z} = a - bi$. È facile vedere che il coniugio rispetta la somma e il prodotto:

$$\overline{z + w} = \overline{z} + \overline{w}, \quad \overline{zw} = \overline{z} \cdot \overline{w} \quad \forall z, w \in \mathbb{C}.$$

Osserviamo che $\mathbb{R} = \{z \in \mathbb{C} |\ z = \overline{z}\}$, mentre i numeri immaginari puri sono invece determinati da $z = -\overline{z}$.

Il prodotto $z\overline{z} = a^2 + b^2 \in \mathbb{R}^+$, e possiamo definire il *modulo* di un numero

complesso, $|\ |: \mathbb{C} \to \mathbb{R}^+$, $z \mapsto |z| = \sqrt{z\overline{z}}$.
È facile vedere che il modulo rispetta il prodotto:

$$|zw| = |z||w| \quad \forall z, w \in \mathbb{C}.$$

Inoltre, $|z| = 0$ se e solo se $z = 0$.

Osserviamo che identificando $\mathbb{C}$ con $\mathbb{R}^2$, tramite $a + bi \mapsto (a, b)$, $\mathbb{R}$ diventa l'asse x, i numeri immaginari puri l'asse y, il coniugio la riflessione rispetto all'asse x, il modulo la distanza dall'origine.

Se ora $z \neq 0$, $|z| \neq 0$ e allora $z\frac{\overline{z}}{|z|^2} = 1$, per cui $\frac{\overline{z}}{|z|^2}$ è l'inverso di z.
Esplicitamente, $(a + bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$.

Spazi vettoriali, sottospazi e applicazioni lineari

Esempio:
■ Applicazioni a valori in un anello.

Sia X un insieme non vuoto e $(\mathcal{A}, +, \cdot)$ un anello.
Poniamo $\mathcal{A}^X = \{f : X \to \mathcal{A}\}$ l'insieme di tutte le applicazioni da X ad $\mathcal{A}$. Operando sui valori, possiamo definire una somma e un prodotto su $\mathcal{A}^X$.
Per $f, g \in \mathcal{A}^X$,

- $f + g : X \to \mathcal{A}$ è data da $(f + g)(x) = f(x) + g(x)\ \forall x \in X$,
- $f \cdot g : X \to \mathcal{A}$ è data da $(f \cdot g)(x) = f(x)g(x)\ \forall x \in X$.

È immediato verificare che $(\mathcal{A}^X, +, \cdot)$ è un anello (0 è l'applicazione nulla (che vale costantemente 0) $0(x) = 0$, 1 è l'applicazione costante $1(x) = 1$, le altre proprietà seguono dalle analoghe su $\mathcal{A}$), e che $\mathcal{A}$ si include in $\mathcal{A}^X$ come le applicazioni costanti. $(\mathcal{A}^X, +, \cdot)$ è quindi un anello che estende $(\mathcal{A}, +, \cdot)$.
Gli invertibili di $\mathcal{A}^X$ sono le applicazioni a valori in $\mathcal{A}'$, quindi anche se $\mathcal{A}$ è un campo, $\mathcal{A}^X$ non lo è se X contiene almeno due elementi.

Su $\mathcal{A}^X$ possiamo definire un altro prodotto (che non è un'operazione!) detto *prodotto per scalari* (che con un leggero abuso indichiamo ancora con $\cdot$)

$$\cdot : \mathcal{A} \times \mathcal{A}^X \to \mathcal{A}^X,\ (a, f) \mapsto a \cdot f,\ \text{dove } (a \cdot f)(x) = af(x)\ \ \forall x \in X.$$

Il prodotto per scalari ha proprietà simili al prodotto in un anello:

1. $\forall a, b \in \mathcal{A}$, $\forall f \in \mathcal{A}^X$, $(ab) \cdot f = a \cdot (b \cdot f)$ (simile all'associatività);
2. $\forall f \in \mathcal{A}^X$, $1 \cdot f = f$ (simile all'esistenza dell'elemento neutro);
3. $\forall a, b \in \mathcal{A}$, $\forall f, g \in \mathcal{A}^X$, $(a + b) \cdot f = a \cdot f + b \cdot f$, $a \cdot (f + g) = a \cdot f + a \cdot g$ (simili alle proprietà distributive).

Osservazioni:
► Per ogni $f \in \mathcal{A}^X$ e per ogni $a \in \mathcal{A}$ abbiamo $0 \cdot f = 0$, $a \cdot 0 = 0$, $(-1) \cdot f = -f$.
► Se $a \in \mathcal{A}'$, allora vale la proprietà di cancellazione $a \cdot f = a \cdot g \Rightarrow f = g$ (ovvero $a \cdot f = 0 \Rightarrow f = 0$).
► Se $\mathcal{A}$ è un campo, il prodotto per scalari "non ha divisori di zero" ($\mathcal{A}^X$ ha una proprietà simile a quella di un domino), ovvero: per $a \in \mathcal{A}$ e $f \in \mathcal{A}^X$, $af = 0$ se e solo se $a = 0$ o $f = 0$.

Formalizziamo questa struttura nella seguente definizione.

Definizione:
Sia $\mathbb{K}$ un campo. Uno *spazio vettoriale su* $\mathbb{K}$ (o $\mathbb{K}$*-spazio vettoriale*) è un insieme V dotato di una operazione $+$ per cui $(V, +)$ è un gruppo abeliano, e di un *prodoto per scalari* $\cdot : \mathbb{K} \times V \to V$, $(\lambda, \underline{v}) \mapsto \lambda \cdot \underline{v}$, tale che

1. $\forall \lambda, \mu \in \mathbb{K}$ e $\forall \underline{v} \in V$, $(\lambda\mu) \cdot \underline{v} = \lambda \cdot (\mu \cdot \underline{v})$;
2. $\forall \underline{v} \in V$, $1 \cdot \underline{v} = \underline{v}$;
3. $\forall \lambda, \mu \in \mathbb{K}$ e $\forall \underline{v} \in V$, $(\lambda + \mu) \cdot \underline{v} = \lambda \cdot \underline{v} + \mu \cdot \underline{v}$;
4. $\forall \lambda \in A$ e $\forall \underline{v}, \underline{w} \in V$, $\lambda \cdot (\underline{v} + \underline{w}) = \lambda \cdot \underline{v} + \lambda \cdot \underline{w}$.

Di solito si omette $\cdot$ e si scrive $\lambda\underline{v}$ al posto di $\lambda \cdot \underline{v}$. Gli elementi di V si dicono *vettori*, quelli di $\mathbb{K}$ *scalari*. Graficamente, per evidenziare la differenza, i vettori verrano sottolineati. Inoltre, quando vorremo evidenziare le operazioni scriveremo $(V, +, \cdot)$.

Conseguenze immediate della definizione:

- $\forall \underline{v} \in V$, $0\underline{v} = \underline{0}$. Infatti,
  $0\underline{v} = (0 + 0)\underline{v} \overset{3}{=} 0\underline{v} + 0\underline{v}$ e si somma a entrambi i membri l'opposto del primo membro.
- $\forall \lambda \in \mathbb{K}$, $\lambda\underline{0} = \underline{0}$. Infatti,
  $\lambda\underline{0} = \lambda(\underline{0} + \underline{0}) \overset{4}{=} \lambda\underline{0} + \lambda\underline{0}$ e si somma a entrambi i membri l'opposto del primo membro.
- $\forall \underline{v} \in V$, $(-1)\underline{v} = -\underline{v}$. Infatti,
  $\underline{v} + (-1)\underline{v} \overset{2}{=} 1\underline{v} + (-1)\underline{v} \overset{3}{=} (1 + (-1))\underline{v} = 0\underline{v} = \underline{0}$.
- Per $\lambda \in \mathbb{K}$, $\underline{v} \in V$ $\lambda\underline{v} = \underline{0}$ se e solo se $\lambda = 0$ o $\underline{v} = \underline{0}$. Infatti,
  se $\lambda \neq 0$, allora esiste $\lambda^{-1} \in \mathbb{K}$ e quindi $\underline{0} = \lambda^{-1}\underline{0} = \lambda^{-1}(\lambda\underline{v}) \overset{1}{=} (\lambda^{-1}\lambda)\underline{v} = 1\underline{v} \overset{2}{=} \underline{v}$.

Definizione:
Dato $(V, +, \cdot)$ spazio vettoriale su $\mathbb{K}$, un sottoinsieme $W \subset V$ è un *sottospazio* se

- contiene $\underline{0}$, $\underline{0} \in W$;
- è chiuso per somma, $\forall \underline{w}_1, \underline{w}_2 \in W$, $\underline{w}_1 + \underline{w}_2 \in W$;
- è chiuso per prodotto per scalari, $\forall \underline{w} \in W$ e $\forall \lambda \in \mathbb{K}$, $\lambda\underline{w} \in W$.

Osservazioni:
▶ Munito della restrizione della somma e del prodotto per scalari di V, W è uno spazio vettoriale su $\mathbb{K}$.
▶ $W = \{\underline{0}\}$ e $W = V$ sono sottospazi (detti "banali").

Esempi:
■ $\mathbb{K}$ dotato delle usuali operazioni è uno spazio vettoriale su $\mathbb{K}$.
In particolare, $\mathbb{C}$ è uno spazio vettoriale su $\mathbb{C}$, ma restringendo gli scalari, è anche uno spazio vettoriale su $\mathbb{R}$. Allo stesso modo, $\mathbb{R}, \mathbb{C}$ e $\mathbb{Q}$ sono spazi vettoriali su $\mathbb{Q}$ e, in generale, se $\mathbb{K}$ è un campo che estende il campo $\mathcal{K}$, ogni spazio vettoriale

su $\mathbb{K}$ è anche uno spazio vettoriale su $\mathcal{K}$.

■ Per ogni insieme non vuoto X, $\mathbb{K}^X$ è uno spazio vettoriale su $\mathbb{K}$.

Data $f : X \to \mathbb{K}$, il *supporto* di f è l'insieme

$$supp(f) = \{x \in X |\ f(x) \neq 0\}.$$

Il sottoinsieme $\mathbb{K}_0^X$ dato dagli elementi di $\mathbb{K}^X$ a supporto finito (tra cui includiamo l'applicazione nulla che è l'unica con supporto vuoto) è un sottospazio. Infatti, $supp(f+g) \subset supp(f) \cup supp(g)$, mentre $supp(\lambda f) = supp(f)$ se $\lambda \neq 0$.

■ $\mathbb{K}^{\mathbb{N}}$ è lo spazio vettoriale su $\mathbb{K}$ delle *successioni* a valori in $\mathbb{K}$.

Una successione si può rappresentare come $a = (a_n)_{n\in\mathbb{N}} = (a_0, a_1, a_2, ...)$ e le operazioni si effettuano "indice per indice": se $b = (b_n)_{n\in\mathbb{N}}$, $a+b = (a_n+b_n)_{n\in\mathbb{N}}$; se $\lambda \in \mathbb{K}$, $\lambda a = (\lambda a_n)_{n\in\mathbb{N}}$. Il sottospazio $\mathbb{K}_0^{\mathbb{N}}$ è dato dalle successioni definitivamente nulle (quelle per cui esiste un $k \in \mathbb{N}$ tale che $a_n = 0$ se $n \geq k$).

■ Per $m, n \in \mathbb{N}^*$ poniamo

$X_{m,n} = \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} = \{(i, j) \in \mathbb{N} \times \mathbb{N} |\ 1 \leq i \leq m, 1 \leq j \leq n\}$.

Ogni elemento f di $\mathbb{K}^{X_{m,n}}$ può essere codificato da una tabella di elementi di $\mathbb{K}$ con m righe e n colonne $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$ (dove usiamo la notazione $i = a\_\_b$ per indicare che l'indice i varia tra gli interi compresi tra a e b), assegnando alla posizione (i, j) intersezione della riga i e della colonna j l'elemeneto di $\mathbb{K}$ $a_{ij} = f((i, j))$, detto *coefficiente di posto* (i, j) della tabella. Una tale tabella è detta *matrice di taglia* $m \times n$ *a coefficienti in* $\mathbb{K}$.

Data una matrice A di taglia $m \times n$, il coefficiente di posto (i, j) di A si indica con $\underset{i\text{—}}{A}\overset{j}{|}$. I coefficienti con primo indice fissato danno una *riga* di A; la i-ma riga di A si indica con $\underset{i\text{—}}{A}$. I coefficienti con secondo indice fissato danno una *colonna* di A; la j-ma colonna di A si indica con $\overset{j}{A}|$. Le righe di A sono matrici di taglia $1 \times n$, le colonne sono matrici di taglia $m \times 1$. Se $R_1, \ldots, R_m$ sono matrici di taglia $1 \times n$, la matrice A di taglia $m \times n$ che ha R_i come i-ma riga si indica con $A = \begin{pmatrix} \underline{R_1} \\ \underline{R_2} \\ \vdots \\ \underline{R_m} \end{pmatrix}$. Se $C^1, \ldots, C^n$ sono matrici di taglia $m \times 1$, la matrice B di taglia $m \times n$ che ha C^j come j-ma colonna si indica con $B = \left(C^1 \middle| C^2 \middle| \cdots \middle| C^n\right)$.

I posti con indici uguali, (i, i), formano la *diagonale* di A.

Graficamente, racchiudiamo la tabella tra parentesi tonde, ad esempio, una matrice di taglia 2×3 è $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. $\underset{2\text{—}}{A}\overset{1}{|} = 4$, $\underset{2\text{—}}{A} = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}$, $\overset{3}{A}| = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$.

la diagonale di A è data dai posti $(1, 1)$ (con coefficiente $\underset{1\text{—}}{A}\overset{1}{|} = 1$) e $(2, 2)$ (con coefficiente $\underset{2\text{—}}{A}\overset{2}{|} = 5$).

La somma di due matrici della stessa taglia (ottenuta dalla somma delle funzioni che esse codificano) effettua l'operazione "posto per posto": se A, B sono

di taglia $m \times n$, $A + B$ è la matrice di taglia $m \times n$ tale che

$$\underset{i}{\underline{(A+B)\big|}}^{j} = \underset{i}{\underline{A\big|}}^{j} + \underset{i}{\underline{B\big|}}^{j}, \quad i = 1\_\_m, j = 1\_\_n,$$

ovvero, usando la notazione $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$, $B = (b_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$,

$$A + B = (a_{ij} + b_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}.$$

Inoltre, il prodotto di una matrice per $\lambda \in \mathbb{K}$ (ottenuto dal prodotto per scalari della funzione che essa codifica) effettua il prodotto "in tutti i posti": se A è di taglia $m \times n$, λA è la matrice di taglia $m \times n$ tale che

$$\underset{i}{\underline{(\lambda A)\big|}}^{j} = \lambda \underset{i}{\underline{A\big|}}^{j}, \quad i = 1\_\_m, j = 1\_\_n,$$

ovvero, usando la notazione $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$,

$$\lambda A = (\lambda a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}.$$

Lo spazio vettoriale su $\mathbb{K}$ ottenuto è detto lo *spazio vettoriale delle matrici di taglia $m \times n$ a coefficienti in $\mathbb{K}$* e si indica con $M(m, n, \mathbb{K})$.
Se $m = n$, le matrici si dicono *quadrate* e scriviamo $M(n, \mathbb{K})$.
Le matrici di $M(m, n, \mathbb{K})$ con l'ultima colonna nulla formano un sottospazio (che, se $n > 1$, è una copia di $M(m, n-1, \mathbb{K})$).

■ Poniamo $\mathbb{K}^m = M(m, 1, \mathbb{K})$. $\mathbb{K}^m = \underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{m \text{ volte}}$, con gli elementi disposti in verticale invece che in orizzontale. $\mathbb{K}^1 = \mathbb{K}$ (formalmente, gli elementi di $\mathbb{K}^1$ sono del tipo (a) con $a \in \mathbb{K}$ e basta togliere le parentesi), e poniamo $\mathbb{K}^0 = \{0\}$. Se $m > 0$, gli elementi di $\mathbb{K}^m$ con l'ultima coordinata nulla formano un sottospazio (che è una copia di $\mathbb{K}^{m-1}$).

■ L'anello dei polinomi $\mathbb{K}[t]$ dotato della usuale somma e della restrizione dell'usuale prodotto ai polinomi di grado minore o uguale a 0 (copia di $\mathbb{K}$ dentro $\mathbb{K}[t]$) è uno spazio vettoriale su $\mathbb{K}$. Per ogni $d \geq 0$, il sottoinsieme $\mathbb{K}_d[t]$ dei polinomi di grado minore o uguale a d è un sottospazio.

■ Con le stesse notazioni del caso $\mathbb{K}^X$, se V è uno spazio vettoriale su $\mathbb{K}$ e X è un insieme non vuoto, allora V^X è uno spazio vettoriale su $\mathbb{K}$. Se $W \subset V$ è un sottospazio, allora W^X si include in V^X come le applicazioni $f : X \to V$ con immagine contenuta in W, ed è un sottospazio.

Definizione:
Siano V e W spazi vettoriali su $\mathbb{K}$. Una $f : V \to W$ si dice *lineare* (o *omomorfismo di spazi vettoriali*) se rispetta la somma e il prodotto per scalari, cioè,

1. $\forall \underline{v}_1, \underline{v}_2 \in V$, $f(\underline{v}_1 + \underline{v}_2) = f(\underline{v}_1) + f(\underline{v}_2)$;

2. $\forall \underline{v} \in V$ e $\forall \mu \in \mathbb{K}$, $f(\mu \underline{v}) = \mu f(\underline{v})$.

Una f che soddisfa 1 si dice *additiva*, una che soddisfa 2 si dice *omogenea*. Se si vuole evidenziare il campo, f si dice $\mathbb{K}$-lineare. Quando diremo che una $f : A \to B$ è lineare, intenderemo che A e B sono spazi vettoriali sullo stesso campo $\mathbb{K}$ e che f è $\mathbb{K}$-lineare.
L'insieme delle applicazioni lineari da V a W si indica con $\mathrm{Hom}(V, W)$.
Le $f : V \to V$ lineari si dicono *endomorfismi* di V e $\mathrm{Hom}(V, V)$ si indica con $\mathrm{End}(V)$, detto lo *spazio degli endomorfismi* di V.

Osservazioni:
▶ Se $f \in \mathrm{Hom}(V, W)$, allora $f(\underline{0}) = \underline{0}$.
Infatti, $f(\underline{0}) = f(\underline{0}+\underline{0}) = f(\underline{0})+f(\underline{0})$ e si somma ad entrambi i membri l'opposto di $f(\underline{0})$.
▶ Se $f \in \mathrm{Hom}(V, W)$, allora per ogni $\underline{v} \in V$, $f(-\underline{v}) = -f(\underline{v})$.
Infatti, $f(-\underline{v}) = f((-1)\underline{v}) = (-1)f(\underline{v}) = -f(\underline{v})$.

Esempi:
■ L'applicazione nulla $0 : V \to W$, $\underline{v} \mapsto \underline{0}$, è lineare.
■ Lapplicazione identità $id_V : V \to V$, $\underline{v} \mapsto \underline{v}$, è lineare. Lo stesso vale per $\lambda id_V : V \to V$, $\underline{v} \mapsto \lambda \underline{v}$.
■ Se $X \subset Y$, ogni $f \in \mathbb{K}^X$ si può *estendere a zero* a Y ponendo per $y \in Y \setminus X$ $f(y) = 0$. Otteniamo $ext : \mathbb{K}^X \to \mathbb{K}^Y$ che è lineare e iniettiva.
■ $f : \mathbb{K}[t] \to \mathbb{K}^{\mathbb{N}}$, $f(a_0 + a_1 t + \cdots + a_k t^k) = (a_0, a_1, \ldots, a_k, 0, 0, \ldots)$ è lineare e iniettiva.
■ L'applicazione *trasposta*, $^\top : M(m, n, \mathbb{K}) \to M(n, m, \mathbb{K})$, $A \mapsto A^\top$, che data una matrice ne scambia le righe con le colonne (in formule: se $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$ $A^\top = (a_{ji})_{\substack{j=1\_\_n \\ i=1\_\_m}}$), è lineare e biunivoca (l'inversa è ancora una trasposta, in quanto $(A^\top)^\top = A$).
■ L'applicazione $vect : M(m, n, \mathbb{K}) \to \mathbb{K}^{mn}$, che data una matrice A riordina i coefficienti in un'unica colonna $vect(A)$ seguendo l'ordine lessicografico (ovvero trasponendo le righe e mettendole in ordine una sotto l'altra) è lineare e biunivoca.
■ L'applicazione *traccia* che ad una matrice associa la somma dei coefficienti sulla diagonale, $\mathrm{tr} : M(m, n, \mathbb{K}) \to \mathbb{K}$, $A \mapsto \sum_{i=1}^{\min(m,n)} \overset{i}{\underset{i\_\_}{A}}$ è lineare e surgettiva.
■ Per ogni $\theta \in \mathbb{R}$, l'applicazione $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, $\binom{a}{b} \mapsto \begin{pmatrix} a\cos(\theta) - b\sin(\theta) \\ a\sin(\theta) + b\cos(\theta) \end{pmatrix}$ è $\mathbb{R}$-lineare e rappresenta la rotazione di centro $\binom{0}{0}$ e angolo θ.
Identificando $\mathbb{R}^2$ con $\mathbb{C}$, la rotazione diventa $r_\theta : \mathbb{C} \to \mathbb{C}$, $z \mapsto e^{i\theta} z$, dove $e^{i\theta} = \cos(\theta) + \sin(\theta) i$, che è $\mathbb{C}$-lineare. (Oss:. $|e^{i\theta}| = 1$, $e^{i\theta_1} e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$ mostra che la circonferenza unitaria $C \subset \mathbb{C}$ munita della restrizione del prodotto su $\mathbb{C}$ è un gruppo.)
■ Date $f, g : V \to W$ lineari, $Z_{f,g} = \{\underline{v} \in V |\ f(\underline{v}) = g(\underline{v})\}$ è un sottospazio di V. Infatti,
- contiene $\underline{0} \in V$, poiché $f(\underline{0}) = \underline{0} = g(\underline{0})$;

- è chiuso per somma, poiché se $\underline{v}_1, \underline{v}_2 \in Z_{f,g}$, $\underline{v}_1 + \underline{v}_2 \in Z_{f,g}$ in quanto $f(\underline{v}_1 + \underline{v}_2) \overset{\text{linearità}}{=} f(\underline{v}_1) + f(\underline{v}_2) \overset{\underline{v}_1, \underline{v}_2 \in Z_{f,g}}{=} g(\underline{v}_1) + g(\underline{v}_2) \overset{\text{linearità}}{=} g(\underline{v}_1 + \underline{v}_2)$.
- è chiuso per prodotto per scalari, poiché se $\underline{v} \in Z_{f,g}$, $\lambda \in \mathbb{K}$, allora $\lambda \underline{v} \in Z_{f,g}$ in quanto $f(\lambda \underline{v}) \overset{\text{linearità}}{=} \lambda f(\underline{v}) \overset{\underline{v} \in Z_{f,g}}{=} \lambda g(\underline{v}) \overset{\text{linearità}}{=} g(\lambda \underline{v})$.
Considerando $V = M(n, \mathbb{K})$, l'applicazione trasposta e l'applicazione identità, $^\top, id_V : V \to V$, il sottoinsieme delle matrici *simmetriche*

$$S(n, \mathbb{K}) = Z_{\top, id_V} = \{A \in M(n, \mathbb{K}) |\ A^\top = A\}$$

è un sottospazio. Allo stesso modo, usando $-id_V$, il sottoinsieme delle matrici *antisimmetriche*

$$A(n, \mathbb{K}) = Z_{\top, -id_V} = \{A \in M(n, \mathbb{K}) |\ A^\top = -A\}$$

è un sottospazio.
■ I sottoinsiemi $\mathcal{D}(\mathbb{R}) \subset \mathcal{C}(\mathbb{R}) \subset \mathbb{R}^\mathbb{R}$ dati rispettivamente dalle funzioni derivabili e dalle funzioni continue sono sottospazi.
L'applicazione derivata $D : \mathcal{D}(\mathbb{R}) \to \mathbb{R}^\mathbb{R}$ è lineare.

Proposizione
$\mathrm{Hom}(V, W)$ è un sottospazio di W^V.

Dimostrazione
• $\mathrm{Hom}(V, W)$ contiene $0 \in W^V$, in quanto l'applicazione nulla è lineare;

• $\mathrm{Hom}(V, W)$ è chiuso per somma, poiché se $f, g : V \to W$ sono lineari, allora $f + g : V \to W$ è lineare in quanto
- è additiva: per ogni $\underline{v}_1, \underline{v}_2 \in V$, $(f+g)(\underline{v}_1 + \underline{v}_2) \overset{\text{def. di } +}{=} f(\underline{v}_1 + \underline{v}_2) + g(\underline{v}_1 + \underline{v}_2) \overset{\text{linearità}}{=} f(\underline{v}_1) + f(\underline{v}_2) + g(\underline{v}_1) + g(\underline{v}_2) \overset{\text{ass. e comm.}}{=} (f(\underline{v}_1) + f(\underline{v}_2)) + (g(\underline{v}_1) + g(\underline{v}_2)) \overset{\text{def. di } +}{=} (f+g)(\underline{v}_1) + (f+g)(\underline{v}_2)$;
- è omogenea: per ogni $\underline{v} \in V$, $\mu \in \mathbb{K}$, $(f+g)(\mu \underline{v}) \overset{\text{def. di } +}{=} f(\mu \underline{v}) + g(\mu \underline{v}) \overset{\text{linearità}}{=} \mu f(\underline{v}) + \mu g(\underline{v}) \overset{4}{=} \mu(f(\underline{v}) + g(\underline{v})) \overset{\text{def. di } +}{=} \mu((f+g)(\underline{v}))$;

• $\mathrm{Hom}(V, W)$ è chiuso per prodotto per scalari, poiché se $f : V \to W$ è lineare e $\lambda \in \mathbb{K}$, allora $\lambda f : V \to W$ è lineare in quanto
- è additiva: per ogni $\underline{v}_1, \underline{v}_2 \in V$, $(\lambda f)(\underline{v}_1 + \underline{v}_2) \overset{\text{def. di } \cdot}{=} \lambda(f(\underline{v}_1 + \underline{v}_2)) \overset{\text{linearità}}{=} \lambda(f(\underline{v}_1) + f(\underline{v}_2)) \overset{4}{=} \lambda f(\underline{v}_1) + \lambda f(\underline{v}_2) \overset{\text{def. di } \cdot}{=} (\lambda f)(\underline{v}_1) + (\lambda f)(\underline{v}_2)$;
- è omogenea: per ogni $\underline{v} \in V$, $\mu \in \mathbb{K}$, $(\lambda f)(\mu \underline{v}) \overset{\text{def. di } \cdot}{=} \lambda(f(\mu \underline{v})) \overset{\text{linearità}}{=} \lambda(\mu f(\underline{v})) \overset{1}{=} (\lambda\mu) f(\underline{v}) \overset{\text{comm.}}{=} (\mu\lambda) f(\underline{v}) \overset{1}{=} \mu(\lambda f(\underline{v})) \overset{\text{def. di } \cdot}{=} \mu((\lambda f)(\underline{v}))$. ◼

Proposizione
Se $f : V \to W$ e $g : W \to Z$ sono lineari, allora la composizione $g \circ f : V \to Z$ è lineare.

Dimostrazione

- $g \circ f$ è additiva: per ogni $\underline{v}_1, \underline{v}_2 \in V$, $(g \circ f)(\underline{v}_2 + \underline{v}_2) = g(f(\underline{v}_1 + v_2)) = g(f(\underline{v}_1) + f(\underline{v}_2)) = g(f(\underline{v}_1) + g(f(\underline{v}_2) = (g \circ f)(\underline{v}_1) + (g \circ f)(\underline{v}_2)$;
- $g \circ f$ è omogenea: per ogni $\underline{v} \in V$, $\mu \in \mathbb{K}$, $(g \circ f)(\mu \underline{v}) = g(f(\mu \underline{v})) = g(\mu f(\underline{v})) = \mu g(f(\underline{v})) = \mu (g \circ f)(\underline{v})$. ■

Osservazioni:
▶ Dato V spazio vettoriale su $\mathbb{K}$, $\mathrm{End}(V)$ è quindi chiuso per composizione $\circ$. Oltre ad essere uno spazio vettoriale su $\mathbb{K}$, $\mathrm{End}(V)$ è quindi dotato dell'ulteriore operazione $\circ$, e $(\mathrm{End}(V), +, \circ)$ è un anello (sottoanello di V^V, in generale non commutativo). È infatti facile verificare che per $f, g : V \to W$, $h : W \to Z$ e $l : U \to V$ lineari,

- $h \circ (f + g) = h \circ f + h \circ g$,
- $(f + g) \circ l = f \circ l + g \circ l$.

Vale inoltre per $\mu \in \mathbb{K}$,

- $h \circ (\mu f) = \mu (h \circ f)$,
- $(\mu f) \circ l = \mu (f \circ l)$.

Il gruppo degli elementi invertibili (rispetto a $\circ$) $\mathrm{End}(V)'$ si indica con $GL(V)$ e si dice *gruppo lineare di* V. $GL(V)$ è un sottogruppo di $\mathrm{S}(V)$, cioè è un gruppo di trasformazioni di V.

Definizione:
$f \in \mathrm{Hom}(V, W)$ si dice *isomorfismo* se è biunivioca e l'inversa insiemistica $f^{-1} : W \to V$ è lineare. In tal caso si dice che V e W sono *isomorfi* (o che V è isomorfo a W traminte f).
(Ricordiamo che se $f : A \to B$ è biunivoca, $f^{-1}(b)$ è l'unico $a \in A$ tale che $f(a) = b$.)

Mostriamo che l'ultima richiesta (f^{-1} è lineare) è conseguenza delle altre due (f è lineare e biunivoca).
Infatti,
- f^{-1} è additiva: per ogni $\underline{w}_1, \underline{w}_2 \in W$, siano $\underline{v}_1, \underline{v}_2 \in V$ tali che $f(\underline{v}_1) = \underline{w}_1$, $f(\underline{v}_2) = \underline{w}_2$ (cioè $\underline{v}_1 = f^{-1}(\underline{w}_1)$, $\underline{v}_2 = f^{-1}(\underline{w}_2)$).
Applicando f^{-1} a $f(\underline{v}_1 + \underline{v}_2) = f(\underline{v}_1) + f(\underline{v}_2) = \underline{w}_1 + \underline{w}_2$ otteniamo
$f^{-1}(\underline{w}_1 + \underline{w}_2) = \underline{v}_1 + \underline{v}_2 = f^{-1}(\underline{w}_1) + f^{-1}(\underline{w}_2)$;
- f^{-1} è omogenea: per ogni $\underline{w} \in W, \mu \in \mathbb{K}$, sia $\underline{v} \in V$ tale che $f(\underline{v}) = \underline{w}$ (cioè $\underline{v} = f^{-1}(\underline{w})$).
Applicando f^{-1} a $f(\mu \underline{v}) = \mu f(\underline{v}) = \mu \underline{w}$ otteniamo $f^{-1}(\mu \underline{w}) = \mu \underline{v} = \mu f^{-1}(w)$.

Esempi:
■ $id_V : V \to V$ è un isomorfismo.
■ Se $f : V \to W$ è un isomorfismo allora $f^{-1} : W \to V$ è un isomorfismo.
■ Gli elementi di $GL(V)$ sono gli isomorfismi da V a V.

■ Poiché la composizione di applicazioni biunivoche è biunivoca e la composizione di applicazioni lineari è lineare, la composizione di isomorfismi è un isomorfismo.
■ La relazione di "essere isomorfi" è dunque una specie di relazione di equivalenza sulla classe degli spazi vettoriali su $\mathbb{K}$.
■ Il sottospazio $\mathbb{K}_0^{\mathbb{N}} \subset \mathbb{K}^{\mathbb{N}}$ delle succesioni definitivamente nulle è isomorfo a $\mathbb{K}[t]$.
■ L'applicazione $vect : M(m,n,\mathbb{K}) \to \mathbb{K}^{mn}$ è un isomorfismo. Quindi se $mn = m'n'$, $M(m,n,\mathbb{K})$ è isomorfo a $M(m',n',\mathbb{K})$.

Definizione:
Sia $f : V \to W$ lineare.
L'*immagine* di f è data da

$$\operatorname{Im} f = \{\underline{w} \in W |\ \exists \underline{v} \in V :\ \underline{w} = f(\underline{v})\}.$$

Il *nucleo* di f è dato da

$$\operatorname{Ker} f = \{\underline{v} \in V |\ f(\underline{v}) = \underline{0}\}.$$

Proposizione
$\operatorname{Im} f$ è un sottospazio di W. $\operatorname{Ker} f$ è un sottospazio di V.

Dimostrazione
Vediamo che $\operatorname{Im} f$ è un sottospazio:
- contiene $\underline{0} \in W$, in quanto $\underline{0} = f(\underline{0})$:
- è chiuso per somma, in quanto se $\underline{w}_1, \underline{w}_2 \in \operatorname{Im} f$, siano $\underline{v}_1, \underline{v}_2 \in V$ tali che $\underline{w}_1 = f(\underline{v}_1), \underline{w}_2 = f(\underline{v}_2)$; allora $\underline{w}_1 + \underline{w}_2 = f(\underline{v}_1) + f(\underline{v}_2) = f(\underline{v}_1 + \underline{v}_2)$;
- è chiuso per prodotto per scalari, in quanto se $\underline{w} \in \operatorname{Im} f$ e $\lambda \in \mathbb{K}$, sia $\underline{v} \in V$ tale che $\underline{w} = f(\underline{v})$; allora $\lambda\underline{w} = \lambda f(\underline{v}) = f(\lambda\underline{v})$.
Per vedere che $\operatorname{Ker} f$ è un sottospazio, consideriamo le applicazioni lineari f e 0, e notiamo che $\operatorname{Ker} f = Z_{f,0}$. ■

Proposizione
Sia $f \in \operatorname{Hom}(V,W)$.
f è surgettiva se e solo se $\operatorname{Im} f = W$.
f è iniettiva se e solo se $\operatorname{Ker} f = \{\underline{0}\}$.

Dimostrazione
La prima proprietà è la definizione di surgettività. Dimostriamo la seconda.
Se f è iniettiva e $\underline{v} \in \operatorname{Ker} f$, da $f(\underline{v}) = \underline{0} = f(\underline{0})$ segue che $\underline{v} = \underline{0}$, per cui $\operatorname{Ker} f = \{\underline{0}\}$.
Se invece $\operatorname{Ker} f = \{\underline{0}\}$, siano $\underline{v}_1, \underline{v}_2 \in V$ tali che $f(\underline{v}_1) = f(\underline{v}_2)$. Allora $f(\underline{v}_1 - \underline{v}_2) = f(\underline{v}_1) - f(\underline{v}_2) = \underline{0}$, ovvero $\underline{v}_1 - \underline{v}_2 \in \operatorname{Ker} f$. Ma allora $\underline{v}_1 - \underline{v}_2 = \underline{0}$, cioè $\underline{v}_1 = \underline{v}_2$. f è quindi iniettiva. ■

Proposizione
Sia $f : V \to W$ lineare e sia $Z \subset V$ un sottospazio.

$f(Z) = \{f(\underline{z})|\ \underline{z} \in Z\}$ è un sottospazio di W.

Dim:
La dimostrazione è analoga a quella fatta per $\operatorname{Im} f = f(V)$.
- $f(Z)$ contiene $\underline{0} \in W$, in quanto $\underline{0} \in Z$ e $\underline{0} = f(\underline{0})$;
- $f(Z)$ è chiuso per somma, in quanto se $\underline{w}_1, \underline{w}_2 \in f(Z)$, siano $\underline{v}_1, \underline{v}_2 \in Z$ tali che $\underline{w}_1 = f(\underline{v}_1), \underline{w}_2 = f(\underline{v}_2)$; allora $\underline{v}_1 + \underline{v}_2 \in Z$ e $\underline{w}_1 + \underline{w}_2 = f(\underline{v}_1) + f(\underline{v}_2) = f(\underline{v}_1 + \underline{v}_2)$;
- $f(Z)$ è chiuso per prodotto per scalari, in quanto se $\underline{w} \in f(Z)$ e $\lambda \in \mathbb{K}$, sia $\underline{v} \in Z$ tale che $\underline{w} = f(\underline{v})$; allora $\lambda\underline{v} \in Z$ e $\lambda\underline{w} = \lambda f(\underline{v}) = f(\lambda\underline{v})$. ❏

Alternativamente, consideriamo la *restrizione* di f a Z, $f_{|_Z} : Z \to W$ ($f_{|_Z}(\underline{z}) = f(\underline{z})$ per ogni $\underline{z} \in Z$). $f_{|_Z}$ è lineare e $\operatorname{Im} f_{|_Z} = f(Z)$.
Osserviamo che $\operatorname{Ker} f_{|_Z} = Z \cap \operatorname{Ker} f$, per cui se $Z \cap \operatorname{Ker} f = \{\underline{0}\}$ (in particolare se f è iniettiva) allora Z è isomorfo a $f(Z)$ (e $f_{|_Z} : Z \to f(Z)$ è un isomorfismo).

Quindi, l'immagine di un sottospazio tramite un'applicazione lineare è un sottospazio. Questo vale anche per le immagini inverse.

Proposizione
Se $U \subset W$ è un sottospazio, allora $f^{-1}(U) = \{\underline{v} \in V|\ f(\underline{v}) \in U\}$ è un sottospazio di V che contiene $\operatorname{Ker} f$.
Dimostrazione
- $f^{-1}(U)$ contiene $\underline{0} \in V$, in quanto $f(\underline{0}) = \underline{0} \in U$;
- $f^{-1}(U)$ è chiuso per somma, in quanto se $\underline{v}_1, \underline{v}_2 \in f^{-1}(U)$, $f(\underline{v}_1 + \underline{v}_2) \in U$ poiché $f(\underline{v}_1 + \underline{v}_2) = f(\underline{v}_1) + f(\underline{v}_2)$, $f(\underline{v}_1), f(\underline{v}_2) \in U$ e U è chiuso per somma;
- $f^{-1}(U)$ è chiuso per prodotto per scalari, in quanto se $\underline{v} \in f^{-1}(U)$ e $\lambda \in \mathbb{K}$, $f(\lambda\underline{v}) \in U$ poiché $f(\lambda\underline{v}) = \lambda f(\underline{v})$, $f(\underline{v}) \in U$ e U è chiuso per prodotto per scalari.
Inoltre, $f(\operatorname{Ker} f) = \{\underline{0}\} \subset U$, quindi $\operatorname{Ker} f \subset f^{-1}(U)$. ❏

Osservazioni:
▶ Per i due sottospazi banali $f(V) = \operatorname{Im} f$, $f(\{\underline{0}\}) = \{\underline{0}\}$, $f^{-1}(W) = V$, $f^{-1}(\{\underline{0}\}) = \operatorname{Ker} f$.
▶ Se $f : V \to W$ e $g : W \to Z$ sono lineari, allora:
$\operatorname{Im}(g \circ f) = g(\operatorname{Im} f) = \operatorname{Im}(g_{|_{\operatorname{Im} f}}) \subset \operatorname{Im} g$
$\operatorname{Ker}(g \circ f) = f^{-1}(\operatorname{Ker} g) \supset \operatorname{Ker} f$.
Notiamo che, in generale, per ogni $Y \subset W$, $f^{-1}(Y) = f^{-1}(Y \cap \operatorname{Im} f)$.

Vediamo altri esempi di spazi vettoriali che si possono costruire a partire da altri spazi vettoriali.

■ Dati V e W due spazi vettoriali su $\mathbb{K}$, possiamo definire una somma ed un prodotto per scalari sul prodotto cartesiano $V \times W$.

Per ogni $(\underline{v}_1, \underline{w}_1), (\underline{v}_2, \underline{w}_2) \in V \times W$, $\lambda \in \mathbb{K}$, poniamo

$$(\underline{v}_1, \underline{w}_1) + (\underline{v}_2, \underline{w}_2) = (\underline{v}_1 + \underline{v}_2, \underline{w}_1 + \underline{w}_2),$$

$$\lambda(\underline{v}_1, \underline{w}_1) = (\lambda\underline{v}_1, \lambda\underline{w}_1).$$

È facile verificare che con tali operazioni $V \times W$ è uno spazio vetoriale su $\mathbb{K}$ ($0 = (0,0)$, $-(\underline{v}, \underline{w}) = (-\underline{v}, -\underline{w})$).
Osservazione: questa procedura si estende al prodotto di una famiglia arbitraria di spazi vettoriali su $\mathbb{K}$.

■ Data $f \in \mathrm{Hom}(V, W)$, il *grafico* di f è dato da

$$\Gamma(f) = \{(\underline{v}, \underline{w}) \in V \times W |\ \underline{w} = f(\underline{v})\}.$$

È facile verificare che $\Gamma(f)$ è un sottospazio di $V \times W$ e che l'applicazione $(id_V, f) : V \to \Gamma(f)$, $\underline{v} \mapsto (\underline{v}, f(\underline{v}))$, è un isomorfismo.

■ Dato V uno spazio vettoriale su $\mathbb{K}$ e $W \subset V$ un sottospazio, definiamo la relazione su V per cui $\underline{v}_1 \sim_W \underline{v}_2$ se $\underline{v}_1 - \underline{v}_2 \in W$. Poiché W è chiuso per somma, $\sim_W$ è una relazione di equivalenza e la somma su V passa al quoziente. Poiché W è chiuso per prodotto per scalari, anche il prodotto per scalari su V passa al quoziente, definendo $\lambda[\underline{v}]_W = [\lambda\underline{v}]_W$. $V\!/\!_W$ è quindi in modo naturale uno spazio vettoriale su $\mathbb{K}$. Inoltre, la proiezione al quoziente $\pi : V \to V\!/\!_W$ è lineare e surgettiva.

Combinazioni Lineari, Operazioni tra Sottospazi.

Dato V uno spazio vettoriale su $\mathbb{K}$ e una famiglia arbitraria non vuota $\mathcal{F}$ di sottospazi di V, l'intersezione degli elementi della famiglia $Z = \bigcap_{W \in \mathcal{F}} W$ è ancora un sottospazio di V.
Infatti,
- poiché $\underline{0} \in W$ per ogni $W \in \mathcal{F}$, $\underline{0} \in Z$;
- se $\underline{z}_1, \underline{z}_2 \in Z$ allora $\underline{z}_1, \underline{z}_2 \in W$ per ogni $W \in \mathcal{F}$, quindi $\underline{z}_1 + \underline{z}_2 \in W$ per ogni $X \in \mathcal{F}$, da cui $\underline{z}_1 + \underline{z}_2 \in Z$;
- se $\underline{z} \in Z$ e $\lambda \in \mathbb{K}$, allora $\underline{z} \in W$ per ogni $W \in \mathcal{F}$, quindi $\lambda \underline{z} \in W$ per ogni $W \in \mathcal{F}$, da cui $\lambda \underline{z} \in Z$.

Definizione:
Dato un sottoinsieme $X \subset V$, sia $\mathcal{F}(X)$ la famiglia dei sottospazi di V che contengono X. $\mathcal{F}(X)$ è non vuota, perché contiene V. L'intersezione della famiglia viene detta *il sottospazio generato da* X e si indica con $\mathrm{Span}(X)$.
Nel caso X sia un insieme finito $X = \{\underline{v}_1, \dots, \underline{v}_n\}$, scriviamo $\mathrm{Span}(\underline{v}_1, \dots, \underline{v}_n)$ al posto di $\mathrm{Span}(\{\underline{v}_1, \dots, \underline{v}_n\})$.

Osservazioni:
▶ $\mathrm{Span}(\emptyset) = \mathrm{Span}(\underline{0}) = \{\underline{0}\}$.
▶ $X \subset \mathrm{Span}(X)$.
▶ $\mathrm{Span}(X) \in \mathcal{F}(X)$ e per ogni $W \in \mathcal{F}(X)$, $\mathrm{Span}(X) \subset W$, per cui $\mathrm{Span}(X)$ è il più piccolo sottospazio (rispetto all'inclusione $\subset$) che contiene X.
▶ $X \subset V$ è un sottospazio se e solo se $\mathrm{Span}(X) = X$.
▶ Se $X \subset Y \subset V$, ogni sottospazio di V che contiene Y contiene anche X, per cui $\mathcal{F}(Y) \subset \mathcal{F}(X)$, e quindi $\mathrm{Span}(X) \subset \mathrm{Span}(Y)$ (ovvero, Span rispetta l'inclusione).
▶ Se $W \subset V$ è un sottospazio, $\mathrm{Span}(X) \subset W$ se e solo se $X \subset W$.

Definizione:
Fissato $X \subset V$, per ogni $f \in \mathbb{K}_0^X$ a supporto finito, ha senso considerare

$$\underline{v}_f = \sum_{\underline{x} \in X} f(\underline{x})\underline{x} = \sum_{\underline{x} \in supp(f)} f(\underline{x})\underline{x} \in V,$$

(in quanto, anche nel caso in cui X sia infinito, la somma ha un numero finito di termini non nulli), dove conveniamo che per l'applicazione nulla (a supporto vuoto) $\underline{v}_0 = \sum_{\underline{x} \in X} 0\underline{x} = \underline{0}$.
Si dice che il vettore $\underline{v}_f$ è una *combinazione lineare* di elementi di X. L'insieme delle combinazioni lineari di elementi di X si indica con $Comb(X)$.

Oss.:
▶ Se $\underline{v}_1, \dots, \underline{v}_k \in V$,

$$Comb(\{\underline{v}_1, \dots, \underline{v}_k\}) = \{x_1\underline{v}_1 + \dots + x_k\underline{v}_k \in V |\ x_1, \dots, x_k \in \mathbb{K}\}.$$

▶ Se $Y \subset X$, $Comb(Y) \subset Comb(X)$.
Infatti ogni combinazione lineare di elementi di Y, $\sum_{\underline{y}\in Y} f(\underline{y})\underline{y}$, è una combinazione lineare di elementi di X, estendendo f a zero su X.
▶ Le applicazioni lineari preservano le combinazioni lineari: se $F \in \mathrm{Hom}(V, W)$ e $\underline{v} \in V$ è combinazione lineare di elementi di X, $\underline{v} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$, allora

$$F(\underline{v}) = F\left(\sum_{\underline{x}\in X} a(\underline{x})\underline{x}\right) \overset{F \text{ additiva}}{=} \sum_{\underline{x}\in X} F(a(\underline{x})\underline{x}) \overset{F \text{ omogenea}}{=} \sum_{\underline{x}\in X} a(\underline{x})F(\underline{x}).$$

Se poniamo $Y = F(X)$ e definiamo $g : Y \to \mathbb{K}$ tale che se $\underline{y} \notin F(supp(f))$, $g(\underline{y}) = 0$, mentre se $\underline{y} \in F(supp(f))$, $g(\underline{y}) = \sum_{\underline{x}\in supp(f):\, F(\underline{x})=\underline{y}} f(\underline{x})$, allora $\underline{v}_f = \underline{v}_g \in Comb(Y)$.
Quindi, $F(Comb(X)) = Comb(F(X))$.

Proposizione
$Comb(X) \in \mathcal{F}(X)$.

Dimostrazione
Per ogni $\underline{x} \in X$, $\underline{x} = 1\underline{x}$, quindi $\underline{x}$ è combinazione lineare di elementi di X (corrisponde alla *funzione caratteristica di* $\underline{x}$, $\chi_{\underline{x}} : X \to \mathbb{K}$ che vale 1 su $\underline{x}$, 0 altrimenti), per cui $X \subset Comb(X)$.
Inoltre, $Comb(X)$ è un sottospazio: contiene $\underline{0}$ in quanto $\underline{0}$ è la combinazione lineare dell'applicazione a supporto vuoto; la somma di due combinazioni lineari è una combinazione lineare (relativa alla somma delle funzioni) e il prodotto per $\lambda \in \mathbb{K}$ di una combinazione lineare è una combinazione lineare (corrispondente alla funzione moltiplicata per λ). Ovvero, l'applicazione $\mathbb{K}_0^X \to V$, $f \mapsto \underline{v}_f$, è lineare con immagine $Comb(X)$. ■

Proposizione
$Comb(X) = \mathrm{Span}(X)$.

Dimostrazione
Ovviamente $\mathrm{Span}(X) \subset Comb(X)$. Se poi $W \in \mathcal{F}(X)$, poiché $X \subset W$ e W è chiuso per somma e prodotto per scalari, tutte le combinazioni lineari di elementi di X appartengono a W, ovvero $Comb(X) \subset W$. Ma allora $Comb(X) \subset \bigcap_{W\in\mathcal{F}(X)} W = \mathrm{Span}(X)$. ■

Se U, W sono sottospazi di V, in generale la loro unione non è un sottospazio. Si ha: $U \cup W$ è un sottospazio se e solo se $U \subset W$ o $W \subset U$.
Infatti, è chiaro che se $U \subset W$ allora $U \cup W = W$ è un sottospazio (e analogamente se $W \subset U$). Viceversa, supponiamo $U \not\subset W$ e $W \not\subset U$. Allora esiste $\underline{u} \in U$ ma $\underline{u} \notin W$ ed esiste $\underline{w} \in W$ ma $\underline{w} \notin U$. Allora $\underline{z} = \underline{u} + \underline{w} \notin U \cup W$. Infatti, se fosse $\underline{z} \in U$, allora $\underline{w} = \underline{z} - \underline{u} \in U$, che non può essere per la scelta di $\underline{w}$. Idem, se fosse $\underline{z} \in W$, allora $\underline{u} = \underline{z} - \underline{w} \in W$ ↯.

Definizione:
Per $U, W \subset V$ sottospazi, il *sottospazio somma* di U e W è

$$U + W = \text{Span}(U \cup W).$$

Poiché ogni combinazione lineare di $U \cup W$ si può sempre scrivere come somma di una combinazione lineare di elementi di U (che è un elemento di U) e di una combinazione lineare di elementi di W (che è un elemento di W), semplicemente raggruppando i termini della combinazione che stanno in U dai rimanenti che quindi stanno in W, e vicevera, la somma di un elemento di U e un elemento di W è una combinazione lineare di elementi di $U \cup W$, abbiamo

$$U + W = \{\underline{u} + \underline{w} \in V |\ \underline{u} \in U, \underline{w} \in W\}.$$

Proposizione
Ogni elemento di $U + W$ si scrive in modo unico come somma di un elemento di U e uno di W se e solo se $U \cap W = \{\underline{0}\}$.

Dimostrazione
Se fosse $\underline{z} \in U \cap W$, $\underline{z} \neq \underline{0}$, allora potremmo scrivere $\underline{0} \in U + W$ in due modi diversi, come $\underline{0} + \underline{0}$ e $\underline{z} + (-\underline{z})$ ↯.
Viceversa, supponiamo che uno $\underline{z} \in U + W$ si possa scrivere come $\underline{z} = \underline{u}_1 + \underline{w}_1 = \underline{u}_2 + \underline{w}_2$ con $\underline{u}_1, \underline{u}_2 \in U$, e $\underline{w}_1, \underline{w}_2 \in W$. Allora $\underline{u}_1 - \underline{u}_2 = \underline{w}_1 - \underline{w}_2 \in U \cap W = \{\underline{0}\}$, per cui $\underline{u}_1 - \underline{u}_2 = \underline{w}_1 - \underline{w}_2 = \underline{0}$, ovvero $\underline{u}_1 = \underline{u}_2$, $\underline{w}_1 = \underline{w}_2$. ❑

Definizione:
Quando $U \cap W = \{\underline{0}\}$, diciamo che U e W sono in *somma diretta* (o che esiste la somma diretta di U e W, o semplicemente che la somma è diretta) e scriviamo $U \oplus W$ al posto di $U + W$.

Se U e W sono in somma diretta, possiamo definire due *proiezioni* della somma sui due fattori, $p_U : U \oplus W \to U$, $p_W : U \oplus W \to W$ nel seguente modo: dato $\underline{v} \in U \oplus W$, scriviamo $\underline{v}$ in modo unico come $\underline{v} = \underline{u} + \underline{w}$ con $\underline{u} \in U$, $\underline{w} \in W$; poniamo $p_U(\underline{v}) = \underline{u}$, $p_W(\underline{v}) = \underline{w}$.
Osserviamo che $p_U + p_W = id_{U \oplus W}$.
Le due proiezioni sono lineari.
Infatti, mostriamo che p_U è lineare, la dimostrazione per p_W è analoga (oppure usiamo $p_W = id_{U \oplus W} - p_U$ e il fatto che somma di lineari è lineare).
Se $\underline{v}_1 = \underline{u}_1 + \underline{w}_1$ e $\underline{v}_2 = \underline{u}_2 + \underline{w}_2$ sono le scrtitture uniche di $\underline{v}_1, \underline{v}_2 \in U \oplus W$, allora la scrittura unica di $\underline{v}_1 + \underline{v}_2$ è $(\underline{u}_1 + \underline{u}_2) + (\underline{w}_1 + \underline{w}_2)$, per cui $p_U(\underline{v}_1 + \underline{v}_2) = \underline{u}_1 + \underline{u}_2 = p_U(\underline{v}_1) + p_U(\underline{v}_2)$.
Se ora $\mu \in \mathbb{K}$ e $\underline{v} = \underline{u} + \underline{w}$ è la scrittura unica di $\underline{v} \in U \oplus W$, allora la scrittura unica di $\mu\underline{v}$ è $\mu\underline{u} + \mu\underline{w}$, per cui $p_U(\mu\underline{v}) = \mu\underline{u} = \mu p_U(\underline{v})$.

Definizione:
Dato un sottospazio W di V, un *supplementare* di W è un sottospazio U di V tale che $U \oplus W = V$.

Se U è un supplementare di W, allora la proiezione al quoziente $\pi : V \to {}^{V}\!/_{W}$ ristretta ad U dà un isomorfismo tra U e ${}^{V}\!/_{W}$.
Infatti, ogni $\underline{v} \in V$ si scrive in modo unico come $\underline{v} = \underline{w} + \underline{u}$ con $\underline{w} \in W, \underline{u} \in U$ e quindi $[\underline{v}] = [\underline{u}] = \pi(\underline{u})$ mostra la surgettività di $\pi_{|_U}$ e $\operatorname{Ker}\pi \cap U = W \cap U = \{\underline{0}\}$ mostra l'iniettività di $\pi_{|_U}$.

Se U' è un altro supplementare di W, allora $F = (\pi_{|_{U'}})^{-1} \circ \pi_{|_U} : U \to U'$ è un isomorfismo "canonico" (o "naturale", ovvero non dipende da alcuna scelta arbitraria).
Esplicitamente, dato $\underline{u} \in U$, scriviamo $\underline{u}$ in modo unico come $\underline{u} = \underline{w} + \underline{u}'$ con $\underline{w} \in W$, $\underline{u}' \in U'$, allora $F(\underline{u}) = \underline{u}'$. (È facile vedere anche direttamente che la F così definita è lineare e un isomorfismo).
Considerando le proiezioni date dalla somma diretta $V = W \oplus U'$, $p_W : V \to W$, $p_{U'} : V \to U'$, $F = p_{U'}{}_{|_U}$.

$$\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)\text{“=”}M(m,n,\mathbb{K})$$

Vediamo come tramite le matrici possiamo rappresentare le applicazioni lineari da $\mathbb{K}^n$ a $\mathbb{K}^m$.

Sia $f\in\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)$.
Per $j=1\_n$, poniamo $C^j=f(\underline{e}_j)\in\mathbb{K}^m$ e organizziamo i C^j in una matrice $M_f\in M(m,n,\mathbb{K})$ la cui colonna j-ma è C^j, $M_f=(C^1|C^2|\cdots|C^n)$.
Osserviamo che per ogni $\underline{v}\in\mathbb{K}^m$, $\underline{v}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$,

$$\begin{aligned} f(\underline{v}) &= f(x_1\underline{e}_1+x_2\underline{e}_2+\cdots+x_n\underline{e}_n)\\ &= x_1f(\underline{e}_1)+x_2f(\underline{e}_2)+\cdots+x_nf(\underline{e}_n)\\ &= x_1C^1+x_2C^2+\cdots+x_nC^n, \end{aligned}$$

per cui l'applicazione f è completamente determinata dalla matrice M_f, detta *matrice associata a f*. Ovvero l'applicazione $M:\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)\to M(m,n,\mathbb{K})$, $f\mapsto M_f$, è iniettiva.

Per $A=(A^1|A^2|\cdots|A^n)\in M(m,n,\mathbb{K})$ e $x_1,\cdots,x_n\in\mathbb{K}$, usiamo la notazione $x_1A^1+x_2A^2+\cdots+x_nA^n=A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$. Quindi $f(\underline{v})=M_f\underline{v}$ per ogni $\underline{v}\in\mathbb{K}^n$.

Viceversa, data $A\in M(m,n,\mathbb{K})$, l'applicazione $L_A:\mathbb{K}^n\to\mathbb{K}^m$, $\underline{v}\mapsto A\underline{v}$, è lineare (detta l'*applicazione lineare data da A*).
Infatti, scrivendo $A=(A^1|A^2|\cdots|A^n)$, se $\underline{v}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$, $\underline{w}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$,
$L_A(\underline{v}+\underline{w})=A(\underline{v}+\underline{w})=A\begin{pmatrix}x_1+y_1\\ \vdots\\ x_n+y_n\end{pmatrix}=(x_1+y_1)A^1+\cdots+(x_n+y_n)A^n=$
$(x_1A^1+\cdots+x_nA^n)+(y_1A^1+\cdots+y_nA^n)=A\underline{v}+A\underline{w}=L_A(\underline{v})+L_A(\underline{w})$.
Se poi $\mu\in\mathbb{K}$,
$L_A(\mu\underline{v})=A(\mu\underline{v})=A\begin{pmatrix}\mu x_1\\ \vdots\\ \mu x_n\end{pmatrix}=\mu x_1A^1+\cdots+\mu x_nA^n=\mu(x_1A^1+\cdots+x_nA^n)=$
$\mu(A\underline{v})=\mu L_A(\underline{v})$.

Poiché $L_A(\underline{e}_i)=A^i$ per ogni $i=1,\dots,n$, $M_{L_A}=A$ e quindi M è anche surgettiva (e dunque biunivoca), e $L:M(m,n,\mathbb{K})\to\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)$, $A\mapsto L_A$, è la sua inversa (è facile verificare direttamente che anche $L_{M_f}=f$).
Possiamo quindi identificare $M(m,n,\mathbb{K})$ e $\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)$ (o dire che, almeno come insiemi, $M(m,n,\mathbb{K})$ “=” $\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)$), poiché possiamo passare da uno all'altro insieme in modo naturale (tramite le applicazioni M e L).
Ma questa identificazione rispetta le operazioni di spazio vettoriale, ovvero M (come pure la sua inversa L) è lineare, e quindi abbiamo che gli spazi vettoriali $(M(m,n,\mathbb{K}),+,\cdot)$ e $(\mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m),+,\cdot)$ sono isomorfi.

Infatti, se $f, g \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$, M_{f+g} si costruisce usando per colonne gli $(f+g)(\underline{e}_i) = f(\underline{e}_i) + g(\underline{e}_i)$, per cui $M_{f+g} = M_f + M_g$. Inoltre, se $\mu \in \mathbb{K}$, $M_{\mu f}$ si costruisce usando per colonne gli $(\mu f)(\underline{e}_i) = \mu f(\underline{e}_i)$, per cui $M_{\mu f} = \mu M_f$.

Poiché anche L è lineare, abbiamo $L_{A+B} = L_A + L_B$, $L_{\mu A} = \mu L_A$ per ogni $A, B \in M(m, n, \mathbb{K})$, $\mu \in \mathbb{K}$ (che possono essere facilmente verificate direttamente).

Osservazioni:
▶ L'applicazione nulla $0 : \mathbb{K}^n \to \mathbb{K}^m$ corrisponde alla matrice nulla di taglia $m \times n$.
▶ L'applicazione identità $id_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$ corrisponde alla matrice quadrata di taglia $n \times n$ $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$.

Gli isomorfismi L e M possono essere usati per dare un analogo matriciale della composizione di applicazioni lineari.
Date $A \in M(m, n, \mathbb{K})$ e $B \in M(s, m, \mathbb{K})$, le applicazioni lineari $L_A : \mathbb{K}^n \to \mathbb{K}^m$, $L_B : \mathbb{K}^m \to \mathbb{K}^s$ possono esse composte a dare $L_B \circ L_A : \mathbb{K}^n \to \mathbb{K}^s$ lineare, di cui possiamo prendere la matrice associata, $C = M_{L_B \circ L_A}$ (di taglia $s \times n$).
La colonna j-ma di C è data $C\underline{e}_j$, ovvero da

$$(L_B \circ L_A)(\underline{e}_j) = L_B(L_A(\underline{e}_j)) = L_B(A^j) = BA^j$$

quindi $C = (BA^1|BA^2|\cdots|BA^n)$.
Diamo quindi la seguente definizione del prodotto tra matrici.

Definizione:
L'applicazione

$$* : M(s, m, \mathbb{K}) \times M(m, n, \mathbb{K}) \to M(s, n, \mathbb{K}), \quad (B, A) \mapsto B * A,$$

dove, se $A = (A^1|A^2|\cdots|A^n)$, $B * A = (BA^1|BA^2|\cdots|BA^n)$, si dice *prodotto righe per colonne* e $B * A$ si dice *prodotto* di B per A.

Osservazioni:
▶ Non tutte le matrici si possono moltiplicare tra di loro, è necessario che il numero di colonne della prima sia uguale al numero di righe della seconda.
▶ È facile verificare che per $f, g : \mathbb{K}^n \to \mathbb{K}^m$, $h : \mathbb{K}^m \to \mathbb{K}^s$ e $l : \mathbb{K}^r \to \mathbb{K}^n$ lineari, $\mu \in \mathbb{K}$,

- $(h \circ f) \circ l = h \circ (f \circ l)$,
- $h \circ (f + g) = h \circ f + h \circ g$,
- $(f + g) \circ l = f \circ l + g \circ l$,

- $h \circ (\mu f) = \mu(h \circ f)$,
- $(\mu f) \circ l = \mu(f \circ l)$,

e quindi, passando alle matrici associate, si ha che per $B, C \in M(m, n, \mathbb{K})$, $A \in M(s, m, \mathbb{K})$ e $D \in M(n, r, \mathbb{K})$ valgono le seguenti uguaglianze:

- $(A * B) * D = A * (B * D)$,
- $A * (B + C) = A * B + A * C$,
- $(B + C) * D = B * D + C * D$,
- $A * (\mu B) = \mu(A * B)$,
- $(\mu B) * D = \mu(B * D)$.

ovvero che il prodotto righe per colonne è associativo, distributivo sulla somma e rispetta il prodotto per scalari (facilmente verificabili anche direttamente dalla definizione).
Inoltre da $id_{\mathbb{K}^m} \circ f = f \circ id_{\mathbb{K}^n} = f$, otteniamo $I_m * A = A * I_n = A$ per ogni $A \in M(m, n, \mathbb{K})$, ovvero le matrici identità sono "elementi neutri" del prodotto righe per colonne (anche questo facilmente verificabile dalla definizione).
In particolare $(M(n, \mathbb{K}), +, *)$ è un anello isomorfo (come anello, tramite L) a $(\mathrm{End}(\mathbb{K}^n), +, \circ)$. Il gruppo degli elementi invertibili di $M(n, \mathbb{K})$ si dice il *gruppo lineare (matriciale) classico* e si indica con $GL(n, \mathbb{K})$ ($GL(n, \mathbb{K}) = M(n, \mathbb{K})'$ e corrisponde agli isomorfismi di $\mathbb{K}^n$, $GL(\mathbb{K}^n)$); è dato dalle matrici di taglia $n \times n$ tali che esiste B di taglia $n \times n$ tale che $A * B = B * A = I_n$. B si dice la *matrice inversa di* A e si indica con A^{-1}. $GL(n, \mathbb{K})$ non è commutativo se $n > 1$.

Da ora in avanti, ometteremo $\circ$ e $*$ e scriveremo semplicemente gf per $g \circ f$ e AB per $A * B$.
Ad esempio, $M_{fg} = M_f M_g$ e $L_A L_B = L_{AB}$.

Esplicitiamo ulteriormente come dipende BA da B e A.
Se $A = (A^1 | A^2 | \cdots | A^n) \in M(m, n, \mathbb{K})$ e $B = \begin{pmatrix} B_1 \\ \hline B_2 \\ \hline \vdots \\ \hline B_s \end{pmatrix} \in M(s, m, \mathbb{K})$, dove B_k è la k-ma riga di B, allora posto $C = BA$, $C = (c_{ij})_{\substack{i=1\_\_s \\ j=1\_\_n}}$, $c_{ij} = B_i A^j$ (prodotto righe per colonne di una matrice di taglia $1 \times m$ e una di taglia $m \times 1$).
Ancora piu esplicito, se $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$, $B = (b_{ij})_{\substack{i=1\_\_s \\ j=1\_\_m}}$, $c_{ij} = \sum_{k=1}^{m} b_{ik} a_{kj}$

Definizione:
Data $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}} \in M(m, n, \mathbb{K})$,
• il *nucleo* di A è il nucleo di $L_A : \mathbb{K}^n \to \mathbb{K}^m$, $\mathrm{Ker}\, A = \mathrm{Ker}\, L_A \subset \mathbb{K}^n$;
• l'*immagine* di A è limmagine di L_A, $\mathrm{Im}\, A = \mathrm{Im}\, L_A \subset \mathbb{K}^m$.

$\operatorname{Ker} A$ è dato dagli $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ tali che $L_A(\underline{X}) = A\underline{X} = \underline{0}$, ovvero tali che

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

quindi $\operatorname{Ker} A$ è lo spazio delle soluzioni di un *sistema lineare omogeneo di m equazioni in n incognite*.

Poiché $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1A^1 + x_2A^2 + \cdots + x_nA^n$, $\operatorname{Im} A = \{A\underline{X} |\ \underline{X} \in \mathbb{K}^n\} = \operatorname{Span}(A^1, \ldots, A^n)$ è il sottospazio di $\mathbb{K}^m$ generato dalle colonne di A, indicato anche con $C(A)$.

Decidere se $\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$, $\underline{b} \neq \underline{0}$, appartenga a $\operatorname{Im} A$, corrisponde a studiare l'esistenza di soluzioni per il sistema lineare *non omogeneo* di m equazioni in n incognite $A\underline{X} = \underline{b}$, ovvero

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

Generatori, Indipendenza Lineare, Basi e Dimensione.

Definizione:
Dato uno spazio vettoriale V, un sottoinsieme $X \subset V$ si dice *insieme di generatori* di V se $V = \mathrm{Span}(X)$. Si dice che che X *genera* V e che V è *generato da* X.
V si dice *finitamente generato* se esiste $X \subset V$ finito che genera V.

Osservazioni:
▶ Gli insiemi di generatori esistono, ad esempio V stesso è un insieme di geńeratori di V.
▶ Per ogni $X \subset V$, X è un insieme di generatori di $\mathrm{Span}(X)$.
▶ Se $X \subset Y \subset V$ e X genera V, allora anche Y genera V (poiché $\mathrm{Span}(X) \subset \mathrm{Span}(Y) \subset V$).
▶ Se U, W sono sottospazi di V, un insieme di generatori per la somma $U + W$ è dato dall'unione di un insieme di generatori di U e di un insieme di generatori di W: $U = \mathrm{Span}(X)$, $W = \mathrm{Span}(Y) \Rightarrow U + V = \mathrm{Span}(X \cup Y)$.

Proposizione
Sia X un insieme di generatori di V e sia $\underline{x}_0 \in X$.
Allora, $X \setminus \{\underline{x}_0\}$ genera V se e solo se $\underline{x}_0 \in \mathrm{Span}(X \setminus \{\underline{x}_0\})$.

Dimostrazione
Se $X \setminus \{\underline{x}_0\}$ genera V, tutti i vettori di V, e quindi anche $\underline{x}_0$, sono combinazione lineare di elementi di $X \setminus \{\underline{x}_0\}$.
Viceversa, scriviamo $\underline{x}_0$ come combinazione lineare di elementi di $X \setminus \{\underline{x}_0\}$, $\underline{x}_0 = \sum_{\underline{x} \in X, \underline{x} \neq \underline{x}_0} f(\underline{x})\underline{x}$. Dato $\underline{v} \in V$, scriviamo $\underline{v}$ come combinazione lineare di elementi di X, $\underline{v} = \sum_{\underline{x} \in X} a(\underline{x})\underline{x} = \sum_{\underline{x} \in X, \underline{x} \neq \underline{x}_0} a(\underline{x})\underline{x} + a(\underline{x}_0)\underline{x}_0 = \sum_{\underline{x} \in X, \underline{x} \neq \underline{x}_0} (a(\underline{x}) + a(\underline{x}_0)f(\underline{x}))\underline{x}$ per cui $\underline{v}$ è combinazione lineare di elementi di $X \setminus \{\underline{x}_0\}$. ❏

Esempi.
■ $V = M(m, n, \mathbb{K})$ è finitamente generato.
Infatti, per $i = 1\_\_m, j = 1\_\_n$, sia E_{ij} la matrice di taglia $m \times n$ con tutti i coefficenti nulli eccetto il coefficiente di posto (i, j) che vale 1. Allora, se $A = (a_{ij})_{\substack{i=1\_\_m \\ j=1\_\_n}}$ si ha $A = \sum_{\substack{i=1\_\_m \\ j=1\_\_n}} a_{ij}E_{ij}$. Quindi $X = \{E_{ij} \in V \mid i = 1\_\_m, j = 1\_\_n\}$ genera V.
Osserviamo che la scritttura di una matrice come combinazione lineare degli E_{ij} è unica.

■ Nel caso di $\mathbb{K}^m = M(m, 1, \mathbb{K})$ si scrive $\underline{e}_i = E_{i1}$, $i = 1\_\_m$, e l'insieme (ordinato) $\{\underline{e}_1, \ldots, \underline{e}_n\}$ si dice la *base canonica* di $\mathbb{K}^m$. Quindi, ogni $\underline{v} \in \mathbb{K}^m$ si scrive in modo unico come combinazione lineare della base canonica:
$\underline{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\underline{e}_1 + x_2\underline{e}_2 + \cdots + x_n\underline{e}_n$.

■ Un insieme di generatori per $\mathbb{K}[t]$ è dato da $\{1, t, t^2, \ldots, t^m, \ldots\}$.
$\mathbb{K}[t]$ non è finitamente generato, infatti se $p_1, \ldots, p_k \in \mathbb{K}[t]$, ogni polinomio in $\mathrm{Span}(p_1, \ldots, p_k)$ ha grado minore o uguale al massimo dei gradi dei p_i.
Se $d \geq 0$, $\mathbb{K}_d[t]$ è generato da $\{1, t, t^2, \ldots, t^d\}$.

Definizione:
Sia V uno spazio vettoriale su $\mathbb{K}$ e $X \subset V$ un sottoinsieme non vuoto.
Diciamo che X è *linearmente indipendente* (o che i vettori di X sono *linearmente indipendenti*) se la combinazione lineare con supporto vuoto (cioè con tutti i coefficienti nulli) è l'unica combinazione lineare di elementi di X che rappresenta $\underline{0} \in V$ come elemento di $\mathrm{Span}(X)$.

Osservazioni:
► Se $\underline{0} \in X$, X non è linearmente indipendente.
► $X = \{\underline{v}\}$ è linearmente indipendente se e solo se $\underline{v} \neq \underline{0}$.
► Se X è linearmente indipendente, allora ogni $Y \subset X$, $Y \neq \emptyset$, è linearmente indipendente (infatti ogni combinazione lineare di elementi di Y è anche una combinazione lineare di elementi di X).

Proposizione
Sia $X \subset V$ un sottoinsieme non vuoto e sia $\underline{x}_0 \in V \setminus X$.
Allora, $X \cup \{\underline{x}_0\}$ è linearmente indipendente se e solo se X è linearmente indipendente e $\underline{x}_0 \notin \mathrm{Span}(X)$.

Dimostrazione
Supponiamo $X \cup \{\underline{x}_0\}$ sia linearmente indipendente. $X \subset X \cup \{\underline{x}_0\}$, quindi X è linearmente indipendente. Se fosse $\underline{x}_0 \in \mathrm{Span}(X)$, potemmo scrivere $\underline{x}_0 = \sum_{\underline{x} \in X} a(\underline{x})\underline{x}$, e quindi $\underline{0} = \underline{x}_0 + \sum_{\underline{x} \in X} -a(\underline{x})\underline{x}$ che mostra che $X \cup \{\underline{x}_0\}$ non sarebbe linearmente indipendente ↯.
Viceversa, supponiamo X sia linearmente indipendente e $\underline{x}_0 \notin \mathrm{Span}(X)$. Data una combinazione lineare nulla di elementi di $X \cup \{x_0\}$, $\underline{0} = \sum_{\underline{x} \in X} a(\underline{x})\underline{x} + a(\underline{x}_0)\underline{x}_0$, se fosse $a(\underline{x}_0) \neq 0$, allora $\underline{x}_0 = \sum_{\underline{x} \in X} -a(\underline{x}_0)^{-1}a(\underline{x})\underline{x} \in \mathrm{Span}(X)$ ↯.
Quindi $a(\underline{x}_0) = 0$ e $\underline{0} = \sum_{\underline{x} \in X} a(\underline{x})\underline{x}$ è una combinazione lineare nulla di elementi di X, per cui $a(\underline{x}) = 0$ per ogni $\underline{x} \in X$. ◪

Osservazioni:
► X non è linearmente indipendente se esiste una combinazione lineare nulla di elementi di X, $\underline{0} = \sum_{\underline{x} \in X} a(\underline{x})\underline{x}$, ed esiste $\underline{x}_0 \in X$ tale che $a(\underline{x}_0) \neq 0$. In questo caso, $\underline{x}_0 = \sum_{\underline{x} \neq \underline{x}_0} (-a(\underline{x}_0)^{-1}a(\underline{x}))\underline{x}$ e $\underline{x}_0 \in \mathrm{Span}(X \setminus \{\underline{x}_0\})$.
Viceversa, se esiste $\underline{x}_0 \in X$ tale che $\underline{x}_0 \in \mathrm{Span}(X \setminus \{\underline{x}_0\})$, $\underline{x}_0 = \sum_{\underline{x} \neq \underline{x}_0} b(\underline{x})\underline{x}$, allora $\underline{0} = \underline{x}_0 + \sum_{\underline{x} \neq \underline{x}_0} (-b(\underline{x}))\underline{x}$ mostra che X non è linearmente indipendente.
► Nel caso X sia finto, possiamo raffinare il criterio:
$X = \{\underline{x}_1, \ldots, \underline{x}_n\}$ non è linearmente indipendente se e solo se esiste un $i = 1 \_ n$ tale che $\underline{x}_i \in \mathrm{Span}(\underline{x}_1, \ldots, \underline{x}_{i-1})$ (esiste un $\underline{x}_i$ che è combinazione lienare dei precedenti. Notare che per $i = 1$ questo implica $\underline{x}_1 = \underline{0}$).

Se infatti esiste un $\underline{x}_i$ che è combinazione lineare dei precedenti, $\{\underline{x}_1, \dots, \underline{x}_i\}$ non è linearmente indipendente e nemmeno X lo è.
Viceversa, data una combinazione lineare nulla con supporto non vuoto $\underline{0} = a_1\underline{x}_1 + \cdots + a_n\underline{x}_n$, sia i il massimo indice per cui $a_i \neq 0$, allora $\underline{x}_i = -(a_i^{-1}a_1\underline{x}_1 + \cdots + a_i^{-1}a_{i-1}\underline{x}_{i-1})$ è combinazione lineare dei precedenti.
► Quindi, $\{\underline{x}_1, \dots, \underline{x}_n\}$ è linearmente indipendente se e solo se per ogni $i = 1\_n$, $\underline{x}_i \notin \mathrm{Span}(\underline{x}_1, \dots, \underline{x}_{i-1})$ (ogni $\underline{x}_i$ non è conbinazione lineare dei precedenti. Notare che questo implica $\underline{x}_1 \neq \underline{0}$).
Questo si può vedere direttamente reiterando la proposizione precedente che in questo caso diventa: $\{\underline{x}_1, \dots, \underline{x}_n, \underline{y}\}$ è linearmente indipendente se e solo se $\{\underline{x}_1, \dots, \underline{x}_n\}$ è linearmente indipendente e $\underline{y} \notin \mathrm{Span}(\underline{x}_1, \dots, \underline{x}_n)$.

Esempi:
- $X = \{E_{ij} |\ i = 1\_m, j = 1\_n\} \subset M(m, n, \mathbb{K})$ è linearmente indipendente.
- La base canonica di $\mathbb{K}^m$ è linearmente indipendente.
- L'insieme dei monomi $\{1, t, t^2, \dots, t^k, \dots\} \subset \mathbb{K}[t]$ è linearmente indipendente.

Proposizione
X è linearmente indipendente se e solo se ogni $\underline{v} \in \mathrm{Span}(X)$ si scrive in modo unico come combinazione lineare di elementi di X.

Dimostrazione
Se X è linearmente indipendente e $\underline{v} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x} = \sum_{\underline{x}\in X} b(\underline{x})\underline{x}$, allora $\sum_{\underline{x}\in X}(a(\underline{x}) - b(\underline{x}))\underline{x} = \underline{0}$, per cui $a(\underline{x}) = b(\underline{x})$ per ogni $\underline{x} \in X$.
Viceversa, se $\underline{0} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$, poiché $\underline{0} = \sum_{\underline{x}\in X} 0\underline{x}$ e la scrittura è unica, $a(\underline{x}) = 0$ per ogni $\underline{x} \in X$. ◼

Osservazioni:
► Se $X \subset V$ è linearmente indipendente, ogni sottoinsieme proprio $Y \subsetneq X$ non genera V.
Infatti, se Y generasse V, scelto $\underline{x}_0 \in X \setminus Y$, $\underline{x}_0$ sarebbe combinazione lineare di elementi di Y e avremmo $\underline{x}_0 = \sum_{\underline{y}\in Y} a(\underline{y})\underline{y}$ e quindi $\underline{0} = \underline{x}_0 + \sum_{\underline{y}\in Y} -a(\underline{y})\underline{y}$ sarebbe una combinazione lineare nulla di elementi di X (estendendo a ponendo $a(\underline{x}_0) = -1$ e poi estendendo a zero a X) a supporto non vuoto ↯.
► Se $X \subset V$ genera V, ogni soprainsieme proprio $Y \supsetneq X$ non è linearmente indipendente.
Infatti, scelto $\underline{y}_0 \in Y \setminus X$, $\underline{y}_0$ è combinazione lineare di elementi di X per cui abbiamo $\underline{y}_0 = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$. Ma allora $\underline{0} = y_0 + \sum_{\underline{x}\in X} -a(\underline{x})\underline{x}$ è una combinazione lineare nulla di elementi di Y (estendendo a ad $X \cup \{\underline{y}_0\}$ ponendo $a(\underline{y}_0) = -1$ e poi estendendo a zero a Y) a supporto non vuoto.
► I due enunciati sopra sono uno la negazione dell'altro.

Definizione:
Un sottoinsieme ordinato $X \subset V$ si dice una *base* di V se è linearmente indipendente e genera V.

Osserviamo che ogni sottoinsieme linearmente indipendente di generatori per V fornisce tante basi quanti modi di ordinarlo.

Gli esempi precedenti danno basi standard per $M(m,n,\mathbb{K})$ (gli E_{ij} si ordinano usualmente in ordine lessicografico), $\mathbb{K}^m$ (la base canonica è ordinata dal pedice) e $\mathbb{K}[t]$ (i t^k si ordinano per grado).

Osservazioni:
▶ Se X è linearmente indipendente, allora X è una base di $\mathrm{Span}(X)$.
▶ X è una base di V se e solo se ogni $\underline{v} \in V$ si scrive in modo unico come combinazione lineare di elementi di X (l'esistenza è equivalente ad essere generatori, $\mathrm{Span}(X) = V$, l'unicità ad essere linearmente indipendenti).

▶ X è una base di V se e solo se è un insieme massimale linearmente indipendente (rispetto all'inclusione $\subset$, ovvero X è linearmente indipendente e ogni $Y \supsetneq X$ non è linearmente indipendente).
Infatti, supponiamo X sia una base e sia $Y \supsetneq X$. Scegliamo $\underline{y} \in Y \setminus X$ e, poiché X genera V, scriviamo $\underline{y}$ come combinazione lineare degli elementi di X, $\underline{y} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$. ma allora $\underline{0} = \underline{y} + \sum_{\underline{x}\in X} -a(\underline{x})\underline{x}$ è una combinazione lineare di elementi di Y a supporto non vuoto, quindi Y non è linearmente indipendente.
Viceversa, supponiamo X sia linearmente indipendente e massimale. Basta dimostrare che X genera V. Sia allora $\underline{v} \in V$. Consideriamo $Y = X \cup \{\underline{v}\}$. Se $Y = X$, $\underline{v} \in X$ e quindi $\underline{v} \in \mathrm{Span}(X)$. Se $Y \supsetneq X$ allora Y non è linearmente indipendente ed esite una combinazione lineare nulla di elementi di Y, $\underline{0} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x} + a(\underline{v})\underline{v}$ a supporto non vuoto. Se fosse $a(\underline{v}) = 0$, allora avremmo $\underline{0} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$ da cui, essendo X linearmente indipendente, $a(\underline{x}) = 0$ per ogni $\underline{x} \in X$, e la combinazione lineare avrebbe supporto vuoto. Quindi, $a(\underline{v}) \neq 0$ e $\underline{v} = \sum_{\underline{x}\in X} -a(\underline{v})^{-1}a(\underline{x})\underline{x} \in \mathrm{Span}(X)$. Si ottiene $\mathrm{Span}(X) = V$ come voluto.

▶ X è una base di V se e solo se è un insieme minimale di generatori per V (rispetto all'inclusione $\subset$, ovvero X genera V e ogni $Y \subsetneq X$ non genera V).
Infatti, supponiamo che X sia una base e sia $Y \subsetneq X$. Scegliamo $\underline{x} \in X \setminus Y$. Se Y generasse V, avremmo che $\underline{x}$ sarebbe combinazione lineare di elementi di Y, $\underline{x} = \sum_{\underline{y}\in Y} a(\underline{y})\underline{y}$, ma allora $\underline{0} = \underline{x} + \sum_{\underline{y}\in Y} -a(\underline{y})\underline{y}$ sarebbe una combinazione lineare nulla di elementi di X a supporto non vuoto. 🗲
Viceversa, supponiamo X sia un insieme di generatori per V e minimale. Basta mostrare che X è linearmente indipendente. Sia allora $\underline{0} = \sum_{\underline{x}\in X} a(\underline{x})\underline{x}$ una combinazione lineare nulla di elementi di X. Se esistesse un $\underline{x}_0 \in X$ tale che $a(\underline{x}_0) \neq 0$, allora $\underline{x}_0 = \sum_{\underline{x}\neq\underline{x}_0} -a(\underline{x}_0)^{-1}a(\underline{x})\underline{x} \in \mathrm{Span}(X \setminus \{\underline{x}_0\})$ e quindi $\mathrm{Span}(X \setminus \{\underline{x}_0\}) = \mathrm{Span}(X) = V$, ovvero $X \setminus \{\underline{x}_0\}$ genererebbe V. 🗲

Sia V uno spazio vettoriale finitamente generato e sia $X = \{\underline{v}_1, \ldots, \underline{v}_n\} \subset V$ un insieme finito (e ordinato) di generatori per V.
Vogliamo mostrare che eliminando, se necessario, alcuni elementi di X, l'insieme

rimanente è una base di V (dove conveniamo che ogni sottoinsieme di X abbia l'ordinamento indotto dall'ordinamento di X). Questa procedura si dice *estrazione di una base* ed è costruttiva, per cui viene presentata sotto forma di algoritmo.
Uno *stato* di X è un sottoinsieme Y di X a sua volta unione di due sottoinsiemi $Y = R \cup T$, dove ogni elemento di R precede (nell'ordinamento di Y) ogni elemento di T (in particolare $R \cap T = \emptyset$). Indichiamo lo stato Y come $\{R|T\}$. Ad esempio, se $X = \{\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{v}_4, \underline{v}_5\}$, $\{\underline{v}_1|\underline{v}_4, \underline{v}_5\}$ è uno stato di X. La barra di separazione sta ad indicare il diverso ruolo dei due sottoinsiemi R e T.
L'algoritmo esegue successivamente n passi $p_1, \ldots, p_n$. Ogni passo p_j ha come input uno stato iniziale $\{R_{j-1}|T_{j-1}\}$ e produce uno stato finale $\{R_j|T_j\}$, che è lo sato iniziale del passo successivo p_{j+1}. Il passo p_1 ha per stato iniziale $\{R_0|T_0\} = \{\emptyset|\underline{v}_1, \ldots, \underline{v}_n\}$. Ad ogni passo, R_j si ottiene da R_{j-1} aggiungendo al più un elemento, mentre T_j si ottiene da T_{j-1} eliminando l'elemento di indice più basso. Quindi, $T_j = \{\underline{v}_{j+1}, \ldots, \underline{v}_n\}$ e il passo p_n ha uno stato finale del tipo $\{R_n|\emptyset\}$, dove R_n è la base cercata. Possiamo pensare a R_{j-1} come ai vettori di X già selezionati per appartenere alla base finale nei passi precedenti, e a R_j come ai vettori selezionati dal passo p_j.
Procediamo per induzione e supponiamo di aver eseguito i passi $p_1, \ldots, p_{j-1}$, per cui $\{R_{j-1}|T_{j-1}\}$ è lo sato iniziale del passo p_j.
Il paso p_j esegue il seguente test su $\underline{v}_j$ (il primo vettore dopo la barra):
$\underline{v}_j \in \mathrm{Span}(R_{j-1})$?
Se la risposta è *sì*, allora $R_j = R_{j-1}$, $T_j = T_{j-1} \setminus \{\underline{v}_j\}$ (in pratica, la barra rimane ferma e si scarta $\underline{v}_j$).
Se la risposta è *no*, allora $R_j = R_{j-1} \cup \{\underline{v}_j\}$, $T_j = T_{j-1} \setminus \{\underline{v}_j\}$ (in pratica, la barra si sposta di un passo verso destra).
Osserviamo che questo descrive anche la base dell'induzione: poiché $\mathrm{Span}(\emptyset) = \{\underline{0}\}$, il passo p_1 esegue il test $\underline{v}_1 = \underline{0}$? Se sì $\underline{v}_1$ viene scartato, altrimenti è il primo vettore della base cercata.
Mostriamo adesso, sempre per induzione, che ogni R_j non vuoto è linearmente indipendente, mentre ogni $R_j \cup T_j$ genera V.
Infatti, se $R_{j-1} = \emptyset$ (cioè, tutti i vettori fino a $\underline{v}_{j-1}$ sono stati scartati poiché nulli), esendo $R_j = \{\underline{v}_j\}$ non vuoto, $\underline{v}_j \neq \underline{0}$ e quindi R_j è linearmente indipendente. Se invece R_{j-1} è non vuoto, allora è linearmente indipendente per ipotesi induttiva e poiché $R_j = R_{j-1}$, oppure si ottiene aggiungendo a R_{j-1} $\underline{v}_j \notin \mathrm{Span}(R_{j-1})$, R_j è linearmente indipendente.
$R_0 \cup T_0 = X$ che genera V. Supponiamo $R_{j-1} \cup T_{j-1}$ generi V, allora $R_j \cup T_j = R_{j-1} \cup T_{j-1}$ oppure si ottiene da $R_{j-1} \cup T_{j-1}$ scartando $\underline{v}_j \in \mathrm{Span}(R_{j-1} \cup T_{j-1} \setminus \{\underline{v}_j\})$, e quindi $R_j \cup T_j$ genera V.
Quindi $R_n = R_n \cup T_n$ è un insieme di generatori di V linearmente indipendente, ovvero una base di V.

Osservazione.
L'algoritmo di estrazione di una base funziona bene anche su una lista ordinata di generatori di V (che premette delle ripetizioni tra i vettori $\underline{v}_i$).

Sia $X = \{\underline{v}_1, \ldots, \underline{v}_n\} \subset V$ un insieme finito (e ordinato) di generatori di V e sia $Z = \{\underline{w}_1, \ldots, \underline{w}_m\} \subset V$ un insieme finito (e ordinato) linearmente indipendente.
Mostriamo che aggiungendo, se necesario, a Z alcuni elementi di X si può estendere Z a base di V. Questa procedura si dice *completamento a base* ed è costruttiva.
Consideriamo infatti la lista ordinata $X' = (\underline{w}_1, \ldots, \underline{w}_m, \underline{v}_1, \ldots, \underline{v}_n)$. Poiché $Z \cup X \supset X$, X' è una lista di generatori di V e possiamo applicare ad X' l'algoritmo di estrazione di una base. Poiché gli elementi di Z sono linearmente indipendenti, l'algoritmo non li scarta, per cui la base estratta sarà del tipo $\{Z, X''\}$, con $X'' \subset X$ come voluto.

Riassumiamo nella seguente proposizione

Proposizione
Sia V uno spazio vettoriale finitamente generato.
Da ogni insieme finito di generatori di V si può estrarre una base di V.
Ogni sottoinsieme finito di V linearmente indipendente può essere completato ad una base di V.

Osservazioni:
▶ Se $W \subset V$ è un sottospazio finitamente generato, sia $Z = \{\underline{w}_1, \ldots, \underline{w}_m\}$ una base di W. Completiamo Z a $\mathcal{B} = \{\underline{w}_1, \ldots, \underline{w}_m, \underline{v}_{m+1}, \ldots, \underline{v}_n\}$ base di V e poniamo $Y = \{\underline{v}_{m+1}, \ldots, \underline{v}_n\}$ e $U = \mathrm{Span}(Y)$. Allora U è un supplementare di W.
Infatti, poiché Z genera W e Y genera U, un insieme di generatori per $U + W$ è dato dalla loro unione $\mathcal{B}$, quindi $U + W = \mathrm{Span}(\mathcal{B}) = V$. Se poi $\underline{v} \in U \cap W$, allora $\underline{v}$ si può scrivere sia come combinazione lineare di elementi di Z, sia come combinazione lineare di elementi di Y, $\underline{v} = \sum_{i=1}^{m} a_i \underline{w}_i = \sum_{i=m+1}^{n} b_i \underline{v}_i$, ma allora $\underline{0} = \sum_{i=1}^{m} a_i \underline{w}_i + \sum_{i=m+1}^{n} -b_i \underline{v}_i$ è una combinazione lineare nulla di elementi di $\mathcal{B}$, che è una base, per cui tutti gli a_i e tutti i b_i sonon nulli. Ma allora $\underline{v} = \underline{0}$, e $U \cap W = \{\underline{0}\}$.

Proposizione
Sia V uno spazio vettoriale finitamente generato e siano $X = \{\underline{v}_1, \ldots, \underline{v}_n\} \subset V$ un insieme di generatori di V e $Z \subset V$ un insieme linearmente indipendente.
Allora Z è finito e contiene al più n elementi.

Dimostrazione:
Supponiamo che Z contenga più di n elementi.
Scegliamo $\underline{w}_1 \in Z$ e consideriamo l'insieme $Y = \{\underline{w}_1\} \cup X$. Poiché contiene X, Y genera V (ma i vettori $\underline{w}_1, \underline{v}_1, \ldots, \underline{v}_n$ non sono linearmente indipendenti).
Applichiamo allora parzialmente l'algoritmo di estrazione di una base alla lista di vettori $(\underline{w}_1, \underline{v}_1, \ldots, \underline{v}_n)$, fermandoci non appena viene eliminato un vettore. Sicuramente $\underline{w}_1$ viene mantenuto e viene scartato uno dei $\underline{v}_i$, diciamo $\underline{v}_{i(1)}$, ottenendo $X_1 = X \setminus \{\underline{v}_{i(1)}\}$ per cui $\{\underline{w}_1\} \cup X_1$ genera V.

Adesso scegliamo un $\underline{w}_2 \in Z \setminus \{\underline{w}_1\}$ e consideriamo $Y = \{\underline{w}_1, \underline{w}_2\} \cup X_1$ e operiamo come prima: scartiamo $\underline{v}_{i(2)}$ da X_1 ottenendo X_2 tale che $\{\underline{w}_1, \underline{w}_2\} \cup X_2$ genera V. Iteriamo aggiungendo un $\underline{w}_j \in Z$ (e togliendo un $\underline{v}_{i(j)}$) alla volta fino a che abbiamo esaurito tutti gli elementi di X e otteniamo $\{\underline{w}_1, \dots, \underline{w}_n\} \subsetneq Z$ che genera V. Ma allora Z non può essere linearmente indipendente ↯. ❑

Abbiamo quindi che in uno spazio vettoriale finitamente generato, gli insiemi linearmente indipendenti sono insiemi finiti.
In particolare, ogni base di V contiene un numero finito di elementi (e tale numero non può essere maggiore del numero di elementi di un insieme di generatori, mentre non può essere minore del numero di elementi di un insieme linearmente indipendente).
Se $B = \{\underline{v}_1, \dots, \underline{v}_n\}$ e $B' = \{\underline{w}_1, \dots, \underline{w}_m\}$ sono due basi di V, poiché B è linearmente indipendente e B' genera V, $n \leq m$. Analogamente, poiché B' è linearmente indipendente e B genera V, $m \leq n$.
Otteniamo la seguente proposizione che motiva la definizione successiva.

Proposizione
Sia V uno spazio vettoriale finitamente generato. Ogni base di V contiene lo stesso numero di elementi.

Definizione:
Sia $V \neq \{\underline{0}\}$ uno spazio vettoriale finitamente generato. La *dimensione* di V, $\dim V$, è il numero di elementi di una qualsiasi base di V.
Si pone $\dim\{\underline{0}\} = 0$.
Quando scriviamo $\dim V = n$ (o scritture analoghe), intenderemo che V è finitamente generato e che n è un numero naturale.

Osservazioni:
▶ $\dim \mathbb{K}^n = n$, $\dim M(m, n, \mathbb{K}) = mn$, $\dim \mathbb{K}_d[t] = d + 1$.
▶ La dimensione è ben definita anche per spazi non finitamente generati (in tal caso coincide con la *cardinalità* di una qualsiasi base, che hanno tutte la stessa cardinalità), ma la dimostrazione è più complessa e conivolge il lemma di Zorn. Ad esempio, $\mathbb{K}[t]$ non ha dimensione finita, ma ogni base è numerabile ($\mathbb{K}[t]$ ha dimensione numerabile). $\mathbb{R}$, come spazio vettoriale su $\mathbb{Q}$ non ha dimensione finita e nemmeno numerabile.

Proposizione
Se $\dim V = n$ e $\underline{v}_1, \dots, \underline{v}_n \in V$, allora
$\{\underline{v}_1, \dots, \underline{v}_n\}$ è una base di $V \iff \underline{v}_1, \dots, \underline{v}_n$ sono linearmente indipendenti $\iff \underline{v}_1, \dots, \underline{v}_n$ sono generatori di V.

Dimostrazione
Le implicazioni verso destra sono ovvie, vediamo le altre.
Supponiamo $\underline{v}_1, \dots, \underline{v}_n$ siano linearmente indipendenti. Se non fossero una base, allora non sarebbero generatori di V, cioè esisterebbe $\underline{v} \in V$ tale che

$\underline{v} \notin \mathrm{Span}(\underline{v}_1, \ldots, \underline{v}_n)$. Ma allora, $\underline{v}_1, \ldots, \underline{v}_n, \underline{v}$ sarebbero linearmente indipendenti e $\dim V \geq n+1$ ↯.
Supponiamo adesso $\underline{v}_1, \ldots, \underline{v}_n$ siano generatori di V. Se non fossero una base, allora non sarebbero linearmente indipendenti, cioè esisterebbe un j tale che $\underline{v}_j \in \mathrm{Span}(\underline{v}_1, \ldots, \underline{v}_{j-1}, \underline{v}_{j+1}, \ldots, \underline{v}_n)$, ma allora $\underline{v}_1, \ldots, \underline{v}_{j-1}, \underline{v}_{j+1}, \ldots, \underline{v}_n$ sarebbero generatori di V e quindi $\dim V \leq n-1$ ↯. ❑

Osservazioni:
► Se $\dim V = n > 0$, allora da ogni insieme di generatori X di V (anche infinito) si può estrarre una base di V.
Infatti, scegliamo $\underline{v}_1 \in X$. Se $X \subset \mathrm{Span}(\underline{v}_1)$, allora $V = \mathrm{Span}(X) \subset \mathrm{Span}(\underline{v}_1) \subset V$, e quindi $V = \mathrm{Span}(\underline{v}_1)$, $n = 1$ e $\underline{v}_1$ dà una base di V. Altrimenti, sia $\underline{v}_2 \in X \setminus \mathrm{Span}(\underline{v}_1)$. $\underline{v}_1, \underline{v}_2$ sono linearmente indipendenti e, come sopra, se $X \subset \mathrm{Span}(\underline{v}_1, \underline{v}_2)$, allora $V = \mathrm{Span}(\underline{v}_1, \underline{v}_2)$, $n = 2$ e $\underline{v}_1, \underline{v}_2$ danno una base di V. Altrimenti, possiamo reiterare scegliendo $\underline{v}_3 \in X \setminus \mathrm{Span}(\underline{v}_1, \underline{v}_2)$ tale che $\underline{v}_1, \underline{v}_2, \underline{v}_3$ sono linearmente indipendenti. Poiché $\dim V = n$, il procedimento deve terminare in n passi producendo n vettori linearmente indipendenti che quindi danno una base di V.

Proposizione
Se $\dim V = n$ e $W \subset V$ è un sottospazio, allora W è finitamente generato e $\dim W \leq n$.

Dimostrazione
Se $W = \{0\}$ non c'è niente da dimostrare. Altrimenti, osserviamo che ogni insieme $X \subset W$ linearmente indipendente rimane linearmente indipendente anche se pensato come sottoinsieme di V. Quindi X contiene al più $n = \dim V$ elementi. ❑

Osservazioni:
► Il sottospazio nullo è l'unico sottospazio di dimensione minima 0.
► Se $\dim V = n$, V è l'unico sottospazio di dimensione massima n. Infatti, se $W \subset V$ è un sottospazio e $\dim W = \dim V = n$, allora data una base di W, $\{\underline{w}_1, \ldots, \underline{w}_n\}$, se questa non fosse una base di V allora non sarebbe un insieme di generatori di V (essendo linearmente indipendente), e potrebbe essere completato ad una base di V che avrebbe più di n elementi ↯.
Useremo spesso questo fatto nella forma seguente: dati $U, W \subset V$ due sottospazi tali che $U \subset W$ e $\dim U = \dim W = m$, allora $U = W$.

Coordinate e Matrici Associate

Proposizione
Sia V uno spazio vettoriale su $\mathbb{K}$, $\dim V = n$, e sia $\{\underline{v}_1, \dots, \underline{v}_n\}$ una base di V. Sia W uno spazio vettoriale su $\mathbb{K}$ e siano $\underline{w}_1, \dots, \underline{w}_n \in W$.
Allora esiste un'unica $f : V \to W$ lineare tale che $f(\underline{v}_i) = \underline{w}_i$ per ogni $i = 1\_n$.

Dimostrazione
Per $\underline{v} \in V$, scriviamo in modo unico $\underline{v}$ come combinazione lineare di $\underline{v}_1, \dots, \underline{v}_n$, $\underline{v} = \sum_{i=1}^n a_i \underline{v}_i$, con $a_i \in \mathbb{K}$ per ogni $i = 1\_n$, e poniamo

$$f(\underline{v}) = \sum_{i=1}^n a_i \underline{w}_i.$$

Mostriamo che f è lineare. Dati $\underline{v}, \underline{w} \in V$, scriviamo $\underline{v}$ e $\underline{w}$ in modo unico come combinazione lineare di $\underline{v}_1, \dots, \underline{v}_n$, $\underline{v} = \sum_{i=1}^n a_i \underline{v}_i$ e $\underline{w} = \sum_{i=1}^n b_i \underline{v}_i$, con $a_i, b_i \in \mathbb{K}$ $i = 1\_n$. Abbiamo $\underline{v} + \underline{w} = \sum_{i=1}^n (a_i + b_i)\underline{v}_i$ e questo è l'unico modo di esprimere $\underline{v} + \underline{w}$ come combinazione lineare di $\underline{v}_1, \dots, \underline{v}_n$, per cui

$$f(\underline{v} + \underline{w}) = \sum_{i=1}^n (a_i + b_i)\underline{w}_i = \sum_{i=1}^n a_i \underline{w}_i + \sum_{i=1}^n b_i \underline{w}_i = f(\underline{v}) + f(\underline{w}).$$

Se poi $\mu \in \mathbb{K}$, $\mu\underline{v} = \sum_{i=1}^n (\mu a_i)\underline{v}_i$ e questo è l'unico modo di esprimere $\mu\underline{v}$ come combinazione lienare di $\underline{v}_1, \dots, \underline{v}_n$, per cui

$$f(\mu\underline{v}) = \sum_{i=1}^n (\mu a_i)\underline{w}_i = \mu \sum_{i=1}^n a_i \underline{w}_i = \mu f(\underline{v}).$$

Inoltre, per ogni j, $\underline{v}_j = 0\underline{v}_1 + \cdots 0\underline{v}_{j-1} + 1\underline{v}_j + 0\underline{v}_{j+1} + \cdots + 0\underline{v}_n$ è l'unico modo di esprimere $\underline{v}_j$ come combinazione lienare di $\underline{v}_1, \dots, \underline{v}_n$, per cui

$$f(\underline{v}_j) = 0\underline{w}_1 + \cdots 0\underline{w}_{j-1} + 1\underline{w}_j + 0\underline{w}_{j+1} + \cdots + 0\underline{w}_n = \underline{w}_j$$

come voluto.
Sia adesso $F : V \to W$ lineare tale che $F(\underline{v}_i) = \underline{w}_i$, $i = 1\_n$. Per ogni $\underline{v} \in V$, scriviamo in modo unico $\underline{v}$ come combinazione lineare di $\underline{v}_1, \dots, \underline{v}_n$, $\underline{v} = \sum_{i=1}^n a_i \underline{v}_i$. Allora

$$F(\underline{v}) = F(\sum_{i=1}^n a_i \underline{v}_i) = \sum_{i=1}^n F(a_i \underline{v}_i) = \sum_{i=1}^n a_i F(\underline{v}_i) = \sum_{i=1}^n a_i \underline{w}_i = f(\underline{v}),$$

per cui $F = f$. ❏

Osservazioni:
▶ L'immagine della f costruita sopra è generata da $\underline{w}_1, \dots, \underline{w}_n$ e quindi ha dimensione finita $\dim \operatorname{Im} f = \dim \operatorname{Span}(\underline{w}_1, \dots, \underline{w}_n) \leq n$.
▶ Allo stesso modo, se $f \in \operatorname{Hom}(V, W)$,

$f(\mathrm{Span}(\underline{v}_1,\dots,\underline{v}_k)) = \mathrm{Span}(f(\underline{v}_1),\dots,f(\underline{v}_k))$.
▶ Otteniamo anche un criterio di uguaglianza per applicazioni lineari:
$f, g \in \mathrm{Hom}(V,W)$ sono uguali se e solo se f e g hanno gli stessi valori su una base di V.

Proposizione
Due spazi vettoriali su $\mathbb{K}$, V e W, di dimensione finita sono isomorfi se e solo se $\dim V = \dim W$.

Dimostrazione
Se $F : V \to W$ è un isomorfismo, fissiamo $\mathcal{B} = \{\underline{v}_1,\dots,\underline{v}_n\}$ una base di V, e poniamo $F(\mathcal{B}) = \{F(\underline{v}_1),\dots,F(\underline{v}_n)\}$. Mostriamo che $F(\mathcal{B})$ è una base di W.
Infatti, poiché F è surgettiva, $W = \mathrm{Im}\, F = \mathrm{Span}(F(\underline{v}_1),\dots,F(\underline{v}_n))$, per cui $F(\mathcal{B})$ genera W. Se poi $\underline{0} = \sum_{i=1}^n a_i F(\underline{v}_i)$ è una combinazione lineare nulla di elementi di $F(\mathcal{B})$, allora $\underline{0} = \sum_{i=1}^n a_i F(\underline{v}_i) = F(\sum_{i=1}^n a_i \underline{v}_i)$, ovvero $\sum_{i=1}^n a_i \underline{v}_i \in \mathrm{Ker}\, F$, ma poiché F è iniettiva, $\mathrm{Ker}\, F = \{\underline{0}\}$, per cui $\underline{0} = \sum_{i=1}^n a_i \underline{v}_i$. Essendo $\underline{v}_1,\dots,\underline{v}_n$ linearmente indipendenti, $a_i = 0$ per ogni i e quindi $F(\mathcal{B})$ è linearmente indipendente.
Viceversa, supponiamo $\dim V = \dim W = n$, fissiamo una base di V, $\underline{v}_1,\dots,\underline{v}_n$ e una base di W, $\underline{w}_1,\dots,\underline{w}_n$ e consideriamo l'unica $f \in \mathrm{Hom}(V,W)$ tale che $f(\underline{v}_i) = \underline{w}_i$, $i = 1\_n$. Poiché $\underline{w}_1,\dots,\underline{w}_n$ sono generatori di W, $\mathrm{Im}\, f = \mathrm{Span}(\underline{w}_1,\dots,\underline{w}_n) = W$ e f è surgettiva. Se poi $\underline{v} = \sum_{i=1}^n a_i \underline{v}_i \in \mathrm{Ker}\, f$, allora $\underline{0} = f(\underline{v}) = \sum_{i=1}^n a_i \underline{w}_i$, ed essendo $\underline{w}_1,\dots,\underline{w}_n$ linearmente indipendenti, $a_i = 0$ per ogni i, e dunque $\underline{v} = 0$. Poiché $\mathrm{Ker}\, f = \{0\}$, f è anche iniettiva e quindi un isomorfismo. ◼

Osservazioni:
▶ Se $F : V \to W$ è lineare ed esiste una base $\mathcal{B}$ di V tale che $F(\mathcal{B})$ è una base di W, allora F è un isomorfismo.
Viceversa, se $F : V \to W$ è un isomorfismo, per ogni base $\mathcal{B}$ di V, $F(\mathcal{B})$ è una base di W.
▶ La proposizione vale anche in dimensione infinita.
▶ Se $A \in M(n,\mathbb{K})$, allora A è invertibile $\iff$ $L_A \in \mathrm{End}(\mathbb{K}^n)$ è un isomorfismo $\iff$ le colonne di A (che generano $\mathrm{Im}\, A = \mathrm{Im}\, L_A$ essendo l'immagine tramite L_A della base canonica di $\mathbb{K}^n$) sono una base di $\mathbb{K}^n$ $\iff$ le colonne di A (che sono n) sono linearmente indipendenti $\iff$ le colonne di A generano $\mathbb{K}^n$.

Definizione:
Sia V uno spazio vettoriale su $\mathbb{K}$, $\dim V = n$, e sia $\mathcal{B} = \{\underline{v}_1,\dots,\underline{v}_n\}$ una base di V. Consideriamo $\mathbb{K}^n$ e la base canonica $\mathcal{C}an_n = \{\underline{e}_1,\dots,\underline{e}_n\}$. L'isomorfismo ottenuto mandando $\mathcal{B}$ in $\mathcal{C}an_n$ si dice *isomorfismo di passaggio alle coordinate nella base* $\mathcal{B}$ e si indica con $[\ ]_{\mathcal{B}} : V \to \mathbb{K}^n$. Se $\underline{v} \in V$, $\underline{v} = \sum_{i=1}^n a_i \underline{v}_i$, allora $[v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ che si dicono le *coordinate di* $\underline{v}$ *rispetto alla base* $\mathcal{B}$.

Definizione:
Siano V, W spazi vettoriali su $\mathbb{K}$, $\dim V = n$, $\dim W = m$. Sia $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ una base di V e sia $\mathcal{D} = \{\underline{w}_1, \ldots, \underline{w}_m\}$ una base di W.
Data $f \in \mathrm{Hom}(V, W)$, la composizione $[\ ]_{\mathcal{D}} \circ f \circ [\ ]_{\mathcal{B}}^{-1} \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ si rappresenta con una matrice $[\ ]_{\mathcal{D}} \circ f \circ [\ ]_{\mathcal{B}}^{-1} = L_M$ con $M \in M(m, n, \mathbb{K})$. La matrice M è detta la *matrice associata ad* f (o che *rappresenta* f) *nelle basi* $\mathcal{B}$ *e* $\mathcal{D}$ e si indica con $M_{\mathcal{D}}^{\mathcal{B}}(f)$.

Osservazioni:
▶ Per ogni $\underline{v} \in V$ abbiamo, per definizione, $[f(\underline{v})]_{\mathcal{D}} = M_{\mathcal{D}}^{\mathcal{B}}(f)[\underline{v}]_{\mathcal{B}}$, per cui $M_{\mathcal{D}}^{\mathcal{B}}(f) = \Big([f(\underline{v}_1)]_{\mathcal{D}} \Big| [f(\underline{v}_2)]_{\mathcal{D}} \Big| \cdots \Big| [f(\underline{v}_n)]_{\mathcal{D}}\Big)$.
▶ Nel caso di $f : \mathbb{K}^n \to \mathbb{K}^m$, la matrice M_f è la matrice associata a f nelle basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$, $M_f = M_{\mathcal{C}an_m}^{\mathcal{C}an_n}(f)$.

Consideriamo l'applicazione $M_{\mathcal{D}}^{\mathcal{B}} : \mathrm{Hom}(V, W) \to M(m, n, \mathbb{K})$, $f \mapsto M_{\mathcal{D}}^{\mathcal{B}}(f)$.
$M_{\mathcal{D}}^{\mathcal{B}}$ è lineare. Infatti, se $f, g \in \mathrm{Hom}(V, W)$, allora

$$\begin{aligned}
M_{\mathcal{D}}^{\mathcal{B}}(f+g) &= \Big([(f+g)(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [(f+g)(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \Big([f(\underline{v}_1) + g(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [f(\underline{v}_n) + g(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \Big([f(\underline{v}_1)]_{\mathcal{D}} + [g(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [f(\underline{v}_n)]_{\mathcal{D}} + [g(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \Big([f(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [f(\underline{v}_n)]_{\mathcal{D}}\Big) + \Big([g(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [g(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= M_{\mathcal{D}}^{\mathcal{B}}(f) + M_{\mathcal{D}}^{\mathcal{B}}(g)
\end{aligned}$$

Se poi $\mu \in \mathbb{K}$,

$$\begin{aligned}
M_{\mathcal{D}}^{\mathcal{B}}(\mu f) &= \Big([(\mu f)(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [(\mu f)(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \Big([\mu f(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [\mu f(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \Big(\mu[f(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| \mu[f(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \mu\Big([f(\underline{v}_1)]_{\mathcal{D}} \Big| \cdots \Big| [f(\underline{v}_n)]_{\mathcal{D}}\Big) \\
&= \mu M_{\mathcal{D}}^{\mathcal{B}}(f)
\end{aligned}$$

$M_{\mathcal{D}}^{\mathcal{B}}$ è surgettiva. Infatti, data $M \in M(m, n, \mathbb{K})$, allora l'applicazione lineare $f = [\ ]_{\mathcal{D}}^{-1} \circ L_M \circ [\ ]_{\mathcal{B}} \in \mathrm{Hom}(V, W)$ è tale che $M_{\mathcal{D}}^{\mathcal{B}}(f) = M$.
$M_{\mathcal{D}}^{\mathcal{B}}$ è iniettiva. Infatti, se $f, g \in Hom(V, W)$ sono tali che $M_{\mathcal{D}}^{\mathcal{B}}(f) = M_{\mathcal{D}}^{\mathcal{B}}(g)$, allora $[f(\underline{v}_i)]_{\mathcal{D}} = [g(\underline{v}_i)]_{\mathcal{D}}$ per ogni i, e quindi $f(\underline{v}_i) = g(\underline{v}_i)$ per ogni i. Poiché f e g hanno gli stessi valori su una base di V, sono uguali.
Riassumiamo nella seguente proposizione.

Proposizione $M_{\mathcal{D}}^{\mathcal{B}} : Hom(V, W) \to M(m, n, \mathbb{K})$ è un isomorfismo.

Abbiamo quindi $\dim \mathrm{Hom}(V, W) = \dim V \dim W$.

Se $f : V \to W$ e $g : W \to Z$ sono lineari, e fissiamo basi $\mathcal{B}, \mathcal{D}, \mathcal{C}$ di V, W, Z rispettivamente, allora

$$M_{\mathcal{C}}^{\mathcal{B}}(g \circ f) = M_{\mathcal{C}}^{\mathcal{D}}(g) M_{\mathcal{D}}^{\mathcal{B}}(f)$$

Infatti, $M_{\mathcal{C}}^{\mathcal{B}}(g \circ f)$ è la matrice che rappresenta $[\ ]_{\mathcal{C}} \circ (g \circ f) \circ [\ ]_{\mathcal{B}}^{-1}$, e abbiamo

$$\begin{aligned} [\ ]_{\mathcal{C}} \circ (g \circ f) \circ [\ ]_{\mathcal{B}}^{-1} &= [\ ]_{\mathcal{C}} \circ (g \circ ([\ ]_{\mathcal{D}}^{-1} \circ [\ ]_{\mathcal{D}}) \circ f) \circ [\ ]_{\mathcal{B}}^{-1} \\ &= ([\ ]_{\mathcal{C}} \circ g \circ [\ ]_{\mathcal{D}}^{-1}) \circ ([\ ]_{\mathcal{D}} \circ f \circ [\ ]_{\mathcal{B}}^{-1}) \\ &= L_{M_{\mathcal{C}}^{\mathcal{D}}(g)} \circ L_{M_{\mathcal{D}}^{\mathcal{B}}(f)} \\ &= L_{M_{\mathcal{C}}^{\mathcal{D}}(g) M_{\mathcal{D}}^{\mathcal{B}}(f)}. \end{aligned}$$

Consideriamo il caso particolare di $id_V : V \to V$.
La matrice associata a id_V nelle basi $\mathcal{B}$ e $\mathcal{B}'$ di V, $M_{\mathcal{B}'}^{\mathcal{B}}(id_V) = \left([\underline{v}_1]_{\mathcal{B}'} \middle| \cdots \middle| [\underline{v}_n]_{\mathcal{B}'}\right)$, è la *matrice del cambio di base* o *del cambio di coordinate* da $\mathcal{B}$ a $\mathcal{B}'$, poiché lega le coordinate di un vettore nella base $\mathcal{B}'$ alle sue coordinate nella base $\mathcal{B}$: per ogni $\underline{v} \in V$

$$[\underline{v}]_{\mathcal{B}'} = M_{\mathcal{B}'}^{\mathcal{B}}(id_V)[\underline{v}]_{\mathcal{B}}.$$

Se adesso abbiamo $f : V \to W$ lineare e due basi $\mathcal{D}$, $\mathcal{D}'$ di W, osservando che $f = id_W \circ f \circ id_V$ e passando alle matrici associate otteniamo

$$M_{\mathcal{D}'}^{\mathcal{B}'}(f) = M_{\mathcal{D}'}^{\mathcal{D}}(id_W) M_{\mathcal{D}}^{\mathcal{B}}(f) M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$$

che esprime come cambia la matrice associata a f quando si cambiano le basi.

Osservazioni:
▶ $M_{\mathcal{B}}^{\mathcal{B}}(id_V) = I_n$.
▶ $M_{\mathcal{B}}^{\mathcal{B}'}(id_V) M_{\mathcal{B}'}^{\mathcal{B}}(id_V) = M_{\mathcal{B}'}^{\mathcal{B}}(id_V) M_{\mathcal{B}}^{\mathcal{B}'}(id_V) = I_n$, quindi le matrici di cambio di base sono invertibili.
▶ Più in generale, se $f : V \to W$ è un isomorfismo, da $f \circ f^{-1} = id_W$, $f^{-1} \circ f = id_V$, abbiamo $M_{\mathcal{D}}^{\mathcal{B}}(f) M_{\mathcal{B}}^{\mathcal{D}}(f^{-1}) = I_n$, $M_{\mathcal{B}}^{\mathcal{D}}(f^{-1}) M_{\mathcal{D}}^{\mathcal{B}}(f) = I_n$, quindi le matrici associate ad un isomorfismo sono invertibili.

Nel caso degli endomorfismi, se $\dim V = n$ e $\mathcal{B}$ è una fissata base di V, allora $M_{\mathcal{B}}^{\mathcal{B}} : End(V) \to M(n, \mathbb{K})$ è un isomorfismo di anelli e si restringe a $M_{\mathcal{B}}^{\mathcal{B}} : GL(V) \to GL(n, \mathbb{K})$ isomorfismo di gruppi. In particolare, per ogni $A \in GL(n, \mathbb{K})$ esiste un unico isomorfismo f di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f) = A$.
Mostriamo che fissata $A \in GL(n, \mathbb{K})$, esiste un'unica base $\mathcal{B}'$ di V tale che $A = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$.
Infatti, siano $A^1, \ldots, A^n$ le colonne di A, e per ogni $i = 1\_n$ sia $\underline{w}_i = [\ ]_{\mathcal{B}}^{-1}(A^i)$. Osserviamo che $\underline{w}_1, \ldots, \underline{w}_n$ sono linearmente indipendenti, infatti se abbiamo $\underline{0} = \sum_{i=1}^n a_i \underline{w}_i$, passando alle coordinate nella base $\mathcal{B}$,

$$\underline{0} = \sum_{i=1}^n a_i [\underline{w}_i]_{\mathcal{B}} = \sum_{i=1}^n a_i A^i,$$

e le colonne della matrice invertibile A sono linearmente indipendenti, quindi tutti i coefficienti sono nulli.
Allora $\mathcal{B}' = \{\underline{w}_1, \dots, \underline{w}_n\}$, che contiene n vettori, è una base di V, e $M_{\mathcal{B}}^{\mathcal{B}'}(id_V) = \left([\underline{w}_1]_{\mathcal{B}} \middle| \cdots \middle| [\underline{w}_n]_{\mathcal{B}}\right) = \left(A^1 \middle| \cdots \middle| A^n\right) = A$. L'unicità deriva dal fatto che i $\underline{w}_i$ sono univocamente determinati da A.
Applicando lo stesso ragionamento a A^{-1}, otteniamo anche che esiste un'unica base $\mathcal{B}''$ di V tale che $A = M_{\mathcal{B}''}^{\mathcal{B}}(id_V)$.

Fissata una base $\mathcal{B}$ di V, abbiamo dunque una doppia interpretazione di una matrice quadrata invertibie: matrice associata ad un isomorfismo di V letto nel sistema di coordinate dato da $\mathcal{B}$; oppure matrice del cambio di coordinate, essendo liberi di specificare il ruolo di $\mathcal{B}$ come base iniziale o finale. Nel primo caso, possiamo pensare che lo spazio V si "muove" e lo stiamo osservando da un sistema di riferimento fisso; nel secondo, lo spazio non si "muove" e stiamo cambiando il sistema di riferimento.

Formule sulle Dimensioni

Proposizione
$f : V \to W$ lineare, $\dim V = n$, allora

$$\dim V = \dim \operatorname{Ker} f + \dim \operatorname{Im} f,$$

detta *formula delle dimensioni* per f.

Dimostrazione
$\operatorname{Ker} f \subset V$ è finitamente generato, e sia $k = \dim \operatorname{Ker} f$.
Sia $\mathcal{D} = \{\underline{u}_1, \dots, \underline{u}_k\}$ una base di $\operatorname{Ker} f$. $\mathcal{D}$ è linearmente indipendente e possiamo completare $\mathcal{D}$ a base $\mathcal{B} = \{\underline{u}_1, \dots, \underline{u}_k, \underline{v}_{k+1}, \dots, \underline{v}_n\}$ di V.
Basta allora dimostrare che $\mathcal{C} = \{f(\underline{v}_{k+1}), \dots, f(\underline{v}_n)\}$ è una base di $\operatorname{Im} f$.
Sia $\underline{w} \in \operatorname{Im} f$, allora esiste $\underline{v} \in V$ tale che $f(\underline{v}) = \underline{w}$. Scriviamo $\underline{v}$ come combinazione lineare della base $\mathcal{B}$, $\underline{v} = \sum_{i=1}^{k} a_i \underline{u}_i + \sum_{i=k+1}^{n} a_i \underline{v}_i$, $a_i \in \mathbb{K}$. Allora
$\underline{w} = f(\underline{v}) = f(\sum_{i=1}^{k} a_i \underline{u}_i + \sum_{i=k+1}^{n} a_i \underline{v}_i) = \sum_{i=1}^{k} a_i f(\underline{u}_i) + \sum_{i=k+1}^{n} a_i f(\underline{v}_i) = \sum_{i=k+1}^{n} a_i f(\underline{v}_i)$.
Quindi, $\mathcal{C}$ genera $\operatorname{Im} f$.
Se poi $a_i \in \mathbb{K}$ sono tali che $\underline{0} = \sum_{i=k+1}^{n} a_i f(\underline{v}_i)$ è una combinazione lineare nulla di elementi di $\mathcal{C}$, $\underline{0} = f(\sum_{i=k+1}^{n} a_i \underline{v}_i)$, ovvero $\sum_{i=k+1}^{n} a_i \underline{v}_i \in \operatorname{Ker} f$. Allora $\sum_{i=k+1}^{n} a_i \underline{v}_i = \sum_{j=1}^{k} b_j \underline{u}_j$, $b_j \in \mathbb{K}$, da cui $\underline{0} = \sum_{j=1}^{k} -b_j \underline{u}_j + \sum_{i=k+1}^{n} a_i \underline{v}_i$ è una combinazione lineare nulla di elementi di $\mathcal{B}$, che è una base, quindi tutti gli a_i (e tutti i b_j) sono nulli.
Quindi, $\mathcal{C}$ è linearmente indipendente. ∎

Osservazioni:
▶ Nella dimostrazione abbiamo di fatto dato un modo per costruire una base di $\operatorname{Im} f$.

Proposizione
$f : V \to W$ lineare, $\dim V = \dim W = n$, allora f è iniettiva se e solo se f è surgettiva.

Dimostrazione
f iniettiva $\iff \operatorname{Ker} f = \{\underline{0}\} \iff \dim \operatorname{Ker} f = 0 \iff \dim \operatorname{Im} f = \dim V = \dim W \iff \operatorname{Im} f = W \iff f$ surgettiva. ∎

Se $f : V \to W$ è lineare e $U \subset V$ è un sottospazio, considerando la restrizione di f a U, $f_{|_U} : U \to W$, abbiamo $\dim f(U) = \dim U - \dim(U \cap \operatorname{Ker} f)$. In particolare, f iniettiva $\Rightarrow \dim f(U) = \dim U$ per ogni sottospazio U di V.

Dimostriamo adesso la *formula di Grassmann.*

Proposizione
Sia V uno spazio vettoriale su $\mathbb{K}$ e siano $U, W \subset V$ due sottospazi di dimensione

finita. Allora

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

Dimostrazione
Poniamo $m = \dim U$, $n = \dim W$. $U \cap W \subset U$ è finitamente generato, e sia $k = \dim U \cap W$.
Sia $\mathcal{D} = \{\underline{z}_1, \dots, \underline{z}_k\}$ una base di $U \cap W$ e completiamo tale base sia a base $\mathcal{B} = \{\underline{z}_1, \dots, \underline{z}_k, \underline{u}_{k+1}, \dots, \underline{u}_m\}$ di U, sia a base $\mathcal{C} = \{\underline{z}_1, \dots, \underline{z}_k, \underline{w}_{k+1}, \dots, \underline{w}_n\}$ di W.
Basta allora dimostrare che $\mathcal{E} = \{\underline{z}_1, \dots, \underline{z}_k, \underline{u}_{k+1}, \dots, \underline{u}_m, \underline{w}_{k+1}, \dots, \underline{w}_n\}$ è una base di $U+W$ (infatti contiene $m+n-k$ elementi).
Se $\underline{v} \in U+W$, scriviamo $\underline{v} = \underline{u} + \underline{w}$ con $\underline{u} \in U$ e $\underline{w} \in W$. Scriviamo poi $\underline{u}$ come combinazione lineare della base $\mathcal{B}$, $\underline{u} = \sum_{i=1}^{k} a_i \underline{z}_i + \sum_{i=k+1}^{m} a_i \underline{u}_i$ e scriviamo $\underline{w}$ come combinazione lineare della base $\mathcal{C}$, $\underline{w} = \sum_{i=1}^{k} b_i \underline{z}_i + \sum_{i=k+1}^{n} b_i \underline{w}_i$, $a_i, b_i \in \mathbb{K}$.
Allora $\underline{v} = \sum_{i=1}^{k} (a_i + b_i)\underline{z}_i + \sum_{i=k+1}^{m} a_i \underline{u}_i + \sum_{i=k+1}^{n} b_i \underline{w}_i$ è combinazione lineare di elementi di $\mathcal{E}$. Quindi, $\mathcal{E}$ genera $U+W$.
Se poi $a_i, b_i, c_i \in \mathbb{K}$ sono tali che $\underline{0} = \sum_{i=1}^{k} a_i \underline{z}_i + \sum_{i=k+1}^{m} b_i \underline{u}_i + \sum_{i=k+1}^{n} c_i \underline{w}_i$ è una combinazione lineare nulla di elementi di $\mathcal{E}$, $U \ni \sum_{i=1}^{k} a_i \underline{z}_i + \sum_{i=k+1}^{m} b_i \underline{u}_i = \sum_{i=k+1}^{n} -c_i \underline{w}_i \in W$, è un elemeno di $U \cap W$, per cui $\sum_{i=k+1}^{n} -c_i \underline{w}_i = \sum_{i=1}^{k} d_i \underline{z}_i$, $d_i \in \mathbb{K}$. Allora $\underline{0} = \sum_{i=1}^{k} d_i \underline{z}_i + \sum_{i=k+1}^{n} c_i \underline{w}_i$ è una combinazione lineare nulla di elementi di $\mathcal{C}$, che è una base, quindi tutti i c_i (e tutti i d_i) sono nulli. Quindi $\sum_{i=1}^{k} a_i \underline{z}_i + \sum_{i=k+1}^{m} b_i \underline{u}_i = \underline{0}$ è una combinazione lineare nulla di elementi di $\mathcal{B}$, che è una base, quindi tutti gli a_i e tutti i b_i sono nulli. Per cui, $\mathcal{E}$ è linearmente indipendente. ❑

Osservazioni:
► Nella dimostrazione abbiamo di fatto dato un modo per costruire una base di $U+W$.
► Se la somma tra U e W è diretta, allora $\dim U \oplus W = \dim U + \dim W$ e l'unione di una base di U e di una base di W dà una base di $U \oplus W$.
► Se quindi U è un sottospazio supplementare del sottospazio W, $V = W \oplus U$, allora $\dim U = \dim V - \dim W$.

Sia V uno spazio vettoriale di dimensione finita n, e sia $W \subset V$ un sottospazio di dimensione m. Consideriamo lo spazio quoziente V/W e la proiezione al quoziente $\pi : V \to V/W$ (che ricordiamo essere lineare).
Poiché $\operatorname{Im}\pi = V/W$ e $\operatorname{Ker}\pi = W$, abbiamo $\dim V/W = n - m$.

Se $\{v_1, \dots, v_n\}$ è una base di V e $\{w_1, \dots, w_m\}$ è una base di W, allora $\{(v_1, 0), \dots, (v_n, 0), (0, w_1), \dots, (0, w_m\}$ è una base di $V \times W$, per cui $\dim V \times W = \dim V + \dim W$.

Se U, W sono sottospazi di V, consideriamo $F : U \times W \to U+W$, $(\underline{u}, \underline{w}) \mapsto \underline{u} + \underline{w}$. F è lineare e iniettiva e il suo nucleo è dato dalle coppie $(\underline{u}, \underline{w}) \in U+W$ tali

che $\underline{u} = -\underline{w}$, ovvero $\operatorname{Ker} F = \{(\underline{v}, -\underline{v}) |\ \underline{v} \in U \cap W\}$ che è isomorfo a $U \cap W$.
Dalla formula delle dimensioni per F segue la formula di Grassmann.
Viceversa, anche la formula delle dimensioni segue dalla formula di Grassmann. Infatti, sia $f : V \to W$ lineare e consideriamo in $V \times \operatorname{Im} f$ il grafico di f, $\Gamma(f) = \{(\underline{v}, f(\underline{v})) |\ \underline{v} \in V\}$, che è un sottospazio isomorfo a V.
Osserviamo che $Z = V \times \{\underline{0}\}$ è un sottospazio e $V \times \operatorname{Im} f = Z + \Gamma(f)$. Infatti ogni $(\underline{v}, f(\underline{w})) = (\underline{v} - \underline{w}, 0) + (\underline{w}, f(\underline{w}))$. Inoltre, $Z \cap \Gamma(f) = \{(\underline{v}, \underline{0}) |\ f(\underline{v}) = \underline{0}\}$ è isomorfo a $\operatorname{Ker} f$. Applicando la formula di Grassmann ai sottospazi Z e $\Gamma(f)$, si ottiene la formula delle dimensioni per f.

L'equivalenza Destra-Sinistra

Siano V e W due spazi vettoriali di dimensione finita n e m rispettivamente. Definiamo su $\mathrm{Hom}(V,W)$ la relazione *destra-sinistra*: $f_1 \sim_{DS} f_2$ se esistono $F \in GL(V)$ e $G \in GL(W)$ tali che $f_2 = G \circ f_1 \circ F$.
Mostriamo che $\sim_{DS}$ è una relazione di equivalenza: Poiché $id_V \in GL(V)$ e $id_W \in GL(W)$, $f = id_W \circ f \circ id_V$ mostra che è riflessiva.
Poiché se $F \in GL(V)$, $G \in GL(W)$ allora $F^{-1} \in GL(V)$, $G^{-1} \in GL(W)$, $f_2 = G \circ f_1 \circ F \iff f_1 = G^{-1} \circ f_2 \circ F^{-1}$ mostra che è simmetrica.
Poiché se $F, H \in GL(V)$, $G, L \in GL(W)$ allora $F \circ H \in GL(V)$, $L \circ G \in GL(W)$, $f_2 = G \circ f_1 \circ F$, $f_3 = L \circ f_2 \circ H \Rightarrow f_3 = (L \circ G) \circ f_1 \circ (F \circ H)$ mostra che è transitiva.

Osservazioni:
▶ Consideriamo $GL(V) \times GL(W)$ come gruppo di trasformazioni di $\mathrm{Hom}(V,W)$, identificando la coppia (F,G) con la trasformazione

$$\tau_{F,G} : \mathrm{Hom}(V,W) \to \mathrm{Hom}(V,W), \ f \mapsto G \circ f \circ F.$$

La relazione destra-sinistra è la relazione indotta da tale gruppo.

Analogamente, definiamo la relazione destra-sinistra per matrici.
Se $A, B \in M(m,n,\mathbb{K})$, diciamo che A è equivalente destra-sinistra a B e scriviamo $A \sim_{DS} B$ se esistono $M \in GL(m,\mathbb{K})$, $N \in GL(n,\mathbb{K})$ tali che $B = MAN$.
Le verifiche sono analoghe a quanto visto sopra.

Abbiamo diverse formulazioni equivalenti dell'equivalenza destra-sinistra:

Proposizione
Date $f_1, f_2 \in \mathrm{Hom}(V,W)$, i seguenti fatti sono equivalenti:

1. $f_1 \sim_{DS} f_2$.
2. Per ogni $\mathcal{B}$ base di V e $\mathcal{D}$ base di W, $M_{\mathcal{D}}^{\mathcal{B}}(f_1) \sim_{DS} M_{\mathcal{D}}^{\mathcal{B}}(f_2)$.
3. Esistono $\mathcal{B}$ base di V e $\mathcal{D}$ base di W tali che $M_{\mathcal{D}}^{\mathcal{B}}(f_1) \sim_{DS} M_{\mathcal{D}}^{\mathcal{B}}(f_2)$.
4. Esistono $\mathcal{B}, \mathcal{B}'$ basi di V e $\mathcal{D}, \mathcal{D}'$ basi di W tali che $M_{\mathcal{D}}^{\mathcal{B}}(f_1) = M_{\mathcal{D}'}^{\mathcal{B}'}(f_2)$.

Dimostrazione

$1 \Rightarrow 2$. Se $f_2 = G \circ f_1 \circ F$ con $F \in GL(V)$ e $G \in GL(W)$, allora le matrici associate a F e G sono invertibili e $M_{\mathcal{D}}^{\mathcal{B}}(f_2) = M_{\mathcal{D}}^{\mathcal{D}}(G) M_{\mathcal{D}}^{\mathcal{B}}(f_1) M_{\mathcal{B}}^{\mathcal{B}}(F)$.

$2 \Rightarrow 3$. Ovvio.

$3 \Rightarrow 1$. Se $M_{\mathcal{D}}^{\mathcal{B}}(f_2) = M M_{\mathcal{D}}^{\mathcal{B}}(f_1) N$ con N, M invertibili, allora esistono unici $F \in GL(V)$ e $G \in GL(W)$ tali che $M_{\mathcal{B}}^{\mathcal{B}}(F) = N$ e $M_{\mathcal{D}}^{\mathcal{D}}(G) = M$. Si ha quindi

$$M_{\mathcal{D}}^{\mathcal{B}}(f_2) = M_{\mathcal{D}}^{\mathcal{D}}(G) M_{\mathcal{D}}^{\mathcal{B}}(f_1) M_{\mathcal{B}}^{\mathcal{B}}(F) = M_{\mathcal{D}}^{\mathcal{B}}(G \circ f_1 \circ F)$$

per cui $f_2 = G \circ f_1 \circ F$.

$3 \Rightarrow 4$. Se $M_{\mathcal{D}}^{\mathcal{B}}(f_2) = M M_{\mathcal{D}}^{\mathcal{B}}(f_1) N$ con N, M invertibili, allora esistono uniche $\mathcal{B}'$ base di V e $\mathcal{D}'$ base di W tali che $N = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$ e $M = M_{\mathcal{D}'}^{\mathcal{D}}(id_W)$. Allora,

$$M_{\mathcal{D}}^{\mathcal{B}}(f_2) = M_{\mathcal{D}'}^{\mathcal{D}}(id_W) M_{\mathcal{D}}^{\mathcal{B}}(f_1) M_{\mathcal{B}}^{\mathcal{B}'}(id_V) = M_{\mathcal{D}'}^{\mathcal{B}'}(id_W \circ f_1 \circ id_V) = M_{\mathcal{D}'}^{\mathcal{B}'}(f_1).$$

$4 \Rightarrow 3$. Se $M_{\mathcal{D}}^{\mathcal{B}}(f_1) = M_{\mathcal{D}'}^{\mathcal{B}'}(f_2)$ allora

$$M_{\mathcal{D}}^{\mathcal{B}}(f_1) = M_{\mathcal{D}'}^{\mathcal{D}}(id_W) M_{\mathcal{D}}^{\mathcal{B}}(f_2) M_{\mathcal{B}}^{\mathcal{B}'}(id_V).$$

Osservazioni:
▶ La dimensione dell'immagine è un *invariante* per la relazione destra-sinistra, cioè $f_1 \sim_{DS} f_2 \Rightarrow \dim \operatorname{Im} f_1 = \dim \operatorname{Im} f_2$.
Infatti, se abbiamo $f_2 = G \circ f_1 \circ F$ con F e G isomorfismi, otteniamo allora $\operatorname{Im} f_2 = \operatorname{Im}(G \circ f_1 \circ F) = G(f_1(F(V))) = G(f_1(V))$, poiché essendo F un isomorfismo, F è surgettiva. Poiché G è iniettiva, mantiene le dimensioni dei sottospazi, quindi $\dim \operatorname{Im} f_2 = \dim G(f_1(V)) = \dim f_1(V) = \dim \operatorname{Im} f_1$.
▶ Nel caso matriciale, data $A \in M(m, n, \mathbb{K})$, l'immagine di A è l'immagine di $L_A : \mathbb{K}^n \to \mathbb{K}^m$, che è generata dalle colonne di A.

Si definisce il *rango* di A come $\operatorname{rnk} A = \dim \operatorname{Im} A = \dim \operatorname{Span}(A^{\overset{1}{|}}, \ldots, A^{\overset{n}{|}})$.
Si ha quindi $A \sim_{DS} B \Rightarrow \operatorname{rnk} A = \operatorname{rnk} B$.

Mostriamo adesso che la dimensione dell'immagine è un invariante *completo* per la relazione destra-sinistra, cioè $f_1 \sim_{DS} f_2$ se e solo se $\dim \operatorname{Im} f_1 = \dim \operatorname{Im} f_2$. Ovvero, data l'applicazione $\operatorname{Hom}(V, W) \to \mathbb{N}$, $f \mapsto \dim \operatorname{Im} f$, la relazione destra-sinsitra è la relazione di equivalenza data da tale applicazione.
Nel caso matriciale, avremo che $A \sim_{DS} B$ se e solo se $\operatorname{rnk} A = \operatorname{rnk} B$, e quindi che data l'applicazione $\operatorname{rnk} : M(m, n, \mathbb{K}) \to \mathbb{N}$, $A \mapsto \operatorname{rnk} A$ la relazione destra-sinsitra è la relazione di equivalenza data da rnk. Notiamo che il rango di una matrice è facilmente calcolabile e quindi abbiamo un modo effettivo per decidere l'equivalenza destra-sinistra.

Proposizione
Data $f \to W$ lineare, esistono una basa $\mathcal{B}$ di V e una base $\mathcal{D}$ di W tali che $M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, dove $r = \dim \operatorname{Im} f$.

Dimostrazione
In modo analogo a quanto visto nella dimostrazione della formula delle dimensioni, costruiamo una base di $\operatorname{Im} f$ estendendo una base $\{\underline{u}_1, \ldots, \underline{u}_{n-r}\}$ di $\operatorname{Ker} f$ a base di V con i vettori $\underline{v}_1, \ldots, \underline{v}_r$ e prendendo le immagini tramite f dei $\underline{v}_i$.
Scegliamo come base di V $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_r, \underline{u}_1, \ldots, \underline{u}_{n-r}\}$ e scegliamo come base di W un completamento $\mathcal{D}$ di $\{f(\underline{v}_1), \ldots, f(\underline{v}_r)\}$.
Allora per ogni $i = 1 \_\_ r$, $[f(\underline{v}_i)]_{\mathcal{D}} = \underline{e}_i$, mentre per ogni $j = 1 \_\_ n - r$, $[f(u_j)]_{\mathcal{D}} = \underline{0}$, per cui $M_{\mathcal{D}}^{\mathcal{B}}(f)$ ha la forma voluta.

Nel caso matriciale, la proposizione diventa:
Per ogni $A \in M(m,n,\mathbb{K})$, esistono $M \in GL(m,\mathbb{K})$, $N \in GL(n,\mathbb{K})$ tali che $MAN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, dove $r = \operatorname{rnk} A$.
Osserviamo che $\operatorname{rnk} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = r$
La matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M(m,n,\mathbb{K})$ si dice *forma nomale di taglia $m \times n$ e rango r per la relazione destra-sinsitra* e si indica con $D(r,m,n)$.

Osserviamo che la forma normale ottenuta nella proposizione, dipende solo dalla taglia del blocco identità in alto a sinistra (oltre che dalle dimensioni di V e W), ovvero dal suo rango $r = \dim \operatorname{Im} f$.
Se quindi $\dim \operatorname{Im} f_1 = \dim \operatorname{Im} f_2$, allora f_1 e f_2 hanno la stesa forma normale in opportune basi di V e di W (possibilmente diverse per f_1 e f_2) e quindi, per la caratterizzazione 4 della relazione destra-sinistra, $f_1 \sim_{DS} f_2$.

Elenchiamo alcune proprietà dell'applicazione trasposta di facile verifica:

- $(AB)^\top = A^\top B^\top$.
- Se A è invertibile, passando alla trasposta in $AA^{-1} = A^{-1}A = I$ otteniamo $A^\top (A^{-1})^\top = (A^{-1})^\top A^\top = I^\top = I$, da cui $A^\top$ è invertibile con inversa $(A^{-1})^\top$.
- Se $D = D(r,m,n) \in M(m,n,\mathbb{K})$ è la forma normale di rango r, allora $D^\top$ è la forma normale $D(r,n,m)$ di rango r in $M(n,m,\mathbb{K})$.

Proposizione
Per ogni $A \in M(m,n,\mathbb{K})$, $\operatorname{rnk} A = \operatorname{rnk} A^\top$.

Dimostrazione
Se $r = \operatorname{rnk} A$, sappiamo che la forma normale di rango r in $M(m,n,\mathbb{K})$ si scrive come $D(r,m,n) = MAN$ con M, N invertibili. Passando alla trasposta, $N^\top A^\top M^\top = D(r,m,n)^\top = D(r,n,m)$, per cui $A^\top \sim_{DS} D(r,n,m)$ e quindi hanno lo stesso rango, $\operatorname{rnk} A^\top = \operatorname{rnk} D(r,n,m) = r$. ■

L'Algoritmo di Gauss

Descriviamo adesso l'algoritmo di Gauss *rispetto alle righe*, che permetterà di calcolare esplicitamente il rango di una matrice.
L'algoritmo riceve in entrata una matrice $A \in M(m,n,\mathbb{K})$ e dà in uscita una matrice $G_R(A) \in M(m,n,\mathbb{K})$, passando attraverso m modifiche operate su A, $A_0 = A \to A_1 \to A_2 \to \cdots \to A_m = G_R(A)$, dove $A_j \in M(m,n,\mathbb{K})$ per ogni j.

La matrice in uscita ha una forma speciale, è *a gradini rispetto alle righe*:
una matrice $M \in M(m,n,\mathbb{K})$ si dice a gradini rispetto alle righe se

- Esiste un $0 \le r \le m$ tale che $\underset{i\text{—}}{M} = 0$ se e solo se $i > r$. Ovvero, le prime r righe di M sono non nulle e le rimanenti sono nulle.
- Per ogni $i = 1\_r$, esiste $1 \le k(i) \le n$ tale che $\underset{i\text{—}}{M}\overset{j}{|} = 0$ se $j < k(i)$ ed inoltre $\underset{i\text{—}}{M}\overset{k(i)}{|} = 1$. Ovvero, il primo coefficiente non nullo nella riga $\underset{i\text{—}}{M}$, detto *pivot* della riga, vale 1 e si trova nel posto $(i,k(i))$.
- La succesione $(k(1),k(2),\dots,k(r))$ degli indici di colonna dei pivot delle r righe non nulle di M è strettamente crescente.

Ad esempio, la seguente matrice 5×6 è a gradini rispetto alle righe con $r = 3$ e succesisone degli indici di colonna dei pivot $(2,3,5)$:

$$\left(\begin{array}{c|cccccc} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & \multicolumn{4}{|c}{1 \quad 2 \quad 3 \quad 4} \\ 0 & 0 & 0 & 0 & \multicolumn{2}{|c}{1 \quad 2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Otteniamo quindi una applicazione $G_R : M(m,n,\mathbb{K}) \to M(m,n,\mathbb{K})$ con immagine le matrici a gradini e vedremo che se M è a gradini, $G_R(M) = M$, e quindi $G_R^2 = G_R$.

L'algoritmo usa tre tipi di *operazioni elementari sulle righe*:

1. scambiare due righe, $R_i \leftrightarrow R_j$ (se $i = j$ la matrice non cambia);
2. moltiplicare una riga per uno scalare λ non nullo, $R_i \to \lambda R_i$ (se $\lambda = 1$ la matrice non cambia);
3. sommare ad una riga un multiplo di un'altra, $R_i \to R_i + cR_j$, $j \neq i$ (se $c = 0$ la matrice non cambia).

Descriviamo adesso l'agoritmo.
Data $A_0 = A$ applichiamo la seguente procedura (che chiameremo "passo base") per produrre A_1.

Se $A_0 = 0$, allora $A_1 = A_0$.
Altrimenti:
- sia j il minimo indice per cui la colonna $A^{j}| \neq 0$;
- sia i il minimo indice per cui $\underset{i}{A^{j}|} \neq 0$;
- eseguiamo su A_0 l'operazione elementare di tipo 1 $R_i \leftrightarrow R_1$, ovvero portiamo la riga i al primo posto. Otteniamo così una matrice A'_0 con le prime $j-1$ colonne nulle e il primo coefficiente λ della colonna j è non nullo;
- eseguiamo su A'_0 l'operazione elementare di tipo 2 $R_1 \to \lambda^{-1}R_1$, ovvero, dividiamo la prima riga per λ. Otteniamo così una matrice A''_0 con le prime $j-1$ colonne nulle e il primo coefficiente della colonna j è uguale a 1,
- per ogni $i = 2\_m$, eseguiamo su A''_0 l'operazione elementare di tipo 3 $R_i \to R_i - \underset{i}{A''^{j}_0|}R_1$, ovvero alla riga i sottraiamo la prima riga moltiplicata per il coefficiente j-mo della riga i. Otteniamo così una matrice A_1 con le prime $j-1$ colonne nulle, il primo coefficiente della colonna j è 1 e gli altri coefficienti della colonna j sono nulli, ovvero la colonna j è $\underline{e}_1 \in \mathbb{K}^m$.
Vediamo un esempio di passo base:

$$A_0 = \begin{pmatrix} 0 & 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 6 & 8 & 0 & 2 \\ 0 & 1 & 1 & -1 & -2 & -3 \\ 0 & 3 & 1 & 0 & 5 & 6 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_1} \begin{pmatrix} 0 & 2 & 6 & 8 & 0 & 2 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & 4 & 5 \\ 0 & 1 & 1 & -1 & -2 & -3 \\ 0 & 3 & 1 & 0 & 5 & 6 \end{pmatrix} \xrightarrow{R_1 \to \frac{1}{2}R_1}$$

$$\begin{pmatrix} 0 & 1 & 3 & 4 & 0 & 1 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & 4 & 5 \\ 0 & 1 & 1 & -1 & -2 & -3 \\ 0 & 3 & 1 & 0 & 5 & 6 \end{pmatrix} \xrightarrow[\substack{R_3 \to R_3 - 0R_1 \\ R_4 \to R_4 - 1R_1 \\ R_5 \to R_5 - 3R_1}]{\substack{R_2 \to R_2 - 0R_1}} \begin{pmatrix} 0 & 1 & 3 & 4 & 0 & 1 \\ 0 & 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & -2 & -5 & -2 & -4 \\ 0 & 0 & -8 & -13 & 5 & 3 \end{pmatrix} = A_1$$

Se abbiamo costruito A_i, per ottenere A_{i+1}:
- sia A'_i la matrice di taglia $(m-i) \times n$ ottenuta da A_i eliminando le prime i righe;
- applichiamo il "passo base" a A'_i ed otteniamo A''_i;
- aggiungiamo in alto alla matrice A''_i le prime i righe della matrice A_i per ottenere A_{i+1}.
Iterando m volte otteniamo $A_m = G_R(A)$.
Osserviamo che possiamo interrompere l'algoritmo non appena otteniamo una $A'_i = 0$.
È chiaro che $G_R(A)$ è a gradini rispetto alle righe, ed è chiaro che se A è già a gradini rispetto alle righe, allora $G_R(A) = A$.

L'algoritmo di Gauss si può ulteriormente prolungare eseguendo altre operazioni di tipo 3 su $G_R(A)$, in modo da ottenere una matrice ancora a gradini e con i pivot nelle stesse posizioni di $G_R(A)$, ma che in ogni colonna con un pivot, il pivot sia l'unico coefficiente non nullo (ovvero, la colonna che contiene l'i-mo

pivot è $\underline{e}_i \in \mathbb{K}^m$; una tale matrice si dice a gradini *ridotta* rispetto alle righe): basta operare come nell'algoritmo usuale usando una riga che contiene un pivot sulle righe precedenti alla riga che contiene il pivot (invece che sulle successive come nell'algoritmo usuale). Tale matrice si indica con $\tilde{G}_R(A)$ e l'algoritmo esteso si dice *algoritmo di Gauss completo sulle righe*.

Scambiando il ruolo delle righe e delle colonne, possiamo definire anche un algoritmo di Gauss sulle colonne G_C, usando le operazioni elementari sulle colonne, che produce una matrice a gradini rispetto alle colonne (trasposta di una matrice a scalini rispetto alle righe). A tutti gli effetti, applicare l'algoritmo sulle colonne è la stessa cosa che applicare l'algoritmo sulle righe di $A^\top$ e poi trasporre il risultato: $G_C(A) = (G_R(A^\top))^\top$. Lo stesso vale per l'algoritmo completo sulle colonne $\tilde{G}_C$.

Per $A \in M(m,n,\mathbb{K})$, indichiamo con $R(A)$ il sottospazio di $M(1,n,\mathbb{K})$ generato dalle righe di A, e con $C(A)$ il sottospazio di $\mathbb{K}^m$ generato dalle colonne di A: $R(A) = \operatorname{Span}(\underset{1}{\underline{A}}, \dots, \underset{m}{\underline{A}})$, $C(A) = \operatorname{Span}(\overset{1}{A|}, \dots, \overset{n}{A|}) = \operatorname{Im} A$.
Mostriamo che se A' è ottenuta da A eseguendo un'operazione elementare per riga, allora $R(A') = R(A)$; analogamente, che se A'' è ottenuta da A eseguendo un'operazione elementare per colonna, allora $C(A'') = C(A)$.
Vediamolo nel caso delle righe, la dimostrazione nel caso delle colonne è completamente analoga.
Eseguendo un'operazione di tipo 1, l'isieme delle righe non cambia, e quindi non cambia il sottospazio generato.
Eseguendo un'operazione di tipo 2 si passa da $R = \{\underset{1}{\underline{A}}, \dots, \underset{i}{\underline{A}}, \dots, \underset{m}{\underline{A}}\}$ a $R' = \{\underset{1}{\underline{A}}, \dots, \underset{i}{\underline{\lambda A}}, \dots, \underset{m}{\underline{A}}\}$. Poiché ogni elemento di R' è combinazione lineare di R, $\operatorname{Span}(R') \subset \operatorname{Span}(R)$. Scrivendo $\underset{i}{\underline{A}} = \lambda^{-1}(\underset{i}{\underline{\lambda A}})$, anche ogni elemento di R è combinazione lineare di R', per cui $\operatorname{Span}(R) \subset \operatorname{Span}(R')$. Quindi $R(A') = R(A)$.
Eseguendo un'operazione di tipo 3 si passa da $R = \{\underset{1}{\underline{A}}, \dots, \underset{i}{\underline{A}}, \dots, \underset{m}{\underline{A}}\}$ a $R' = \{\underset{1}{\underline{A}}, \dots, \underset{i}{\underline{A + cA}}, \dots, \underset{m}{\underline{A}}\}$. Poiché ogni elemento di R' è combinazione lineare di R, $\operatorname{Span}(R') \subset \operatorname{Span}(R)$. Scrivendo $\underset{i}{\underline{A}} = (\underset{i}{\underline{A}} + \underset{j}{\underline{cA}}) - \underset{j}{\underline{cA}}$, anche ogni elemento di R è combinazione lineare di R', per cui $\operatorname{Span}(R) \subset \operatorname{Span}(R')$. Quindi $R(A') = R(A)$.

Poiché l'algoritmo di Gauss (anche completo) esegue un numero finito di operazioni elementari, abbiamo

$$R(A) = R(G_R(A)) = R(\tilde{G}_R(A)),\ C(A) = C(G_C(A)) = C(\tilde{G}_C(A)).$$

Osserviamo che per una matrice M a gradini rispetto alle righe, le r righe non nulle sono linearmente indipendenti: la combinazione lineare $a_1\underset{1}{\underline{M}} + \cdots + a_r\underset{r}{\underline{M}}$ ha esattamente a_1 nella colonna del primo pivot; quindi se la combinazione lineare è nulla, allora $a_1 = 0$. Ma allora, la combinazione lineare ha esattamente a_2 nella colonna del secondo pivot e quindi $a_2 = 0$. Reitarando, tutti i coefficienti a_i sono nulli. (Osserviamo che per una matrice a gradini ridotta rispetto alle

righe è ancora più evidente, poiché nella colonna dell'i-mo pivot la combinazione lineare ha esattamente a_i). Si ha quindi che $\dim R(M)$ è uguale al numero di pivot di M.
Lo stesso vale per una matrice N a gradini rispetto alle colonne: le colonne non nulle sono linearmente indipendenti e $\dim C(N) = \operatorname{rnk} N$ è uguale al numero di pivot di N.

Quindi le righe non nulle di $G_R(A)$ (e di $\tilde{G}_R(A)$), come pure le colonne non nulle di $G_C(A)$ (e di $\tilde{G}_C(A)$) danno una base di $R(A)$ e di $C(A)$ rispettivamente.
Abbiamo quindi che $\dim R(A)$ è uguale al numero di pivot di $G_R(A)$, e che $\dim C(A) = \operatorname{rnk} A$ è uguale al numero di pivot di $G_C(A)$.

Usando il fatto noto che $\operatorname{rnk} A = \operatorname{rnk} A^\top$, abbiamo

$$\operatorname{rnk} A = \operatorname{rnk} A^\top = \operatorname{rnk} G_C(A^\top) = \operatorname{rnk}(G_R(A))^\top = \operatorname{rnk} G_R(A).$$

Poiché $(G_R(A))^\top$ è a gradini rispetto alle colonne e ha lo stesso numero di pivot di $G_R(A)$, $G_R(A)$ e $G_C(A)$ hanno lo stesso numero di pivot.
Quindi, $\dim C(A) = \dim R(A) = \operatorname{rnk} A$, ovvero, il numero massimo di colonne di A linearmente indipendenti coincide con il numero massimo di righe di A linearmente indipendenti, entrambi uguali a $\operatorname{rnk} A$.

Osservazioni:
► Se $G_R(A)$ ha i pivot nelle colonne $k(1), \dots, k(r)$, allora le colonne di A, $A^{\overset{k(1)}{|}}, \dots, A^{\overset{k(r)}{|}}$ danno una base di $C(A)$. Infatti, eseguendo l'algoritmo di Gauss sulle righe della matrice $B = (A^{\overset{k(1)}{|}}|\cdots|A^{\overset{k(r)}{|}})$ otteniamo $G_R(B) = (G_R(A)^{\overset{k(1)}{|}}|\cdots|G_R(A)^{\overset{k(r)}{|}})$ (con r pivot), per cui $\operatorname{rnk} B = r = \operatorname{rnk} A$. Quindi le colonne di B, che sono r, sono una base di $C(B)$. Inoltre, poiché $C(B) \subset C(A)$ ed hanno la stessa dimensione, $C(B) = C(A)$.

Sia $A\underline{X} = \underline{b}$, dove $A \in M(m, n, \mathbb{K})$ e $\underline{b} \in \mathbb{K}^m$ un sistema lineare di m equazioni in n incognite $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. A si dice *matrice dei coefficienti del sistema*, $\underline{b}$ il *termine noto*, $\underline{X}$ l'*incognita*.
La matrice $M = (A|\underline{b})$ si dice la *matrice completa del sistema*. Il sistema si dice *risolubile* se esiste una *soluzione*, cioè un $\underline{v} \in \mathbb{K}^n$ tale che $A\underline{v} = \underline{b}$.
Se il sistema è *omogeneo*, cioè se $\underline{b} = \underline{0}$, allora l'insieme delle soluzioni è $\operatorname{Ker} A$ ed il sistema è risolubile ($\underline{X} = \underline{0}$ è una soluzione).
In generale, il sistema è risolubile se e solo se $\underline{b} \in \operatorname{Im} A$, ovvero se e solo se $\underline{b}$ è combinazione lineare delle colonne di A. Quindi, se esiste una soluzione le colonne di M appartengono a $\operatorname{Im} A$ e quindi $\operatorname{Im} M \subset \operatorname{Im} A$. Viceversa, se $\operatorname{Im} M \subset \operatorname{Im} A$, $\underline{b} \in \operatorname{Im} M$ appartiene a $\operatorname{Im} A$ ed esiste una soluzione. Poiché $\operatorname{Im} A \subset \operatorname{Im} M$, abbiamo che il sistema è risolubile se e solo se $\operatorname{Im} M = \operatorname{Im} A$ e quindi se e solo se $\dim \operatorname{Im} M = \dim \operatorname{Im} A$.
Abbiamo dunque il seguente criterio di risolubilità:

Criterio di Rouché-Capelli:
il sistema lineare $A\underline{X} = \underline{b}$ è risolubile se e solo se $\operatorname{rnk}(A|\underline{b}) = \operatorname{rnk} A$.

L'algoritmo di Gauss ci dà un modo effettivo per calcolare il rango di una matrice, e quindi un modo effettivo di valutare la risolubilità dei sistemi lineari.

Osserviamo che se facciamo una operazione elementare sulle righe di M, ottenendo $M' = (A'|\underline{b}')$, allora il sistema lineare $A'\underline{X} = \underline{b}'$ è *equivalente* al sistema $A\underline{X} = \underline{b}$, cioè i due sistemi hanno le stesse soluzioni.
Infatti, eseguire una operazione elementare sulle righe di M equivale a eseguire la stessa operazione elementare sulle equazioni del sistema. In particolare, una operazione di tipo 1 scambia di posto due equazioni, e quindi i due sistemi hanno di fatto le stesse equazioni. Una operazione di tipo 2, invece, rimpiazza una equazione $a_1x_1 + \cdots + a_nx_n = b$ con un suo multiplo non nullo $\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b$ ed è chiaro che le soluzioni delle due equazioni sono uguali. Una operazione di tipo 3, cambia una equazione $a_1x_1 + \cdots + a_nx_n = b$ sommandole un multiplo di un'altra equazione $\alpha_1x_1 + \cdots + \alpha_nx_n = \beta$ ottenendo $(a_1 + c\alpha_1)x_1 + \cdots + (a_n + c\alpha_n)x_n = b + c\beta$; è allora chiaro che il sistema di due equazioni

$$\begin{cases} \alpha_1x_1 + \cdots + \alpha_nx_n &= \beta \\ a_1x_1 + \cdots + a_nx_n &= b \end{cases}$$

e il sistema di due equazioni

$$\begin{cases} \alpha_1x_1 + \cdots + \alpha_nx_n &= \beta \\ (a_1 + c\alpha_1)x_1 + \cdots + (a_n + c\alpha_n)x_n &= b + c\beta \end{cases}$$

sono equivalenti.

Supponiamo che il sistema $A\underline{X} = \underline{b}$ sia risolubile e che $\operatorname{rnk} A = r$.
Se eseguiamo sulla matrice M l'algoritmo di Gauss completo rispetto alle righe, ottenendo $(\tilde{G}_R(A)|\underline{b}')$, allora il sistema $\tilde{G}_R(A)\underline{X} = \underline{b}'$ è equivalente al sistema $A\underline{X} = \underline{b}$.
Osserviamo che quest'ultimo sistema è facilmente risolvibile. Infatti, la j-ma riga non nulla (che contiene il j-mo pivot nella colonna $k(j)$) dà un'equazione del tipo $x_{k(j)} + a_{j,k(j)+1}x_{k(j)+1} + \cdots + a_{j,n}x_n = b'_j$ da cui si ricava facilmente $x_{k(j)}$ in funzione delle variabili successive, $x_{k(j)} = b'_j - a_{j,k(j)+1}x_{k(j)+1} - \cdots - a_{j,n}x_n$. Osserviamo però che in tale equazione non compaiono altre variabili corrispondenti a pivot.
Quindi, possiamo ricavare tute le r variabili corrispondenti a pivot in funzione delle altre $n - r$ (che sono libere di variare su tutto $\mathbb{K}$).

La generica soluzione sarà quindi del tipo:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{k(1)-1} \\ x_{k(1)} \\ x_{k(1)+1} \\ \vdots \\ x_{k(2)} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{k(1)-1} \\ b'_1 - a_{1,k(1)+1}x_{k(j)+1} - \cdots - a_{1,n}x_n \\ x_{k(1)+1} \\ \vdots \\ b'_2 - a_{2,k(2)+1}x_{k(2)+1} - \cdots - a_{2,n}x_n \\ \vdots \\ x_n \end{pmatrix} =$$

$$= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b'_1 \\ 0 \\ \vdots \\ b'_2 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} x_1 \\ \vdots \\ x_{k(1)-1} \\ -a_{1,k(1)+1}x_{k(1)+1} - \cdots - a_{1,n}x_n \\ x_{k(1)+1} \\ \vdots \\ -a_{2,k(2)+1}x_{k(2)+1} - \cdots - a_{2,n}x_n \\ \vdots \\ x_n \end{pmatrix}$$

Tralasciando il termine costante, l'altro termine si riscrive, separando i contributi di ogni singola variabile che non corrisponde a un pivot, come:

$$x_1 \begin{pmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_{k(1)-1} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{k(1)+1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -a_{1,k(1)+1} \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_{j(2)+1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -a_{1,j(2)+1} \\ 0 \\ \vdots \\ -a_{2,j(2)+1} \\ \vdots \\ 0 \end{pmatrix}$$

$$+ \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -a_{1,n} \\ 0 \\ \vdots \\ -a_{2,n} \\ \vdots \\ 1 \end{pmatrix}$$

in cui riconosciamo le combinazioni lineari di $n-r$ vettori linearmente indipendenti (trascurando le posizioni dei pivot, abbiamo vettori della base canonica di $\mathbb{K}^{n-r}$). Lo spazio delle soluzioni si ottiene dunque sommando un vettore fissato ai vettori di uno sottospazio vettoriale di $\mathbb{K}^n$ di dimensione $n-r$.

In generale, se $W \subset \mathbb{K}^n$ è un sottoinsieme e $\underline{v} \in \mathbb{K}^n$, il *traslato di W tramite il vettore $\underline{v}$* è dato da $\underline{v} + W = \{\underline{v} + \underline{w} |\ \underline{w} \in W\}$. Osserviamo che $\underline{v} + W$ è in bigezione con W.
Se W è un sottospazio, $\underline{v}+W$ si dice *sottospazio affine di giacitura* W e poniamo $\dim(\underline{v} + W) = \dim W$.

Supponiamo che il sistema $A\underline{X} = \underline{b}$ sia risolubile. Allora le soluzioni del sistema lineare sono un traslato di $\operatorname{Ker} A$.
Infatti, fissiamo una soluzione particolare $\underline{y}$, $A\underline{y} = \underline{b}$, e mostriamo che le soluzioni del sistema sono $\underline{y} + \operatorname{Ker} A$. Se $\underline{v} \in \operatorname{Ker} A$, allora $A(\underline{y}+\underline{v}) = A\underline{y} + A\underline{v} = \underline{b}$ mostra che $\underline{y} + \underline{v}$ è soluzione del sistema. Viceversa, se $\underline{x}$ è soluzione del sistema, $\underline{v} = \underline{x} - \underline{y} \in \operatorname{Ker} A$, infatti $A\underline{v} = A(\underline{x} - \underline{y}) = A\underline{x} - A\underline{y} = \underline{b} - \underline{b} = 0$. E allora $\underline{x} = \underline{y} + \underline{v} \in \underline{y} + \operatorname{Ker} A$.

Riassumendo:

Proposizione
il sistema lineare $A\underline{X} = \underline{b}$ in n incognite è risolubile se e solo se $\operatorname{rnk} A = \operatorname{rnk}(A|\underline{b})$. In tal caso, le soluzioni sono un sottospazio affine di $\mathbb{K}^n$ di giacitura $\operatorname{Ker} A$ e dimensione $n - \operatorname{rnk} A$.

Osservazione:
▶ Vogliamo riottenere per altra via, senza usare quanto già sappiamo dalla vecchia dimostrazione, che $\operatorname{rnk} A^\top = \operatorname{rnk} A$.
Osserviamo che $C(A^\top) = R(A)$, quindi $\operatorname{rnk} A^\top = \dim C(A^\top) = \dim R(A)$ che è uguale al numero di pivot r di $G_R(A)$.
Ora, ragionando come sopra per il sistema lineare omogeneo $A\underline{X} = \underline{0}$, otteniamo che lo spazio delle soluzioni, date da $\operatorname{Ker} A$, ha dimensione $n - r$. Ma dalla formula delle dimensioni, $n = \dim \mathbb{K}^n = \operatorname{rnk} A + \dim \operatorname{Ker} A$, da cui $\operatorname{rnk} A = n - \dim \operatorname{Ker} A = n - (n - r) = r = \operatorname{rnk} A^\top$, come voluto.

Una matrice $E \in M(n, \mathbb{K})$ si dice *elementare*, se si ottiene dalla matrice I_n eseguendo una operazione elementare (sulle righe o sulle colonne).
Osserviamo che ogni matrice elementare ottenuta con una operazione di colonna si ottiene anche con una operazione di riga:
• la matrice $T_{i,j}$ ottenuta da I_n eseguendo $R_i \leftrightarrow R_j$ è uguale alla matrice ottenuta da I_n eseguendo $C_i \leftrightarrow C_j$.

$$T_{i,j} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ j \\ \\ i \\ \\ \\ \end{matrix}$$

Allo stesso modo, la matrice $D_{\lambda,i}$ ottenuta da I_n eseguendo $R_i \to \lambda R_i$ è uguale alla matrice ottenuta da I_n eseguendo $C_i \to \lambda C_i$

$$D_{\lambda,i} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \end{matrix}$$

la matrice $L_{i,c,j}$ ottenuta da I_n eseguendo $R_i \to R_i + cR_j$ è uguale alla matrice ottenuta da I_n eseguendo $C_j \to C_j + cC_i$.

$$L_{i,c,j} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & c & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ j \\ \\ i \end{matrix}$$

È facile verificare che

- $T_{i,j}^2 = I_n$,
- $D_{\lambda,i} D_{\frac{1}{\lambda},i} = D_{\frac{1}{\lambda},i} D_{\lambda,i} = I_n$,
- $L_{i,c,j} L_{i,-c,j} = L_{i,-c,j} L_{i,c,j} = I_n$,

per cui le matrici elementari sono invertibili.

Data $A \in M(m,n,\mathbb{K})$, se B è ottenuta eseguendo una operazione elementare sulle righe di A e E è la matrice elementare ottenuta eseguendo la stessa operazione elementare sulle rige di I_m, allora $B = EA$.
Analogamente, se B è ottenuta eseguendo una operazione elementare sulle colonne di A e E è la matrice elementare ottenuta eseguendo la stessa operazione elementare sulle colonne di I_n, allora $B = AE$.
Per la verifica, basta vedere qual è l'effetto delle moltiplicazioni $\underline{e}_i^\top A$ ($\underline{e}_i$ un vettore della base canonica di $\mathbb{K}^m$) e $A\underline{e}_j$ ($\underline{e}_j$ un vettore della base canonica di $\mathbb{K}^n$): è immediato vedere che $\underline{e}_i^\top A = \underset{i\text{—}}{A}$, la i-ma riga di A, $A\underline{e}_j = \overset{j}{A|}$, la j-ma colonna di A.

Otteniamo che $G_R(A)$ e $\tilde{G}_R(A)$ si ottengono da A moltiplicando a sinistra per un numero finito di matrici elementari. Analogamente, $G_C(A)$ e $\tilde{G}_C(A)$ si ottengono da A moltiplicando a destra per un numero finito di matrici elementari.

Osservazioni:
▶ Sia $A \in M(n,\mathbb{K})$.
A è invertibile $\iff$ $\operatorname{rnk} A = n$ $\iff$ applicando l'algoritmo di Gauss sulle

righe $G_R(A)$ è triangolare superiore e gli elementi sulla diagonale sono uguali a 1.
In tal caso, $\tilde{G}_R(A) = I_n$. Quindi esistono k matrici elementari di taglia $n \times n$ $E_1, \ldots, E_k$ tali che $E_k E_{k-1} \cdots E_1 A = I_n$. Poiché E_k è invertibile, abbiamo $E_{k-1} \cdots E_1 A = E_k^{-1}$, da cui $E_{k-1} \cdots E_1 A E_k = I_n$. Reiterando abbiamo $A E_k E_{k-1} \cdots E_1 = I_n$, per cui $A^{-1} = E_k E_{k-1} \cdots E_1$.
Osserviamo che le E_i sono facilmente determinabili poiché corrispondono alle operazioni eseguite su A dall'algoritmo di Gauss completo sulle righe.
▶ Abbiamo anche dimostrato che ogni matrice invertibile è prodotto di matrici elementari.
▶ Dallo studio dell'equivalenza destra-sinistra, per ogni $A \in M(m, n, \mathbb{K})$ di rango r esistono $M \in GL(m, \mathbb{K})$, $N \in GL(n, \mathbb{K})$ tali che $MAN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.
Per determinare delle possibili M e N, eseguiamo su A l'algoritmo di Gauss sulle righe e poi quello sulle colonne. È immediato vedere che $G_C(G_R(A)) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.
Allora, esistono matrici elementari di taglia $m \times m$, $E_1, \ldots, E_k$, ed esistono matrici elementari di taglia $n \times n$, $F_1, \ldots, F_h$, tali che $(E_k \cdots E_1) A (F_1 \cdots F_h) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ quindi possiamo prendere $M = E_k \cdots E_1$, $N = F_1 \cdots F_h$.
▶ Dato un sistema lineare $A\underline{X} = \underline{b}$, con matrice completa $M = (A|\underline{b})$, fare un'operazione elementare sulle righe di M significa moltiplicare a M a sinistra per una matrice elementare E, ottenendo $(EA|E\underline{b})$ che corrisponde al sistema lineare $EA\underline{X} = E\underline{b}$.
È chiaro quindi che i due sistemi lineari sono equivalenti poiché $E(A\underline{X} - \underline{b}) = 0$ se e solo se $(A\underline{X} - \underline{b}) = 0$, essendo E invertibile.
Il metodo risolutivo dato dall'algoritmo di Gauss si può reinterpretare dicendo che esiste una matrice invertibile Q per cui il sistema lineare $QA\underline{X} = Q\underline{b}$ è facilmente risolvibile.

Lo spazio vettoriale duale

Definizione:
Sia V uno spazio vettoriale su $\mathbb{K}$, il *duale* di V è lo spazio vettoriale su $\mathbb{K}$

$$V^* = \mathrm{Hom}(V, \mathbb{K}).$$

Gli elementi di V^* si dicono *funzionali* su V.

Se $\dim V = n$, allora per ogni base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ di V abbiamo l'isomorfismo $M^{\mathcal{B}}_{\mathcal{C}an} : V^* \to M(1, n, \mathbb{K})$, dove $\mathcal{C}an$ è la base canonica di $\mathbb{K}$ (data dal solo numero 1), quindi $\dim V^* = n$.
Usando la base standard $E_{11}, \dots, E_{1n}$ di $M(1, n, \mathbb{K})$, e prendendone le controimmagini tramite l'isomorfismo, otteniamo un'unica base $\mathcal{B}^* = \{v_1^*, \dots, v_n^*\}$ di V^* detta *base duale di* $\mathcal{B}$. Si ha quindi che, per ogni $i = 1\_n$, $M^{\mathcal{B}}_{\mathcal{C}an}(v_i^*) = E_{1i}$, ovvero $v_i^*(\underline{v}_j) = 0$ se $i \neq j$, $v_i^*(\underline{v}_i) = 1$. Definendo il simbolo $\delta_{ij} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}$, scriviamo $v_i^*(\underline{v}_j) = \delta_{ij}$.
Notiamo che possiamo anche evitare di usare l'isomorfismo per definire la base duale. Fissato i, esiste un unico funzionale $f_i \in V^*$ tale che

$$f_i(\underline{v}_j) = \delta_{ij} \quad \forall j = 1\_n,$$

poiché abbiamo dato i valori su una base di V (notiamo che $f_i = v_i^*$ poiché coincidono su una base di V). Mostriamo allora che $X = \{f_1, \dots, f_n\}$ è una base di V^*. Dati $a_1, \dots, a_n \in \mathbb{K}$, se $0 = a_1 f_1 + \cdots + a_n f_n$ è una combinazione lineare nulla, valutando su $\underline{v}_i$ abbiamo $0 = a_1 f_1(\underline{v}_i) + \cdots + a_n f_n(\underline{v}_i) = a_i f_i(\underline{v}_i) = a_i$. Quindi X è linearmente indipendente. Se poi $f \in V^*$, consideriamo la combinazione lineare $F = f(\underline{v}_1) f_1 + \cdots + f(\underline{v}_n) f_n$. Come prima, valutando su $\underline{v}_i$ abbiamo $F(\underline{v}_i) = f(\underline{v}_1) f_1(\underline{v}_i) + \cdots + f(\underline{v}_n) f_n(\underline{v}_i) = f(\underline{v}_i) f_i(\underline{v}_i) = f(\underline{v}_i)$. Poiché f e F coincidono su una base di V, $f = F$ e quindi X genera V^*.

Osservazioni:
▶ Se V ha dimensione finita, V e V^* sono isomorfi, ma non in modo naturale: non si può costruire un isomorfismo che non dipenda da qualche scelta.
Ad esempio l'isomorfismo $\chi_{\mathcal{B}} : V \to V^*$ che manda la base $\mathcal{B}$ nella base $\mathcal{B}^*$ mandando $\underline{v}_i \mapsto v_i^*$, $i = 1\_n$, dipende dalla scelta della base $\mathcal{B}$.
▶ Se V non ha dimensione finita, V^* non è isomorfo a V.
▶ Notiamo che nel caso $V = \mathbb{K}^n$, scegliendo $\mathcal{B} = \mathcal{C}an$ la base canonica di $\mathbb{K}^n$, e usando l'usuale identificazione di $(\mathbb{K}^n)^*$ con $M(1, n, \mathbb{K})$, la base duale $\mathcal{C}an^*$ corrisponde alla base $\{E_{11}, \dots, E_{1n}\}$ e $\chi_{\mathcal{C}an} : \mathbb{K}^n \to (\mathbb{K}^n)^*$ corrisponde a $^\top : \mathbb{K}^n \to M(1, n, \mathbb{K})$.

Definizione:
Data $F \in \mathrm{Hom}(V, W)$, l'applicazione $F^\top : W^* \to V^*$, $g \mapsto g \circ F$, si dice *trasposta di* F.

Mostriamo che $F^\top \in \mathrm{Hom}(W^*, V^*)$.

Infatti, dati due funzionali $g_1, g_2 \in W^*$,
$F^\top(g_1+g_2) = (g_1+g_2)\circ F = g_1\circ F + g_2 \circ F = F^\top(g_1) + F^\top(g_2)$.
Inoltre, dato $\mu \in \mathbb{K}$,
$F^\top(\mu g_1) = (\mu g_1)\circ F = \mu(g_1 \circ F) = \mu F^\top(g_1)$.

In modo completamente analogo, è immediato verificare che valgono le seguenti proprietà:
• $(F+G)^\top = F^\top + G^\top \quad \forall F, G \in \mathrm{Hom}(V,W)$,
• $(\mu F)^\top = \mu F^\top \quad \forall F \in \mathrm{Hom}(V,W), \lambda \in \mathbb{K}$,
• $id_V^\top = id_{V^*}$.
• $(F\circ G)^\top = G^\top \circ F^\top \quad \forall F \in \mathrm{Hom}(V,W), G \in \mathrm{Hom}(W,Z)$,
Infatti, se $g \in Z^*$,
$(F\circ G)^\top(g) = g\circ(F\circ G) = (g\circ F)\circ G = F^\top(g)\circ G = G^\top(F^\top(g))$.

Quindi $^\top : \mathrm{Hom}(V,W) \to \mathrm{Hom}(W^*, V^*)$ è lineare e se F è un isomorfismo, passando alle trasposte in $F\circ F^{-1} = id_W$ e $F^{-1}\circ F = id_V$, anche $F^\top$ lo è e $(F^\top)^{-1} = (F^{-1})^\top$.

Nel caso di $L_A : \mathbb{K}^n \to \mathbb{K}^m$, $A \in M(m,n,\mathbb{K})$, identificando come al solito $(\mathbb{K}^k)^*$ con $M(1,k,\mathbb{K})$, possiamo pensare $L_A^\top : M(1,m,\mathbb{K}) \to M(1,n,\mathbb{K})$, per cui, se $R \in M(1,n,\mathbb{K})$, $L_A^\top(R) = RA$ e se nei duali scegliamo le basi duali delle basi canoniche otteniamo $M_{\mathcal{C}an^*}^{\mathcal{C}an^*}(L_A^\top) = A^\top = \left(M_{\mathcal{C}an}^{\mathcal{C}an}(L_A)\right)^\top$.
Questo è vero in generale: se V e W hanno dimensione finita, $\mathcal{B}, \mathcal{D}$ sono basi di V e di W, $\mathcal{B}^*, \mathcal{D}^*$ le basi duali di V^* e W^*, allora

$$M_{\mathcal{B}^*}^{\mathcal{D}^*}(F^\top) = \left(M_{\mathcal{D}}^{\mathcal{B}}(F)\right)^\top.$$

Infatti, poniamo $A = M_{\mathcal{D}}^{\mathcal{B}}(F)$. Scriviamo $(F^\top)(w_1^*) = \alpha_1 v_1^* + \cdots + \alpha_n v_n^*$ e valutiamo su $\underline{v}_i$ per ottenere $\alpha_i = F^\top(w_1^*)(\underline{v}_i) = w_1^*(F(\underline{v}_i))$. Ma $F(\underline{v}_i) = a_{1i}\underline{w}_1 + \cdots + a_{ni}\underline{w}_n$, dove gli a_{ji} sono i coefficienti della i-ma colonna di A. Quindi $\alpha_i = a_{1i}$. La prima colonna della matrice $M_{\mathcal{B}^*}^{\mathcal{D}^*}(F^\top)$ è quindi $\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}$ cioè la trasposta della prima riga di A. Lo stesso vale per le altre colonne.

Definizione:
Dato $W \subset V$ un sottospazio, l'*annullatore di* W è dato da

$$Ann(W) = \{g \in V^* |\ g(\underline{w}) = 0 \ \ \forall \underline{w} \in W\}.$$

Equivalentemente,

$$Ann(W) = \{g \in V^* |\ W \subset \mathrm{Ker}\, g\} = \{g \in V^* |\ g_{|_W} = 0\},$$

ovvero, poiché $|_W : V^* \to W^*$ è lineare, $Ann(W) = \mathrm{Ker}\,|_W$. Quindi $Ann(W)$ è un sottospazio di V^*.

Proposizione
Se V ha dimensione finita, e $W \subset V$ è un sottospazio,

$$\dim Ann(W) = \dim V - \dim W.$$

Dimostrazione
Fissata una base di W $\{\underline{w}_1, \dots, \underline{w}_m\}$, estendiamo tale base a base di V ottenendo $\mathcal{B} = \{\underline{w}_1, \dots, \underline{w}_m, \underline{v}_{m+1}, \dots, \underline{v}_n\}$.
Consideriamo la base duale $\mathcal{B}^* = \{w_1^*, \dots, w_m^*, v_{m+1}^*, \dots, v_n^*\}$. Basta allora dimostrare che $H = \{v_{m+1}^*, \dots, v_n^*\}$ è una base di $Ann(W)$. Infatti, poiché $v_i^*(\underline{w}_j) = 0$ per ogni $i = m+1\_n, j = 1\_m$, i v_i^* appartengono a $Ann(W)$. Essendo $H \subset \mathcal{B}^*$, H è linearmente indipendente. Se $g \in Ann(W)$, scriviamo g come combinazione lineare della base $\mathcal{B}^*$, $g = a_1 w_1^* + \dots + a_m w_m^* + b_{m+1} v_{m+1}^* + \dots + b_n v_n^*$. Valutando su $\underline{w}_k$ otteniamo $0 = g(\underline{w}_k) = a_k$, per cui H genera $Ann(W)$. ❑

Osservazioni:
▶ Abbiamo esplicitamente esibito una base di $Ann(W)$.
▶ Alternativamente, osservando che ogni $g : W \to \mathbb{K}$ lineare si estende ad una $\tilde{g} : V \to \mathbb{K}$ lineare (ad esempio, usando la base $\mathcal{B}$ della dimostrazione, basta porre $\tilde{g}(\underline{w}_i) = g(\underline{w}_i)$, $\tilde{g}(\underline{v}_j) = 0$) abbiamo che $|_W : V^* \to W^*$ è surgettiva e basta applicare il teorema delle dimensioni.

Data $F \in \mathrm{Hom}(V, W)$,

$$\mathrm{Ker}\, F^\top = \{g \in W^* |\ g \circ F = 0\} = \{g \in W^* |\ \mathrm{Im}\, F \subset \mathrm{Ker}\, g\} = Ann(\mathrm{Im}\, F).$$

Quindi
$\dim \mathrm{Im}\, F^\top = \dim W^* - \dim Ann(\mathrm{Im}\, F) = \dim W - (\dim W - \dim \mathrm{Im}\, F) = \dim \mathrm{Im}\, F$.
Abbiamo quindi ottenuto per altra via il fatto che per ogni $A \in M(m, n, \mathbb{K})$, $\mathrm{rnk}\, A = \mathrm{rnk}\, A^\top$.

Osserviamo che $\mathrm{Im}\, F^\top = Ann(\mathrm{Ker}\, F)$, infatti hanno la stessa dimensione e se $g \in V^*$, $F^\top(g) \in Ann(\mathrm{Ker}\, F)$, poiché se $\underline{v} \in \mathrm{Ker}\, F$, $F^\top(g)(\underline{v}) = g(F(\underline{v})) = g(\underline{0}) = 0$.

Definizione:
Il *biduale* di V è lo spazio vettoriale su $\mathbb{K}$

$$V^{**} = (V^*)^* = \mathrm{Hom}(V^*, \mathbb{K}).$$

Se V ha dimensione finita, $\dim V = \dim V^* = \dim V^{**}$ e quindi V e V^{**} sono isomorfi. Data una base $\mathcal{B}$ di V, $\chi_{\mathcal{B}^*} \circ \chi_{\mathcal{B}} : V \to V^{**}$ è un isomorfismo, che a priori dipende dalla scelta della base $\mathcal{B}$.
Definiamo una applicazione lineare canonica $\chi : V \to V^{**}$ (cioè che non dipende da alcuna scelta arbitraria):
dato $\underline{v} \in V$, poniamo $\chi(\underline{v}) = val_{\underline{v}} : V^* \to \mathbb{K}$, la valutazione su $\underline{v}$, definita da

$val_{\underline{v}}(g) = g(\underline{v})$ per ogni $g \in V^*$.

Mostriamo che χ è ben definita, ovvero che $\chi(\underline{v}) = val_{\underline{v}}$ è in effetti un funzionale su V^*, cioè che è lineare:
dati $g_1, g_2 \in V^*$,
$val_{\underline{v}}(g_1 + g_2) = (g_1 + g_2)(\underline{v}) = g_1(\underline{v}) + g_2(\underline{v}) = val_{\underline{v}}(g_1) + val_{\underline{v}}(g_2)$;
se poi $\mu \in \mathbb{K}$,
$val_{\underline{v}}(\mu g_1) = (\mu g_1)(\underline{v}) = \mu(g_1(\underline{v})) = \mu val_{\underline{v}}(g_1)$.

Mostriamo adesso che χ è lineare:
dati $\underline{v}_1, \underline{v}_2 \in V$ si ha
$\chi(\underline{v}_1 + \underline{v}_2)(g) = g(\underline{v}_1 + \underline{v}_2) = g(\underline{v}_1) + g(\underline{v}_2) = \chi(\underline{v}_1)(g) + \chi(\underline{v}_2)(g) =$
$= (\chi(\underline{v}_1) + \chi(\underline{v}_2))(g)$ per ogni $g \in V^*$;
se poi $\mu \in \mathbb{K}$,
$\chi(\mu\underline{v}_1)(g) = g(\mu\underline{v}_1) = \mu g(\underline{v}_1) = \mu(\chi(\underline{v}_1)(g)) = (\mu\chi(\underline{v}_1))(g)$ per ogni $g \in V^*$.

Nel caso in cui V abbia dimensione finita, χ è un isomosrfismo. Infatti, avendo V e V^{**} la stessa dimensione, basta vedere che χ è iniettiva. Se $\chi(\underline{v}) = 0$, allora $0 = \chi(\underline{v})(g) = g(\underline{v})$ per ogni $g \in V^*$. Se fosse $\underline{v} \neq 0$, allora potremmo completare a base $\{\underline{v} = \underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ di V, ma allora il primo elemento della base duale v_1^* sarebbe un funzionale non nullo su $\underline{v}$ ↯.

Osservazioni:
▶ La χ è ben definita anche quando V ha dimensione infinita, ma in tal caso è solo iniettiva.
▶ Per ogni base $\mathcal{B}$ di V, $\chi = \chi_{\mathcal{B}^*} \circ \chi_{\mathcal{B}}$ (che quindi non dipende dalla base). Verifichiamo che i due isomorfismi coincidono sulla base $\mathcal{B}$. $(\chi_{\mathcal{B}^*} \circ \chi_{\mathcal{B}})(\underline{v}_i) = v_i^{**}$ è il funzionale su V^* che manda v_i^* in 1 e gli altri v_j^* in 0. $\chi(\underline{v}_i)(v_j^*) = v_j^*(\underline{v}_i) = \delta_{ij}$ ha gli stessi valori sulla base $\mathcal{B}^*$, quindi $(\chi_{\mathcal{B}^*} \circ \chi_{\mathcal{B}})(\underline{v}_i) = \chi(\underline{v}_i)$.

Data $F : V \to W$ lineare, abbiamo $(F^T)^T : V^{**} \to W^{**}$. Indicando con $\chi_V : V \to V^{**}$ e $\chi_W : W \to W^{**}$ le applicazioni canoniche, abbiamo

$$\chi_W \circ F = (F^T)^T \circ \chi_V.$$

Infatti, per ogni $\underline{v} \in V$, e per ogni $g \in W^*$,
$(\chi_W \circ F)(\underline{v}) = \chi_W(F(\underline{v})) = val_{F(\underline{v})}$, per cui $(\chi_W \circ F)(\underline{v})(g) = g(F(\underline{v}))$.
D'altra parte,
$((F^T)^T \circ \chi_V)(\underline{v}) = (F^T)^T(val_v) = val_{\underline{v}} \circ F^T$, per cui
$((F^T)^T \circ \chi_V)(\underline{v})(g) = val_{\underline{v}}(g \circ F) = g(F(\underline{v}))$, come voluto.

Nel caso V e W abbiano dimensione finita, χ_V e χ_W sono isomorfismi canonici, e quindi, a patto di identificare ogni spazio con il proprio biduale usando tali isomorfismi, possiamo pensare che "$(F^T)^T = F$".

Dato $W \subset V$ un sottospazio, $\chi(W)$ e $Ann(Ann(W))$ sono sottospazi di V^{**}.
Abbiamo $\chi(W) \subset Ann(Ann(W))$, infatti, dato $\underline{w} \in W$, e $g \in Ann(W)$, $val_{\underline{w}}(g) = g(\underline{w}) = 0$, e quindi $\chi(\underline{w}) = val_{\underline{w}} \in Ann(Ann(W))$.

Se V ha dimensione finita, $\dim \chi(W) = \dim W = \dim Ann(Ann(W))$, per cui

$$\chi(W) = Ann(Ann(W)).$$

Abbiamo allora che se $U, W \subset V$ sono sottospazi,

$$U = W \iff Ann(U) = Ann(W).$$

Infatti, se $Ann(U) = Ann(W)$ allora $Ann(Ann(U)) = Ann(Ann(W))$, ovvero $\chi(U) = \chi(W)$ e applicando χ^{-1}, $U = W$. L'altra freccia è ovvia.

Considerando l'annullatore come applicazione dai sottospazi di V ai sottospazi di V^*, $Ann : \{\text{Sottospazi di } V\} \to \{\text{Sottospazi di } V^*\}$ è quindi iniettiva.
Vogliamo mostrare che in effetti è biunivoca e per farlo costruiamo la sua inversa.

Se $U \subset V^*$ è un sottospazio, il suo *luogo di zeri* è dato da

$$Z(U) = \{\underline{v} \in V |\ g(\underline{v}) = 0\ \forall g \in U\}.$$

Notiamo che $Z(U) = \bigcap_{g \in U} \operatorname{Ker} g$ che quindi è un sottospazio di V.
Inoltre osserviamo che

$$Z(U) = \{\underline{v} \in V |\ val_{\underline{v}}(g) = 0\ \forall g \in U\} = \{\underline{v} \in V |\ val_{\underline{v}} \in Ann(U)\},$$

ovvero $Z(U) = \chi^{-1}(Ann(U))$ da cui $\chi(Z(U)) = Ann(U) \cap \operatorname{Im} \chi$.
In particolare, se V ha dimensione finita χ è un isomorfismo e abbiamo

$$\chi(Z(U)) = Ann(U),$$

da cui $\dim Z(U) = \dim \chi(Z(U)) = \dim Ann(U) = \dim V^* - \dim U$.

Se $W \subset V$ è un sottospazio, allora $Z(Ann(W)) = W$.
Infatti, $\chi(Z(Ann(W))) = Ann(Ann(W)) = \chi(W)$, e si applica χ^{-1}.

Se $U \subset V^*$ è un sottospazio, allora $Ann(Z(U)) = U$.
Infatti, $Ann(Ann(Z(U))) = \chi(Z(U)) = Ann(U)$, e si conclude osservando che $Ann : \{\text{ Sottospazi di } V^*\} \to \{\text{Sottospazi di } V^{**}\}$ è iniettiva.

Quindi l'applicazione $Z : \{\text{ Sottospazi di } V^*\} \to \{\text{Sottospazi di } V\}$ è l'inversa di $Ann : \{\text{ Sottospazi di } V\} \to \{\text{Sottospazi di } V^*\}$.

Dati $V_1, \dots, V_k, Z$ spazi vettoriali su $\mathbb{K}$, una $\phi : V_1 \times \cdots \times V_k \to Z$ (scriviamo $\phi(\underline{v}_1, \dots, \underline{v}_k)$ al posto del formalmente più corretto $\phi((\underline{v}_1, \dots, \underline{v}_k))$) si dice *lineare nell'i-mo argomento* se per ogni $\underline{v}_{j,0} \in V_j$, $j = 1 \_ k$, $j \neq i$ l'applicazione

$$\begin{gathered}\phi(\underline{v}_{1,0}, \dots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \dots, \underline{v}_{k,0}) : V_i \to Z \\ \underline{v} \mapsto \phi(\underline{v}_{1,0}, \dots, \underline{v}_{i-1,0}, \underline{v}, \underline{v}_{i+1,0}, \dots, \underline{v}_{k,0})\end{gathered}$$

è lineare (ovvero, se comunque si fissano gli altri argomenti, si ottiene una applicazione lineare nell'argomento i-mo).

Definizione:
$\phi : V_1 \times \cdots \times V_k \to Z$ si dice *multilineare* se è lineare in ogni argomento $(i = 1\_k)$.
L'insieme delle applicazioni multilineari da $V_1 \times \cdots \times V_k$ a Z si indica con $Mult(V_1 \times \cdots \times V_k, Z)$.
Per $k = 2$, tali applicazioni si dicono *bilineari* e l'insieme delle applicazioni bilineari da $V_1 \times V_2$ a Z si indica con $Bil(V_1 \times V_2, Z)$.

$Mult(V_1 \times \cdots \times V_k, Z)$ è un sottospazio di $Z^{V_1 \times \cdots \times V_k}$.
Infatti, l'applicazione nulla $0 : V_1 \times \cdots \times V_k \to Z$ è multilineare in quanto per ogni $\underline{v}_{j,0} \in V_j$, $j = 1\_k$, $0(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0}) = 0 : V_i \to Z$ che è lineare per ogni $i = 1\_k$.
Se ϕ_1, ϕ_2 sono multilineari, allora $\phi_1 + \phi_2$ è multilineare in quanto (con le notazioni di cui sopra)

$$(\phi_1 + \phi_2)(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0}) =$$

$$\phi_1(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0}) + \phi_2(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0})$$

che è lineare in quanto somma di lineari.
Se infine ϕ è multilineare e $\lambda \in \mathbb{K}$, allora $\lambda\phi$ è multilineare in quanto

$$(\lambda\phi)(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0}) =$$

$$\lambda(\phi(\underline{v}_{1,0}, \ldots, \underline{v}_{i-1,0}, \cdot, \underline{v}_{i+1,0}, \ldots, \underline{v}_{k,0}))$$

che è lineare in quanto prodotto per scalari di lineare.

Vediamo un'altra applicazione lineare canonica. Definiamo

$$\chi : \mathrm{Hom}(V, W) \to Bil(V \times W^*, \mathbb{K})$$

mandando $f : V \to W$ nell'applicazione bilineare $\phi_f : V \times W^* \to \mathbb{K}$ definita da $\phi_f(\underline{v}, g) = g(f(\underline{v}))$.

Mostriamo che χ è ben definita, ovvero che ϕ_f è bilineare:
fissato $\underline{v}_0 \in V$, $\phi_f(\underline{v}_0, \cdot) : W^* \to \mathbb{K}$; $g \mapsto g(f(\underline{v}_0))$, è la valutazione su $f(\underline{v}_0)$ e quindi lineare; fissato $g_0 \in W^*$, $\phi_f(\cdot, g_0) : V \to \mathbb{K}$, $\underline{v} \mapsto g_0(f(\underline{v}))$, è $g_0 \circ f$ e quindi lineare perché composizione di lineari.

Mostriamo che χ è lineare:
date $f_1, f_2 \in \mathrm{Hom}(V, W)$,
$\chi(f_1 + f_2)(\underline{v}, g) = g((f_1 + f_2)(\underline{v})) = g(f_1(\underline{v}) + f_2(\underline{v})) = g(f_1(\underline{v})) + g(f_2(\underline{v})) = \chi(f_1)(\underline{v}, g) + \chi(f_2)(\underline{v}, g)$ per ogni $\underline{v} \in V, g \in W^*$;
se poi $\mu \in \mathbb{K}$,
$\chi(\mu f_1)(\underline{v}, g) = g((\mu f_1)(\underline{v})) = g(\mu f_1(\underline{v})) = \mu g(f_1(\underline{v})) = \mu\chi(f_1)(\underline{v}, g)$ per ogni $\underline{v} \in V, g \in W^*$.

Nel caso W abbia dimensione finita, mostriamo che χ è un isomorfismo:
è iniettiva, in quanto se $\chi(f) = 0$, $g(f(\underline{v})) = 0$ per ogni g e $\underline{v}$. Se fosse $f \neq 0$, allora esisterebbe $\underline{v}_0 \in V$ tale che $f(\underline{v}_0) \neq 0$, ma allora esisterebbe un funzionale $g_0 \in W^*$ tale che $g_0(f(\underline{v}_0)) \neq 0$ ↯.
Per dimostrare la surgettività, fissiamo una base $\mathcal{D} = \{\underline{w}_1, \dots, \underline{w}_m\}$ di W e sia $\mathcal{D}^* = \{w_1^*, \dots, w_m^*\}$ la base duale; data $\phi \in Bil(V \times W^*, \mathbb{K})$, definiamo

$$f : V \to W, \quad \underline{v} \mapsto \sum_{j=1}^{m} \phi(\underline{v}, w_j^*)\underline{w}_j.$$

f è lineare poiché ϕ è lineare nel primo argomento, e $\chi(f) = \phi$ in quanto fissato $\underline{v} \in V$ coincidono sulla base $\mathcal{D}^*$.

Osservazioni:
▶ Se W ha dimensione infinita, χ è solo iniettiva.
▶ Scegliendo $W = V$ di dimensione finita otteniamo che $\mathrm{End}(V)$ è canonicamente isomorfo a $Bil(V \times V^*, \mathbb{K})$.
▶ Scegliamo $V = \mathbb{K}$ e W di dimensione finita.
$\mathrm{Hom}(\mathbb{K}, W)$ è canonicamente isomorfo a W, tramite gli isomorfismi $f \mapsto f(1)$ e $\underline{w} \mapsto g : \mathbb{K} \to W$, $g : \lambda \mapsto \lambda\underline{w}$.
$Bil(\mathbb{K} \times W^*, \mathbb{K})$ è canonicamente isomorfo a W^{**}, infatti, più in generale, $Bil(\mathbb{K} \times Z, \mathbb{K})$ è canonicamente isomorfo a Z^* tramite gli isomorfismi $\phi \mapsto \phi(1, \cdot)$ e $g \mapsto \varphi : \mathbb{K} \times Z \to \mathbb{K}$, $\varphi : (\lambda, \underline{z}) \mapsto g(\lambda\underline{z})$.
Componendo questi isomorfismi con l'isomorfismo canonico χ tra $\mathrm{Hom}(\mathbb{K}, W)$ e $Bil(\mathbb{K} \times W^*, \mathbb{K})$ si ottiene l'isomorfismo canonico tra W e W^{**}.

Definizione:
$\phi \in Mult(V^k, \mathbb{K})$ si dice *alternante* se per ogni $h = 1\_(k-1)$, per ogni $\underline{v}_i \in V$, $i = 1\_k$, $\underline{v}_h = \underline{v}_{h+1} = \underline{x}$, $\phi(\underline{v}_1, \dots, \underline{v}_{h-1}, \underline{x}, \underline{x}, \underline{v}_{h+2}, \dots, \underline{v}_k) = 0$, ovvero ϕ si annulla ogniqualvolta due argomenti adiacenti sono uguali.
L'insieme delle applicazioni multilineari alternanti su V^k a valori in $\mathbb{K}$, dette anche *k-forme*, si indica con $\Lambda^k(V)$.
È immediato vedere che $\Lambda^k(V)$ è un sottospazio di $Mult(V^k, \mathbb{K})$.

Se $\phi \in Mult(V^k, \mathbb{K})$ è alternante, allora:

a) per ogni $h = 1\_(k-1)$ e per ogni $\underline{v}_i \in V$, $i = 1\_k$,
$\phi(\underline{v}_1, \dots, \underline{v}_{h-1}, \underline{v}_{h+1}, \underline{v}_h, \underline{v}_{h+2}, \dots, \underline{v}_k) = -\phi(\underline{v}_1, \dots, \underline{v}_k)$,
ovvero, scambiando tra loro due argomenti adiacenti ϕ cambia di segno.

b) per ogni $1 \leq h < j \leq k$ e per ogni $\underline{v}_i \in V$, $\underline{v}_h = \underline{v}_j = \underline{x}$,
$\phi(\underline{v}_1, \dots, \underline{v}_{h-1}, \underline{x}, \underline{v}_{h+1}, \dots, \underline{v}_{j-1}, \underline{x}, \underline{v}_{j+1}, \dots, \underline{v}_k) = 0$,
ovvero, se due argomenti (anche non adiacenti) sono uguali ϕ si annulla.

c) per ogni $1 \leq h < j \leq k$ e per ogni $\underline{v}_i \in V$,
$\phi(\underline{v}_1, \dots, \underline{v}_{h-1}, \underline{v}_j, \underline{v}_{h+1}, \dots, \underline{v}_{j-1}, \underline{v}_h, \underline{v}_{j+1}, \dots, \underline{v}_k) = -\phi(\underline{v}_1, \dots, \underline{v}_k)$,
ovvero, scambiando tra loro due argomenti (anche non adiacenti) ϕ cambia di segno.

Per a), evidenziando solo gli argomenti di interesse, h e $h+1$, si ha, usando la linearità nell'h-mo argomento

$$\begin{aligned}0 &= \phi(\ldots, \underline{v}_h + \underline{v}_{h+1}, \underline{v}_h + \underline{v}_{h+1}, \ldots) = \\ &= \phi(\ldots, \underline{v}_h, \underline{v}_h + \underline{v}_{h+1}, \ldots) + \phi(\ldots, v_{h+1}, v_h + v_{h+1}, \ldots)\end{aligned}$$

e usando la linearità nell'$(h+1)$-mo argomento

$$\begin{aligned}0 &= \phi(\ldots, \underline{v}_h, \underline{v}_h, \ldots) + \phi(\ldots, \underline{v}_h, \underline{v}_{h+1}, \ldots) + \\ &\quad + \phi(\ldots, \underline{v}_{h+1}, \underline{v}_h, \ldots) + \phi(\ldots, \underline{v}_{h+1}, \underline{v}_{h+1}, \ldots) = \\ &= \phi(\ldots, \underline{v}_h, \underline{v}_{h+1}, \ldots) + \phi(\ldots, \underline{v}_{h+1}, \underline{v}_h, \ldots).\end{aligned}$$

Per b), scambiando l'argomento h-mo con l'argomento successivo $j-h-1$ volte e usando a) otteniamo

$$\begin{aligned}&\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{x}, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{x}, \underline{v}_{j+1}, \ldots, \underline{v}_k) = \\ &\quad = (-1)^{j-i-1}\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{x}, \underline{x}, \underline{v}_{j+1}, \ldots, \underline{v}_k) = 0.\end{aligned}$$

Per c), scambiando l'argomento h-mo con l'argomento successivo $j-h$ volte e usando a) otteniamo

$$\begin{aligned}&\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{v}_j, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{v}_h, \underline{v}_{j+1}, \ldots, \underline{v}_k) = \\ &\quad = (-1)^{j-h}\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{v}_h, \underline{v}_j, \underline{v}_{j+1}, \ldots, \underline{v}_k)\end{aligned}$$

e scambiando l'argomento $(j-1)$-mo con l'argomento precedente $j-h-1$ volte e usando a) otteniamo

$$\begin{aligned}&\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{v}_j, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{v}_h, \underline{v}_{j+1}, \ldots, \underline{v}_k) = \\ &\quad = (-1)^{2j-2h-1}\phi(\underline{v}_1, \ldots, \underline{v}_{h-1}, \underline{v}_h, \underline{v}_{h+1}, \ldots, \underline{v}_{j-1}, \underline{v}_j, \underline{v}_{j+1}, \ldots, \underline{v}_k) = \\ &\quad = -\phi(\underline{v}_1, \ldots, \underline{v}_k).\end{aligned}$$

Segue dalla proprietà b) che, se $\underline{v}_1, \ldots, \underline{v}_k \in V$ sono linearmente dipendenti, allora per ogni $\phi \in \Lambda^k(V)$, $\phi(\underline{v}_1, \ldots, \underline{v}_k) = 0$.

Infatti, esiste un $1 \leq i \leq k$ tale che $\underline{v}_i \in \mathrm{Span}(\underline{v}_1, \ldots, \underline{v}_{i-1}, \underline{v}_{i+1}, \ldots, \underline{v}_k)$, ovvero, esistono $a_j \in \mathbb{K}$, $j = 1\_k$, $j \neq i$, tali che $\underline{v}_i = \sum_{j\neq i} a_j \underline{v}_j$. Quindi, usando la linearità nell i-mo argomento,

$$\phi(\underline{v}_1, \ldots, \underline{v}_i, \ldots, \underline{v}_k) = \phi(\underline{v}_1, \ldots, \textstyle\sum_{j\neq i} a_j \underline{v}_j, \ldots, \underline{v}_k) = \displaystyle\sum_{j\neq i} a_j \phi(\underline{v}_1, \ldots, \underline{v}_j, \ldots, \underline{v}_k)$$

e si conclude notando che in $\phi(\underline{v}_1, \ldots, \underline{v}_j, \ldots, \underline{v}_k)$ due argomenti sono uguali.

Come corollario abbiamo che $\Lambda^k(V) = \{0\}$ se $k > \dim V$.

Nel seguito vogliamo dimostrare che $\Lambda^n(V) \neq \{0\}$ e che $\dim \Lambda^n(V) = 1$ (ricordiamo che $n = \dim V$). Per farlo, serve capire come cambia il valore di una applicazione multilineare alternante quando si permutano gli argomenti (la proprietà c) ci dice cosa succede con uno scambio semplice di due argomenti, e vogliamo estendere questo risultato ad una permutazione arbitraria).

Sia $S_k = S(\{1, 2, \dots, k\})$ il gruppo simmetrico su $J_k = \{1, 2, \dots, k\}$ (dato dalle applicazioni biunivoche $\sigma : J_k \to J_k$, dette *permutazioni* di $\{1, 2, \dots, k\}$).

Consideriamo il polinomio in k variabili (detto di Vandermonde)

$$P(x_1, \dots, x_k) = \prod_{1 \le i < j \le k} (x_i - x_j).$$

Per $\sigma \in S_k$, definiamo il *segno* di σ come

$$\epsilon(\sigma) = \frac{P(x_{\sigma(1)}, \dots, x_{\sigma(k)})}{P(x_1, \dots, x_k)}.$$

Poiché $P(x_{\sigma(1)}, \dots, x_{\sigma(k)})$ ha gli stessi fattori di $P(x_1, \dots, x_k)$, eventualmente cambiati di segno, $\epsilon(\sigma) = \pm 1$.
Le permutazioni per cui $\epsilon(\sigma) = 1$ si dicono *pari*, le altre si dicono *dispari*.

Osserviamo che, se $\sigma, \tau \in S_k$, allora

$$\epsilon(\sigma \circ \tau) = \frac{P(x_{\sigma(\tau(1))}, \dots, x_{\sigma(\tau(k))})}{P(x_1, \dots, x_k)} = \frac{P(x_{\tau(1)}, \dots, x_{\tau(k)})}{P(x_1, \dots, x_k)} \cdot \frac{P(x_{\sigma(\tau(1))}, \dots, x_{\sigma(\tau(k))})}{P(x_{\tau(1)}, \dots, x_{\tau(k)})} = \epsilon(\sigma)\epsilon(\tau)$$

ovvero $\epsilon : S_k \to \mathbb{Z}' = \{1, -1\}$ è un omomorfismo di gruppi.

Per $1 \le i < j \le k$, definiamo la *trasposizione* $\tau(i, j) \in S_k$, tale che

$$\tau(i,j)(h) = \begin{cases} j \text{ se } h = i \\ i \text{ se } h = j \\ h \text{ se } h \neq i, j \end{cases}$$

(ovvero, $\tau(i, j)$ scambia solo i con j). Osserviamo che $\tau(i, j)^2 = id_{J_k}$.
Definiamo anche $\tau(i, i) = id_{J_k}$, $i = 1 \_ k$.

Data $\sigma \in S_k$, componendo con $\tau_1 = \tau(1, \sigma(1))$ otteniamo $\sigma_1 = \tau_1 \circ \sigma$ tale che $\sigma_1(1) = 1$. Componendo con $\tau_2 = \tau(2, \sigma_1(2))$ otteniamo $\sigma_2 = \tau_2 \circ \sigma_1 = \tau_2 \circ \tau_1 \circ \sigma$ tale che $\sigma_2(1) = 1$ e $\sigma_2(2) = 2$.
Procedendo in modo induttivo, otteniamo $\sigma_i = \tau_i \circ \cdots \circ \tau_1 \circ \sigma$ tale che σ_i fissa $1, 2, \dots, i$.
Allora $\sigma_{k-1} = \tau_{k-1} \circ \cdots \circ \tau_1 \circ \sigma = id_{J_k}$, per cui $\sigma^{-1} = \tau_{k-1} \circ \cdots \circ \tau_1$ e quindi $\sigma = \tau_1 \circ \cdots \circ \tau_{k-1}$.
Eliminando i $\tau_j = id_{J_k}$ abbiamo che ogni permutazione è composizione di un numero finito di trasposizioni (in modo non unico).

Osserviamo che $\epsilon(\tau(i, j)) = -1$.
Infatti, se $h, l \neq i, j$, $h < l$ allora il fattore $(x_h - x_l)$ compare sia al numeratore che al denominatore dell'espressione di $\epsilon(\tau(i, j))$. Se $h \neq i, j$, allora i fattori $(x_h - x_i)(x_h - x_j)$ compaiono (eventualmente con entrambi i fattori cambiati di segno) sia al numeratore che al denominatore.
Otteniamo dunque $\epsilon(\tau(i, j)) = \frac{x_j - x_i}{x_i - x_j} = -1$.

Quindi $\sigma \in S_k$ è pari se e solo se è composizione di un numero pari di trasposizioni. Inoltre, $\epsilon(\sigma) = \epsilon(\sigma^{-1})$ (evidente anche dal fatto che $\epsilon(id_{J_k}) = 1$ e quindi che $\epsilon(\sigma)\epsilon(\sigma^{-1}) = 1$).

Osserviamo infine che se τ è una trasposizione, allora comporre con τ dà una bigezione tra l'insieme delle permutazioni pari e l'insieme delle permutazioni dispari.

Sia $\phi \in \Lambda^k(V)$ una applicazione multilineare alternante, e sia $\sigma \in S_k$ una permutazione. Allora

$$\phi(\underline{v}_{\sigma(1)}, \underline{v}_{\sigma(2)}, \dots, \underline{v}_{\sigma(k)}) = \epsilon(\sigma)\phi(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k).$$

Infatti, se $\sigma(j_1) = 1$ ($j_1 = \sigma^{-1}(1)$), scambiando il primo argomento con il j_1-mo argomento (se $j_1 = 1$ non effettuiamo scambi), abbiamo per la proprietà c)

$$\phi(\underline{v}_{\sigma(1)}, \dots, \underline{v}_{\sigma(j_1)} = \underline{v}_1, \dots, \underline{v}_{\sigma(k)}) = \epsilon(\tau(1, j_1))\phi(\underline{v}_1, \dots, \underline{v}_{\sigma(1)}, \dots, \underline{v}_{\sigma(k)}) = \\ = \epsilon(\tau(1, \sigma^{-1}(1)))\phi(\underline{v}_{\sigma_1(1)}, \underline{v}_{\sigma_1(2)}, \dots, \underline{v}_{\sigma_1(k)})),$$

dove $\sigma_1 = \sigma \circ \tau(1, \sigma^{-1}(1))$.
Se adesso $\sigma_1(j_2) = 2$, scambiamo il secondo argomento con il j_2-mo argomento e otteniamo

$$\phi(\underline{v}_{\sigma_1(1)}, \underline{v}_{\sigma_1(2)}, \dots, \underline{v}_{\sigma_1(j_2)} = \underline{v}_2, \dots, \underline{v}_{\sigma_1(k)}) = \epsilon(\tau(2, j_2))\phi(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_{\sigma_1(k)})$$

e quindi

$$\phi(\underline{v}_{\sigma(1)}, \underline{v}_{\sigma(2)}, \dots, \underline{v}_{\sigma(k)}) = \epsilon(\tau(2, \sigma_1^{-1}(2)))\epsilon(\tau(1, \sigma^{-1}(1)))\phi(\underline{v}_{\sigma_2(1)}, \dots, \underline{v}_{\sigma_2(k)}))$$

dove $\sigma_2 = \sigma \circ \tau(1, \sigma^{-1}(1)) \circ \tau(2, \sigma_1^{-1}(2))$.
Reiterando $k-1$ volte, si ha

$$\phi(\underline{v}_{\sigma(1)}, \underline{v}_{\sigma(2)}, \dots, \underline{v}_{\sigma(k)}) = \epsilon(\tau(k-1, \sigma_{k-2}^{-1}(k-1)) \circ \cdots \circ \tau(1, \sigma^{-1}(1)))\phi(\underline{v}_1, \dots, \underline{v}_k)$$

dove $id_{J_k} (= \sigma_k) = \sigma \circ \tau(1, \sigma^{-1}(1)) \circ \cdots \circ \tau(k-1, \sigma_{k-2}^{-1}(k-1))$.
Quindi $\tau(k-1, \sigma_{k-2}^{-1}(k-1)) \circ \cdots \circ \tau(1, \sigma^{-1}(1)) = \sigma$ come voluto.

Sia $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ una base di V e sia $\mathcal{B}^* = \{v_1^*, \dots, v_n^*\}$ la base duale. Ricordiamo che per ogni $\underline{v} \in V$ si ha $\underline{v} = \sum_{i=1}^n v_i^*(\underline{v})\, \underline{v}_i$.
Data $\phi \in \Lambda^n(V)$ e $\underline{w}_1, \dots, \underline{w}_n \in V$, usando la linearità nel primo argomento abbiamo

$$\phi(\underline{w}_1, \dots, \underline{w}_n) = \phi(\sum_{i_1=1}^n v_{i_1}^*(\underline{w}_1)\, \underline{v}_{i_1}, \dots, \underline{w}_n) = \sum_{i_1=1}^n v_{i_1}^*(\underline{w}_1)\phi(\underline{v}_{i_1}, \underline{w}_2, \dots, \underline{w}_n).$$

Usando adesso la linearità nel secondo argomento abbiamo

$$\phi(\underline{w}_1, \dots, \underline{w}_n) = \sum_{i_1=1}^n v_{i_1}^*(\underline{w}_1) \sum_{i_2=1}^n v_{i_2}^*(\underline{w}_2)\phi(\underline{v}_{i_1}, \underline{v}_{i_2}, \underline{w}_3, \dots, \underline{w}_n)$$

Notiamo che nel caso $i_1 = i_2$, $\phi(\underline{v}_{i_1}, \underline{v}_{i_2}, \dots, \underline{w}_k) = 0$, quindi possiamo scrivere

$$\phi(\underline{w}_1, \dots, \underline{w}_n) = \sum_{\substack{i_1, i_2 = 1 \\ i_1 \neq i_2}}^n v_{i_1}^*(\underline{w}_1) v_{i_2}^*(\underline{w}_2)\phi(\underline{v}_{i_1}, \underline{v}_{i_2}, \underline{w}_3, \dots, \underline{w}_n).$$

Reiterando per tutti gli argomenti otteniamo

$$\phi(\underline{w}_1,\dots,\underline{w}_n)=\sum_{\substack{i_1,\dots,i_n=1\\ \text{distinti}}}^{n} v^*_{i_1}(\underline{w}_1)\cdots v^*_{i_n}(\underline{w}_n)\phi(\underline{v}_{i_1},\underline{v}_{i_2},\dots,\underline{v}_{i_n}).$$

Notiamo che se $1\le i_1,\dots,i_n\le k$ sono distinti, definiscono una permutazione $\sigma\in S_n$ tale che $\sigma(j)=i_j$, $j=1\_n$. Allora possiamo riscrivere

$$\begin{aligned}\phi(\underline{w}_1,\dots,\underline{w}_n)&=\sum_{\sigma\in S_n} v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)\phi(\underline{v}_{\sigma(1)},\dots,\underline{v}_{\sigma(n)})=\\&=\sum_{\sigma\in S_n} v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)\epsilon(\sigma)\phi(\underline{v}_1,\dots,\underline{v}_n)=\\&=\phi(\underline{v}_1,\dots,\underline{v}_n)\sum_{\sigma\in S_n}\epsilon(\sigma)v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n).\end{aligned}$$

Osserviamo che ϕ è completamente determinata dal valore che ha sulla base $\mathcal{B}$, $\phi(\underline{v}_1,\dots,\underline{v}_n)$.
Quindi, se esiste $\bar\phi\in\Lambda^n(V)$ tale che $\bar\phi(\underline{v}_1,\dots,\underline{v}_n)=\alpha\neq 0$, allora per ogni $\varphi\in\Lambda^n(V)$ si ha che $\varphi=\dfrac{\varphi(\underline{v}_1,\dots,\underline{v}_n)}{\alpha}\bar\phi$ (in quanto hanno lo stesso valore sulla base $\mathcal{B}$). Ne consegue che $\dim\Lambda^n(V)\le 1$.

Per vedere che in effetti $\dim\Lambda^n(V)=1$, mostriamo che esiste una applicazione multilineare alternante $\phi_{\mathcal{B}}:V^n\to\mathbb{K}$ tale che $\phi_{\mathcal{B}}(\underline{v}_1,\dots,\underline{v}_n)=1$.
Osserviamo che se tale $\phi_{\mathcal{B}}$ esiste, allora è unica poiché dati $\underline{w}_1,\dots,\underline{w}_n\in V$, $\phi_{\mathcal{B}}(\underline{w}_1,\dots,\underline{w}_n)=\sum_{\sigma\in S_n}\epsilon(\sigma)v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)$.
Dimostriamo allora che l'applicazione

$$F:V^n\to\mathbb{K},\ (\underline{w}_1,\dots,\underline{w}_n)\mapsto\sum_{\sigma\in S_n}\epsilon(\sigma)v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)$$

è multilineare alternante e $F(\underline{v}_1,\dots,\underline{v}_n)=1$.
Fissando tutti gli argomenti eccetto i'i-mo abbiamo

$$F(\underline{w}_1,\dots,\cdot,\dots\underline{w}_n)=\sum_{\sigma\in S_n}\alpha_\sigma v^*_{\sigma(i)}\in V^*$$

poiché combinazione lineare di funzionali su V, dove $\alpha_\sigma=\epsilon(\sigma)\prod_{j\neq i}v^*_{\sigma(j)}(\underline{w}_j)\in\mathbb{K}$ sono costanti (non dipendono dall'i-mo argomento).
F è quindi multilineare.
$F(\underline{v}_1,\dots,\underline{v}_n)=\sum_{\sigma\in S_n}\epsilon(\sigma)v^*_{\sigma(1)}(\underline{v}_1)\cdots v^*_{\sigma(n)}(\underline{v}_n)$ e l'unico termine non nullo della somma è quello relativo a $\sigma=id_{J_n}$, per cui

$$F(\underline{v}_1,\dots,\underline{v}_n)=\epsilon(id_{J_n})v^*_1(\underline{v}_1)\cdots v^*_n(\underline{v}_n)=1.$$

Mostrare che F è alternante non è altrettanto semplice.
Vogliamo mostrare che se esiste un $1\le i<n$ tale che $\underline{w}_i=\underline{w}_{i+1}=\underline{x}$, allora $F(\underline{w}_1,\dots,\underline{x},\underline{x},\dots,\underline{w}_n)=0$.

Per farlo, spezziamo la somma che definisce F in due somme, sommando separatamente sulle permutazioni pari e sulle permutazioni dispari

$$F(\underline{w}_1,\dots,\underline{w}_n)=\sum_{\substack{\sigma\in S_n\\ \text{pari}}} v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)-\sum_{\substack{\sigma\in S_n\\ \text{dispari}}} v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)$$

e usiamo il fatto che la trasposizione $\tau=\tau(i,i+1)$ definisce per composizione l'applicazione $S_n\to S_n$, $\sigma\mapsto\sigma\circ\tau$, che dà una bigezione tra le permutazioni pari e le permutazioni dispari, per riscrivere la seconda somma come una somma sulle permutaizoni pari

$$\sum_{\substack{\sigma\in S_n\\ \text{dispari}}} v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n)=\sum_{\substack{\sigma\in S_n\\ \text{pari}}} v^*_{\sigma(\tau(1))}(\underline{w}_1)\cdots v^*_{\sigma(\tau(n))}(\underline{w}_n).$$

Ora,

$$\begin{aligned}v^*_{\sigma(\tau(1))}(\underline{w}_1)\cdots v^*_{\sigma(\tau(i))}(\underline{w}_i)v^*_{\sigma(\tau(i+1))}(\underline{w}_{i+1})\cdots v^*_{\sigma(\tau(n))}(\underline{w}_n)=\\ =v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(i+1)}(\underline{w}_i)v^*_{\sigma(i)}(\underline{w}_{i+1})\cdots v^*_{\sigma(n)}(\underline{w}_n)\end{aligned}$$

ma $\underline{w}_i=\underline{w}_{i+1}$, quindi

$$\begin{aligned}v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(i+1)}(\underline{w}_i)v^*_{\sigma(i)}(\underline{w}_{i+1})\cdots v^*_{\sigma(n)}(\underline{w}_n)=\\ =v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(i+1)}(\underline{w}_{i+1})v^*_{\sigma(i)}(\underline{w}_i)\cdots v^*_{\sigma(n)}(\underline{w}_n)=\\ =v^*_{\sigma(1)}(\underline{w}_1)\cdots v^*_{\sigma(n)}(\underline{w}_n).\end{aligned}$$

Le due somme quindi coincidono, come voluto.

Abbiamo dunque che $\dim\Lambda^n(V)=1$ e $\phi_{\mathcal{B}}$ ne è una base.

Data $f\in\mathrm{Hom}(W,V)$ e $\phi\in\Lambda^k(V)$, è facile vedere che l'applicazione ottenuta componendo ϕ con f in tutti gli argomenti,

$$\phi_f:W^k\to\mathbb{K},\ \phi_f(\underline{w}_1,\dots,\underline{w}_k)=\phi(f(\underline{w}_1),\dots,f(\underline{w}_k))$$

è multilineare e alternante. È altresì facile vedere che l'applicazione

$$f^*:\Lambda^k(V)\to\Lambda^k(W),\ \phi\mapsto\phi_f,$$

(detta *pullback* tramite f) è lineare ed è un isomorfismo se f lo è.
In particolare, se $k=n$, $W=V$, $f\in\mathrm{End}(V)$, abbiamo che ϕ_f è un multiplo di $\phi_{\mathcal{B}}$, $\phi_f=\phi_f(\underline{v}_1,\dots,\underline{v}_n)\phi_{\mathcal{B}}=\phi(f(\underline{v}_1),\dots,f(\underline{v}_n))\phi_{\mathcal{B}}$.
Ad esempio, usando l'isomorfismo dato dalle coordinate nella base $\mathcal{B}$, otteniamo $[\ ]^*_{\mathcal{B}}:\Lambda^n(\mathbb{K}^n)\to\Lambda^n(V)$ che manda $\phi_{\mathcal{C}an}$ in $\phi_{\mathcal{B}}$, dove $\mathcal{C}an$ è la base canonica di $\mathbb{K}^n$.

Scegliendo $\phi=\phi_{\mathcal{B}}$, si ottiene quella che si chiama la relazione di Binet:

$$\phi_{\mathcal{B}}(f(\underline{w}_1),\dots,f(\underline{w}_n))=\phi_{\mathcal{B}}(f(\underline{v}_1),\dots,f(\underline{v}_n))\phi_{\mathcal{B}}(\underline{w}_1,\dots,\underline{w}_n).$$

Se $\mathcal{B}' = \{\underline{w}_1, \ldots, \underline{w}_n\}$ è una base di V, scegliamo come f l'isomorfismo che manda $\mathcal{B}'$ in $\mathcal{B}$, $f(\underline{w}_i) = \underline{v}_i$, $i = 1\_\_n$. Dalla relazione di Binet si ottiene

$$1 = \phi_{\mathcal{B}}(\underline{v}_1, \ldots, \underline{v}_n) = \phi_{\mathcal{B}}(f(\underline{v}_1), \ldots, f(\underline{v}_n))\phi_{\mathcal{B}}(\underline{w}_1, \ldots, \underline{w}_n),$$

da cui $\phi_{\mathcal{B}}(\underline{w}_1, \ldots, \underline{w}_n) \neq 0$.
Avendo già osservato che per vettori linearmente dipendenti $\phi_{\mathcal{B}}$ si annulla, abbiamo:
$\underline{w}_1, \ldots, \underline{w}_n \in V$ sono una base di V se e solo se $\phi_{\mathcal{B}}(\underline{w}_1, \ldots, \underline{w}_n) \neq 0$
(lo stesso vale cambiando generatore di $\Lambda^n(V)$).

Osservazioni:
▶ Data $\phi \in Mult(V^k, \mathbb{K})$, allora la formula $\sum_{\sigma \in S_k} \epsilon(\sigma)\phi(\underline{v}_{\sigma(1)}, \ldots, \underline{v}_{\sigma(k)})$ definisce un elemento $Alt(\phi)$ di $\Lambda^k(V)$.
▶ Date $\phi \in \Lambda^k(V)$, $\psi \in \Lambda^h(V)$, possiamo definire il loro *prodotto esterno* $\phi \wedge \psi \in \Lambda^{k+h}(V)$ tramite la formula

$$\phi \wedge \psi(\underline{v}_1, \ldots, \underline{v}_{k+h}) = \sum_{\sigma \in S_{k+h}} \epsilon(\sigma)\phi(\underline{v}_{\sigma(1)}, \ldots, \underline{v}_{\sigma(k)})\psi(\underline{v}_{\sigma(k+1)}, \ldots, \underline{v}_{\sigma(k+h)}).$$

È facile vedere che il prodotto esterno è associativo, bilineare e commutativo pesato, ovvero che se $\phi \in \Lambda^k(V)$, $\psi \in \Lambda^h(V)$, $\phi \wedge \psi = (-1)^{k+h}\psi \wedge \phi$.
▶ Osservando che $\Lambda^1(V) = V^*$, ha senso fare il prodotto esterno di funzionali. In particolare, $\phi_{\mathcal{B}} = v_1^* \wedge v_2^* \wedge \cdots \wedge v_n^*$.
▶ Per $k \leq n$, l'insieme $\{v_{i_1}^* \wedge v_{i_2}^* \wedge \cdots \wedge v_{i_k}^* \mid 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$ è una base di $\Lambda^k(V)$ che quindi ha dimensione $\binom{n}{k}$.
▶ In maniera analoga, date $\phi \in Mult(V^k, \mathbb{K})$, $\psi \in Mult(V^h, \mathbb{K})$, possiamo definire il loro *prodotto tensoriale* $\phi \otimes \psi \in Mult(V^{k+h}, \mathbb{K})$ tramite la formula

$$\phi \otimes \psi(\underline{v}_1, \ldots, \underline{v}_{k+h}) = \phi(\underline{v}_1, \ldots, \underline{v}_k)\psi(\underline{v}_{k+1}, \ldots, \underline{v}_{k+h}).$$

È facile vedere che il prodotto tensoriale è associativo e bilineare (ma non commutativo).
Osservando che $Mult(V, \mathbb{K}) = V^*$, ha senso fare il prodotto tensoriale di funzionali.
L'insieme $\{v_{i_1}^* \otimes v_{i_2}^* \otimes \cdots \otimes v_{i_k}^* \mid 1 \leq i_1, i_2, \ldots, i_k \leq n\}$ è una base di $Mult(V^k, \mathbb{K})$ che quindi ha dimensione n^k.
▶ $\phi \wedge \psi = Alt(\phi \otimes \psi)$.
▶ Si possono definire le applicazioni multilineari *simmetriche* su V^k a valori in $\mathbb{K}$: sono quelle $\phi \in Mult(V^k, \mathbb{K})$ tali che $\phi(\underline{v}_{\sigma(1)}, \ldots, \underline{v}_{\sigma(k)}) = \phi(\underline{v}_1, \ldots, \underline{v}_k)$ per ogni $\underline{v}_1, \ldots, \underline{v}_k \in V$ e per ogni permutazione $\sigma \in S_k$. Tali applicazioni danno un sottospazio $Sym^k(V)$ di $Mult(V^k, \mathbb{K})$ di dimensione $\binom{n+k-1}{k}$.
Data $\phi \in Mult(V^k, \mathbb{K})$, allora la formula $\sum_{\sigma \in S_k} \phi(\underline{v}_{\sigma(1)}, \ldots, \underline{v}_{\sigma(k)})$ definisce un elemento $Sym(\phi)$ di $Sym^k(V)$.

Coniugio, Similitudine, Determinante

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n.
Diciamo che due endomorfismi $f, g \in End(V)$ sono *coniugati* se esiste un isomorfismo $h \in GL(V)$ tale che $g = h \circ f \circ h^{-1}$. Scriviamo $f \sim g$.
Poiché $f = id_V \circ f \circ id_V$, $g = h \circ f \circ h^{-1} \iff f = h^{-1} \circ g \circ h$, $g = h_1 \circ f \circ h_1^{-1}$, $l = h_2 \circ g \circ h_2^{-1} \Rightarrow l = (h_2 \circ h_1) \circ f \circ (h_2 \circ h_1)^{-1}$, essere coniugati è una relazione di equivalenza.
Considerando $V = \mathbb{K}^n$, la relazione di coniugazione dà una relazione di equivalenza sulle matrici quadrate $n \times n$, detta *similitudine*:
$A, B \in M(n, \mathbb{K})$ si dicono *simili*, e si scrive $A \sim B$, se esiste $P \in GL(n, \mathbb{K})$ tale che $B = PAP^{-1}$.
Vogliamo sudiare gli insiemi quoziente $End(V)/\sim$, $M(n, \mathbb{K})/\sim$.

Notiamo che la relazione di coniugazione/similitudine è un caso particolare della relazione destra-sinistra e in modo completamente analogo a quanto fatto per la relazione destra-sinsitra (usando, fissata una base $\mathcal{B}$ di V, l'isomorfismo $M_{\mathcal{B}}^{\mathcal{B}} : End(V) \to M(n, \mathbb{K})$) abbiamo la seguente:

Proposizione
Date $f_1, f_2 \in End(V)$, i seguenti fatti sono equivalenti:

1. $f_1 \sim f_2$.
2. Per ogni $\mathcal{B}$ base di V, $M_{\mathcal{B}}^{\mathcal{B}}(f_1) \sim M_{\mathcal{B}}^{\mathcal{B}}(f_2)$.
3. Esiste $\mathcal{B}$ base di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f_1) \sim M_{\mathcal{B}}^{\mathcal{B}}(f_2)$.
4. Esistono $\mathcal{B}, \mathcal{B}'$ basi di V tali che $M_{\mathcal{B}}^{\mathcal{B}}(f_1) = M_{\mathcal{B}'}^{\mathcal{B}'}(f_2)$.

Poiché se $f \sim g$ allora $f \sim_{DS} g$, la dimensione dell'immagine è un invariante per coniugazione. Analogamente, poiché se $A \sim B$ allora $A \sim_{DS} B$, il rango è un invariante per similitudine.
Osserviamo che, per $\lambda \in \mathbb{K}$ e per ogni $h \in GL(V)$, $h \circ (\lambda id_V) \circ h^{-1} = \lambda id_V$ (rispettivamente, per ogni $P \in GL(n, \mathbb{K})$, $P(\lambda I_n)P^{-1} = \lambda I_n$) allora la classe di coniugazione di λid_V (rispettivamente, la classe di similitudine di λI_n) contiene un solo elemento: $[\lambda id_V]_\sim = \{\lambda id_V\}$ (rispettivamente $[\lambda I_n]_\sim = \{\lambda I_n\}$).
Questo mostra che la dimensione dell'immagine (rispettivamente il rango) non è un invariante completo (almeno nel caso in cui $\mathbb{K}$ abbia almeno 3 elementi) come nel caso dell'equivalenza destra-sinistra: servono altri invarianti!
Trovare invarianti completi non sarà così immediato e lo faremo, su un campo arbitrario, solo per alcune classi di endomorfismi/matrici. Su un campo algebricamente chiuso e su $\mathbb{R}$ avremo una descrizione completa delle due relazioni (l'invariante completo sarà dato dalla forma normale di Jordan e dalla forma normale di Jordan reale, rispettivamente), mentre la situazione su altri campi è più complicata e non l'affronteremo (l'invariante completo è dato dalla forma normale razionale).

Un primo tentativo è quello di trovare una applicazione $D : M(n, \mathbb{K}) \to \mathbb{K}$ tale che $D(AB) = D(A)D(B)$ per ogni $A, B \in M(n, \mathbb{K})$. Questa fornirebbe un invariante per similitudine (e usando la proposizione sopra, un invariante per coniugazione):

$$D(P^{-1}AP) = D(P^{-1}A)D(P) = D(P)D(P^{-1}A) = D(PP^{-1}A) = D(A).$$

Se inoltre $D(I_n) \neq 0$, allora avremmo $D(A) \neq 0$ per ogni $A \in GL(n, \mathbb{K})$ (infatti $D(A)D(A^{-1}) = D(I_n)$) e potemmo usare D per decidere se una matrice è invertibile o meno.

Analizziamo ad esempio il caso $n = 2$.
Sia $A \in M(2, \mathbb{K})$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Allora A è invertibile se e solo se rnk $A = 2$, ovvero se e solo se $G_R(A)$ ha due pivot.
Se $G_R(A)$ ha due pivot, allora non può essere $a_{11} = a_{21} = 0$.
Se $a_{11} \neq 0$, allora il primo passo dell'algoritmo di Gauss dà $\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} \\ 0 & a_{22} - \frac{a_{12}a_{21}}{a_{11}} \end{pmatrix}$
per cui $a_{22} - \frac{a_{12}a_{21}}{a_{11}} \neq 0$.
Se $a_{11} = 0$, $a_{21} \neq 0$ e quindi il primo passo dell'algoritmo di Gauss dà $\begin{pmatrix} 1 & \frac{a_{22}}{a_{21}} \\ 0 & a_{12} \end{pmatrix}$
per cui $a_{12} \neq 0$.
In entrambi i casi, $a_{11}a_{22} - a_{21}a_{12} \neq 0$.
Se invece $G_R(A)$ ha meno di due pivot, allora o la prima colonna è nulla oppure, come sopra, $a_{11} \neq 0$, e $a_{22} - \frac{a_{12}a_{21}}{a_{11}} = 0$, o $a_{11} = 0$, $a_{21} \neq 0$ e $a_{12} = 0$.
In tutti i casi, $a_{11}a_{22} - a_{21}a_{12} = 0$.
Quindi l'applicazione $D : M(2, \mathbb{K}) \to \mathbb{K}$, $A \mapsto \underset{1\text{—}}{A^{1}|}\underset{2\text{—}}{A^{2}|} - \underset{2\text{—}}{A^{1}|}\underset{1\text{—}}{A^{2}|}$, ha la propietà: $A \in GL(2, \mathbb{K})$ se e solo se $D(A) \neq 0$.

Cerchiamo di individuare altre proprietà di D e per farlo, usiamo l'isomorfismo $\Phi : M(2, \mathbb{K}) \to \mathbb{K}^2 \times \mathbb{K}^2$ dato da $A \mapsto (A^{1}|, A^{2}|)$, la cui inversa è data da $\Phi^{-1} : \mathbb{K}^2 \times \mathbb{K}^2 \to M(2, \mathbb{K})$, $(\underline{X}, \underline{Y}) \mapsto (\underline{X}|\underline{Y})$ (ovvero, pensiamo una matrice di taglia 2×2 come la coppia data dalle sue colonne).
Allora è immediato verificare che $D \circ \Phi^{-1} : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}$ è bilineare (ovvero, fissata una colonna D è lineare come funzione dell'altra colonna).
Inoltre, se $A = (\underline{X}|\underline{X})$ ha entrambe le colonne uguali, $\underline{X} = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$, $x_1, x_2 \in \mathbb{K}$, allora $D(A) = x_1x_2 - x_2x_1 = 0$, ovvero $D \circ \Phi^{-1}$ è alternante.
Infine, $D \circ \Phi^{-1}(\underline{e}_1, \underline{e}_2) = D(I_2) = 1$, quindi $D \circ \Phi^{-1} = \phi_{\mathcal{C}an} \in \Lambda^2(\mathbb{K}^2)$.

Date $A, B \in M(2, \mathbb{K})$, $D(BA) = \phi_{\mathcal{C}an}(BA^{1}|, BA^{2}|) = (\phi_{\mathcal{C}an})_{L_B}(A^{1}|, A^{2}|)$ che per la regola di Binet vale $\phi_{\mathcal{C}an}(B\underline{e}_1, B\underline{e}_2)\phi_{\mathcal{C}an}(A^{1}|, A^{2}|) = \phi_{\mathcal{C}an}(B^{1}|, B^{2}|)\phi_{\mathcal{C}an}(A^{1}|, A^{2}|)$.

Otteniamo $D(AB) = D(A)D(B)$.

In questo caso concreto, vale la pena mostrarlo direttamente facendo vedere che l'applicazione $D' : M(2,\mathbb{K}) \to \mathbb{K}$, $A \mapsto D(BA)$, è bilineare e alternante nelle colonne e quindi coincide con un multiplo di D. Il risultato segue notando che $D'(I_2) = D(B)$.

$$\begin{aligned}
D'((X+Y|A_2)) &= D((B(X+Y)|BA_2)) = D((BX+BY)|BA_2)) = \\
&= D((BX|BA_2)) + D((BY|BA_2)) = D'((X|A_2)) + D'((Y|A_2)) \\
D'((A_1|X+Y)) &= D((BA_1|B(X+Y))) = D((BA_1|BX+BY)) = \\
&= D((BA_1|BX)) + D((BA_1|BY)) = D'((A_1|X)) + D'((A_1|Y)) \\
D'((\mu A_1|A_2)) &= D((\mu BA_1|BA_2)) = \mu D((BA_1|BA_2)) = \mu D'((A_1|A_2)) = \\
&= D((BA_1|\mu BA_2)) = D'((A_1|\mu A_2)).
\end{aligned}$$

Inolte, $D'((X|X)) = D((BX|BX)) = 0$.

Notiamo che il fatto che una matrice A di taglia 2×2 è invertibile se e solo se $D(A) \neq 0$ (da cui eravamo partiti) segue immediatamente dal fatto che $\phi_{\mathcal{C}an}$ non si annulla sulle basi di $\mathbb{K}^2$ mentre si annulla sulle coppie di vettori linearmente dipendenti.
Rivediamo il risultato teorico in questo caso concreto dimostrandolo usando le proprietà individuate.
Se A è invertibile, abbiamo $AA^{-1} = I_2$ e otteniamo $D(A)D(A^{-1}) = 1$, per cui $D(A) \neq 0$ ed inoltre $D(A^{-1}) = \frac{1}{D(A)}$.
Se A non è invertibile, allora una delle due colonne è multipla dell'altra. Se ad esempio $A^{\overset{1}{|}} = \underline{X}$ e $A^{\overset{2}{|}} = \lambda \underline{X}$, allora $D(A) = D((\underline{X}|\lambda \underline{X})) = \lambda D((\underline{X}|\underline{X})) = 0$.

Oserviamo che $D(AB) = D(BA)$ e, come osservato all'inizio, D è un invariante per similitudine.

Passando dalle matrici agli endomorfismi, se V è uno spazio vettoriale su $\mathbb{K}$ di dimensione 2 e $f \in \mathrm{End}(V)$, poiché le matrici associate a f nelle varie basi di V (stessa base in partenza e in arrvio) sono tutte simili tra di loro, possiamo definire $D : \mathrm{End}(V) \to \mathbb{K}$ ponendo $D(f) = D(M^{\mathcal{B}}_{\mathcal{B}}(f))$, dove $\mathcal{B}$ è una qualsiasi base di V.
Osserviamo che D è invariante per coniugazione e che $f \in GL(V)$ se e solo se $D(f) \neq 0$.

Notiamo che gli invarianti per coniugazione (o similitudine) $\dim \mathrm{Im}\, f$ e $D(f)$ ($\mathrm{rnk}\, A$ e $D(A)$) non sono invarianti completi: ad esempio I_2 e $-I_2$ hanno lo stesso rango e $D(I_2) = D(-I_2)$ ma non sono simili (se $2 \neq 0$ in $\mathbb{K}$).

Le proprietà individuate per D ci danno l'idea per estendere quanto appena fatto al caso di matrici di taglia $n \times n$ ed ottenere una applicazione $M(n,\mathbb{K}) \to \mathbb{K}$ con tutte le buone proprietà di D.

Definizione:
Il *determinante* $n \times n$ è l'applicazione $\det_n : M(n,\mathbb{K}) \to \mathbb{K}$ ottenuta componendo

l'isomorfismo $\Phi : M(n,\mathbb{K}) \to (\mathbb{K}^n)^n$, $A \mapsto (\overset{1}{A|}, \dots, \overset{n}{A|})$, con $\phi_{\mathcal{C}an}$

$$\det{}_n(A) = \phi_{\mathcal{C}an}(\overset{1}{A|}, \dots, \overset{n}{A|}).$$

Di solito si omette l'indice n, essendo chiara la taglia delle matrici di cui si fa il determinante, e di solito si scrive $\det A$ al posto di $\det(A)$, se non crea confusione.

Quindi, dalla definizione di multilineare alternante si ha:
1) det è multilineare nelle colonne (ovvero, fissate $n-1$ colonne, det è lineare nella restante colonna);
2) Se una matrice ha due colonne adiacenti uguali, $A = (\cdots|X|X|\cdots)$, allora $\det A = 0$.
E dalla definizione di $\phi_{\mathcal{C}an}$ si ha
3) $\det I_n = 1$.
Ripercorrendo la discussione fatta sulle applicazioni lineari alternanti otteniamo anche
4) $\det(\cdots|X|Y|\cdots) = -\det(\cdots|Y|X|\cdots)$.
5) $\det(\cdots|X|\cdots|X|\cdots) = 0$.
6) $\det(\cdots|X|\cdots|Y|\cdots) = -\det(\cdots|Y|\cdots|X|\cdots)$.
7) Per ogni $\sigma \in S_n$, indichiamo con $\sigma(A)$ la matrice ottenuta dalla matrice A permutando le colonne secondo σ^{-1}, ovvero $\sigma(A) = (\overset{\sigma(1)}{A|}\overset{\sigma(2)}{A|}|\cdots|\overset{\sigma(n)}{A|})$. Allora $\det\sigma(A) = \epsilon(\sigma)\det A$.
8) Data $A = (a_{ij}) \in M(n,\mathbb{K})$, per calcolare esplicitamente $\det A$ scriviamo ogni colonna come combinazione lineare della base canonica di $\mathbb{K}^n$, poi sviluppiamo totalmente usando la multilinearità in tutte le colonne ed eliminiamo i termini con due colonne uguali che danno contributo nullo. Otteniamo

$$\det A = \sum_{\sigma\in S_n} \epsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

9) det è l'unica applicazione da $M(n,\mathbb{K})$ a $\mathbb{K}$ che soddisfa 1), 2) e 3). Ogni altra applicazione $D' : M(n,\mathbb{K}) \to \mathbb{K}$ che soddisfa 1) e 2) è multipla di *det*: $D' = D'(I_n)\det$.

Osservazioni:
▶ Per $n = 1$ si ha $\det(a) = a$.
▶ Per $n = 2$ si ha $\det A = \epsilon(id_{J_2})\underset{1\text{—}}{\overset{1}{A|}}\underset{2\text{—}}{\overset{2}{A|}} + \epsilon(\tau(1,2))\underset{2\text{—}}{\overset{1}{A|}}\underset{1\text{—}}{\overset{2}{A|}} = \underset{1\text{—}}{\overset{1}{A|}}\underset{2\text{—}}{\overset{2}{A|}} - \underset{2\text{—}}{\overset{1}{A|}}\underset{1\text{—}}{\overset{2}{A|}}$ ovvero, $\det = D$.

In modo del tutto analogo al caso 2×2, possiamo dimostrare le seguenti affermazioni che discendono direttamente dalle proprietà di $\phi_{\mathcal{C}an}$ (o dalle proprietà analoghe per det enunciate con le colonne):

- per ogni $A, B \in M(n,\mathbb{K})$, $\det(BA) = \det B \det A$ (detta relazione di Binet);

- $A \in M(n,\mathbb{K})$ è invertibile se e solo se $\det A \neq 0$ e in tal caso si ha $\det A^{-1} = \frac{1}{\det A}$;
- se $A, B \in M(n,\mathbb{K})$, $B \sim A$ allora $\det B = \det A$ (det è invariante per similitudine);
- per ogni V spazio vettoriale su $\mathbb{K}$ di dimensione n e per ogni $f \in \mathrm{End}(V)$, è ben definito $\det f = \det M_{\mathcal{B}}^{\mathcal{B}}(f)$, dove $\mathcal{B}$ è una qualsiasi base di V, ed è un invariante per coniugazione.

Proposizione
Per ogni $A \in M(n,\mathbb{K})$, $\det A^\top = \det A$.

Dimostrazione
Poniamo $a_{ij} = \underset{i\text{—}}{A}\overset{j}{|}$, per cui $\underset{i\text{—}}{A^\top}\overset{j}{|} = a_{ji}$.
L'applicazione $\sigma \mapsto \sigma^{-1}$ è una bigezione di S_n che preserva la parità, ovvero $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$. Inoltre, riordinando i termini di $a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$ secondo il secondo indice, otteniamo $a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}$. Allora

$$\begin{aligned}\det A^\top &= \sum_{\sigma\in S_n} \epsilon(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = \\ &= \sum_{\sigma\in S_n} \epsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n} = \\ &= \sum_{\sigma^{-1}\in S_n} \epsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n} = \det A.\end{aligned}$$

Diamo adesso un altro modo di definire $\det_n$ per induzione su n (notiamo che questa è una dimostrazione alternativa che esistono applicazioni multilineari alternanti non nulle su $\mathbb{K}^n$).

Definiamo ricorsivamente le seguenti applicazioni $F_n : M(n,\mathbb{K}) \to \mathbb{K}$.
Per $n = 1$, poniamo $F_1((a)) = a$ per ogni $a \in \mathbb{K}$.
Per $n > 1$, supponiamo di disporre di F_{n-1} e fissiamo un arbitrario indice di riga i, $1 \le i \le n$.
Data $A \in M(n,\mathbb{K})$, e dato un indice di colonna j, $1 \le j \le n$, sia $\underset{i\text{—}}{A}\overset{j}{|}$ la sottomatrice di taglia $(n-1)\times(n-1)$ ottenuta da A cancellando la riga i e la colonna j.
Definiamo $F_n(A) = \sum_{j=1}^{n} (-1)^{i+j} \underset{i\text{—}}{A}\overset{j}{|} F_{n-1}(\underset{i\text{—}}{A}\overset{j}{|})$, detto *sviluppo di Laplace rispetto alla i-ma riga.*
Osserviamo che se $n = 2$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, sviluppando rispetto alla prima riga otteniamo
$F_2(A) = \sum_{j=1}^{2}(-1)^{1+j}a_{1j}F_1(\underset{1\text{—}}{A}\overset{j}{|}) = (-1)^2 a_{11}F_1(\underset{1\text{—}}{A}\overset{1}{|}) + (-1)^3 a_{12}F_1(\underset{1\text{—}}{A}\overset{2}{|})$
$= a_{11}F_1(a_{22}) - a_{12}F_1(a_{21}) = a_{11}a_{22} - a_{12}a_{21}$.

Sviluppando rispetto alla seconda riga otteniamo
$F_2(A) = \sum_{j=1}^{2}(-1)^{2+j}a_{2j}F_1(\underset{2}{\overset{j}{A}}) = (-1)^3 a_{21}F_1(\underset{2}{\overset{1}{A}}) + (-1)^4 a_{22}F_1(\underset{2}{\overset{2}{A}})$
$= -a_{21}F_1(a_{12}) + a_{22}F_1(a_{11}) = a_{11}a_{22} - a_{12}a_{21}$.
Quindi $F_2 = \det_2$ e non dipende dalla riga usata per lo sviluppo.

Vediamo adesso, per induzione su n, che F_n verifica le proprietà 1), 2) e 3), cioè che $F_n = \det_n$ e quindi non dipende dalla riga usata per lo sviluppo.

Per $n = 1$ la verifica è diretta. Supponiamo allora $n > 1$.
Per mostrare la multilinearità di F_n nelle colonne, basta mostrare che ogni addendo della somma $(-1)^{i+j}\underset{i}{\overset{j}{A}}F_{n-1}(\underset{i}{\overset{j}{A}})$ è una funzione multilineare delle colonne.
Fissiamo allora un indice di colonna k e mostriamo che $(-1)^{i+j}\underset{i}{\overset{j}{A}}F_{n-1}(\underset{i}{\overset{j}{A}})$ è lineare rispetto alla k-ma colonna.
Se $k = j$, allora $(-1)^{i+j}F_{n-1}(\underset{i}{\overset{j}{A}}) = (-1)^{i+k}F_{n-1}(\underset{i}{\overset{k}{A}})$ non dipende dalla colonna k, esendo costante (in $\underset{i}{\overset{k}{A}}$ si è tolta proprio la colonna k), mentre $\underset{i}{\overset{j}{A}} = \underset{i}{\overset{k}{A}}$ è il coeffieiente i-mo della colonna k-ma di A, che è una funzione lineare della colonna k-ma.
Se $k \neq j$, $(-1)^{i+j}\underset{i}{\overset{j}{A}}$ non dipende dalla colonna k, essendo costante, mentre per ipotesi induttiva, $F_{n-1}(\underset{i}{\overset{j}{A}})$ è lineare nella sua colonna corrispondente alla colonna k-ma di A; tale colonna contiene $n-1$ coefficienti della colonna k-ma di A, che sono funzioni lineari della colonna k-ma. Si conclude ricordando che la composizione di applicazioni lineari è lineare.

Consideriamo adesso una matrice con la colonna k-ma uguale alla colonna $(k+1)$-ma, $A = (\ldots|X|X|\ldots)$.
Abbiamo $\underset{i}{\overset{k}{A}} = \underset{i}{\overset{(k+1)}{A}}$ e $\underset{i}{\overset{k}{A}} = \underset{i}{\overset{k+1}{A}}$. Inoltre, osserviamo che se $j \neq k$, $\underset{i}{\overset{j}{A}}$ ha due colonne uguali, e quindi per ipotesi induttiva $F_{n-1}(\underset{i}{\overset{j}{A}}) = 0$. Abbiamo allora $F_n(A) = \sum_{j=1}^{n}(-1)^{i+j}\underset{i}{\overset{j}{A}}F_{n-1}(\underset{i}{\overset{j}{A}}) = (-1)^{i+k}\underset{i}{\overset{k}{A}}F_{n-1}(\underset{i}{\overset{k}{A}}) +$
$(-1)^{i+k+1}\underset{i}{\overset{k+1}{A}}F_{n-1}(\underset{i}{\overset{k+1}{A}}) = 0$

Per finire, se $A = I_n$, $\underset{i}{\overset{j}{A}} = \delta_{ij}$ e $\underset{i}{\overset{i}{A}} = I_{n-1}$, quindi $F_n(I_n)$ vale
$\sum_{j=1}^{n}(-1)^{i+j}\underset{i}{\overset{j}{A}}F_{n-1}(\underset{i}{\overset{j}{A}}) = (-1)^{i+i}\underset{i}{\overset{i}{A}}F_{n-1}(\underset{i}{\overset{i}{A}}) = F_{n-1}(I_{n-1}) = 1.$

Osservazioni:
▶ Ricordando che le colonne di $A^\top$ sono le righe di A, le proprietà 1) e 2) del determinante (e le altre che ne derivano) sono vere anche per le righe.
▶ Possiamo sviluppare il determinante anche rispetto ad una colonna: fissiamo un indice di colonna j, $1 \le j \le n$, e otteniamo

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} \overset{j}{\underset{i}{|}} A \overset{j}{\underset{i}{|}} \det \overset{\not{j}}{\underset{\not{i}}{|}} A \overset{\not{j}}{\underset{\not{i}}{|}}.$$

▶ Ad esempio, sviluppando reiteratamente rispetto alla prima colonna, se $A \in M(n, \mathbb{K})$ è triangolare superiore, $\det A = \prod_{i=1}^{n} \overset{i}{\underset{i}{|}} A \overset{i}{\underset{i}{|}}$. In particolare A è invertibile se e solo se gli elementi sulla diagonale sono tutti non nulli. Lo stesso vale per le matrici triangolari inferiori.

Data $A \in M(m, n, \mathbb{K})$, siano $1 \le h < m$, $1 \le k < n$. Suddividiamo A in quattro sottomatrici $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, dove A_1 è di taglia $h \times k$ (e quindi le taglie di A_2 A_3, A_4 sono $h \times (n-k)$, $(m-h) \times k$, $(m-h) \times (n-k)$). Se $B \in M(n, s, \mathbb{K})$, facciamo la stessa cosa con gli indici k e $1 \le l < s$ $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$, con B_1 di taglia $k \times l$.
Per determinare il coefficiente $(1,1)$ di AB, usiamo la prima riga di A, composta dalla prima riga di A_1 e dalla prima riga di A_2, $\underset{1}{A} = (\underset{1}{A_1} \mid \underset{1}{A_2})$, e la prima colonna di B, composta dalla prima colonna di B_1 e dalla prima colonna di B_3, $B\overset{1}{|} = \begin{pmatrix} B_1 \overset{1}{|} \\ \hline B_3 \overset{1}{|} \end{pmatrix}$. Quindi $\underset{1}{AB}\overset{1}{|} = \underset{1}{A_1} B_1 \overset{1}{|} + \underset{1}{A_2} B_3 \overset{1}{|}$. Lo stesso vale per gli altri coefficienti di AB, con le dovute modifiche.
Otteniamo una suddivisione di AB del tipo

$$AB = \begin{pmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{pmatrix}.$$

Analogamente, se suddividiamo A in pq blocchi, $A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,p} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,p} \\ \vdots & \vdots & \cdots & \vdots \\ A_{q,1} & A_{q,2} & \cdots & A_{q,p} \end{pmatrix}$
(le sottomatrici con primo indice fissato hanno lo stesso numero di righe, quelle con secondo indice fissato hanno lo stesso numero di colonne) e suddividiamo in modo analogo B in qr blocchi $B_{i,j}$ in modo che tutti i prodotti seguenti abbiano senso, allora otteniamo una suddivisione di AB con pr blocchi del tipo $(AB)_{i,j} = \sum_{k=1}^{q} A_{i,k} B_{k,j}$.

Nel caso A sia quadrata, i blocchi di A con primo indice maggiore del secondo siano nulli, $A_{ij} = 0$ se $i > j$, e i blocchi con indice uguale siano quadrati (per

cui $q = p$), A si dice *triangolare (superiore) a blocchi*.
Se A è triangolare a blocchi e anche i blocchi con primo indice minore del secondo sono nulli, e quindi $A_{i,j} = 0$ se $i \neq j$, A si dice *diagonale a blocchi* e si indica con $A = diag(A_{1,1}, A_{2,2}, \dots, A_{p,p})$.

Se $A \in M(n, \mathbb{K})$ è triangolare a blocchi, il determinante di A è il prodotto dei determinanti dei blocchi lungo la diagonale: $\det A = \prod_{i=1}^{p} \det A_{i,i}$.
Dimostriamolo per induzione su n, osservando che i casi $n = 1, 2$ sono ovvi.

Poniamo $B = A_{1,1} \in M(n_1, \mathbb{K})$, $b_{ij} = \underset{i}{\underset{—}{B^{j}}}|$.
Sviluppiamo il determinante lungo la prima colonna di A, osservando che intervengono solo i coefficienti della prima colonna di B:

$$\det A = \sum_{i=1}^{n_1} (-1)^{1+i} b_{i1} \det \underset{\nwarrow}{\underset{—}{A^{\nwarrow}}}|.$$

La sottomatrice $\underset{\nwarrow}{\underset{—}{A^{\nwarrow}}}| \in M(n-1, \mathbb{K})$ è triangolare a blocchi, con blocco $(1,1)$ $\underset{\nwarrow}{\underset{—}{B^{\nwarrow}}}|$ e blocchi (i,i) $A_{i,i}$ se $i > 1$, quindi per ipotesi induttiva,

$$\det \underset{\nwarrow}{\underset{—}{A^{\nwarrow}}}| = \det \underset{\nwarrow}{\underset{—}{B^{\nwarrow}}}| \prod_{j=2}^{p} \det A_{j,j}.$$

Quindi

$$\det A = \textstyle\sum_{i=1}^{n_1} (-1)^{1+i} b_{i1} \det \underset{\nwarrow}{\underset{—}{B^{\nwarrow}}}| \prod_{j=2}^{p} \det A_{j,j} =$$
$$= \left(\textstyle\prod_{j=2}^{p} \det A_{j,j}\right) \sum_{i=1}^{n_1} (-1)^{1+i} b_{i1} \det \underset{\nwarrow}{\underset{—}{B^{\nwarrow}}}| = \left(\textstyle\prod_{j=2}^{p} \det A_{j,j}\right) \det B.$$

Osserviamo che per le matrici elementari

- $\det T_{i,j} = -1$ (primo tipo);
- $\det D_{\lambda,i} = \lambda$ (secondo tipo);
- $\det L_{i,c,j} = 1$ (terzo tipo).

In particolare, se A' è ottenuta da A eseguendo operazioni del terzo tipo (di riga o di colonna), allora $\det A' = \det A$.
Allora, possiamo usare l'algoritmo di Gauss sulle righe di $A \in M(n, \mathbb{K})$ per calcolarne il determinante: calcoliamo $G_R(A)$ tenendo conto di quante e quali operazioni del primo e secondo tipo eseguiamo. Se $G_R(A)$ non ha n pivot, $\operatorname{rnk} A < n$ e $\det A = 0$. Se ha n pivot, $G_R(A)$ è triangolare superiore e il suo determinante vale 1. Se abbiamo eseguito m operazioni del secondo tipo usando $\lambda_1, \dots, \lambda_m \in \mathbb{K}$, allora $1 = \det G_R(A) = \pm \prod_{i=i}^{m} \lambda_i \det A$, dove abbiamo il segno $+$ se abbiamo eseguito un numero pari di operazioni del primo tipo, il segno $-$ altrimenti.

Consideriamo il sistema lineare $A\underline{X} = \underline{b}$ con $A \in M(n, \mathbb{K})$, $\underline{b} \in \mathbb{K}^n$.
Per ogni indice di colona j, sia $A(\underline{b}, j)$ la matrice ottenuta da A sostituendo la

colonna j con $\underline{b}$.
Supponiamo che $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sia una soluzione. Allora,
$\det A(\underline{b}, j) = \det(\overset{1}{A^|}| \cdots |\underline{b}| \cdots |\overset{n}{A^|}) = \det(\overset{1}{A^|}| \cdots |A\underline{X}| \cdots |\overset{n}{A^|}) =$
$\det(\overset{1}{A^|}| \cdots |x_1\overset{1}{A^|} + \cdots + x_n\overset{n}{A^|}| \cdots |\overset{n}{A^|}) = \det(\overset{1}{A^|}| \cdots |x_j\overset{j}{A^|}| \cdots |\overset{n}{A^|}) = x_j \det A.$
Quindi, se $\det A \neq 0$, il sisema ha un'unica soluzione $x_j = \frac{\det A(\underline{b},j)}{\det A}$.
Il fatto che il sistema ha un'unica soluzione era già noto, visto che se $\operatorname{rnk} A = n$ le soluzioni sono un sottospazio affine di dimensione 0, ovvero A è invertibile e $\underline{X} = A^{-1}\underline{b}$, ma qui si dà un modo esplicito per calcolare direttamente la soluzione attraverso il determinante.
Questa è detta *formula di Cramer* per la risoluzione dei sistemi quadrati con matrice invertibile.

Possiamo usare la formula di Cramer anche nel caso di matrice non quadrata o non invertibile.
Consideriamo il sistema lineare $A\underline{X} = \underline{b}$ con $A \in M(m, n, \mathbb{K})$, $\underline{b} \in \mathbb{K}^m$ e matrice completa M. Sia $r = \operatorname{rnk} A = \operatorname{rnk} M$ (il sistema è quindi risolubile).
Selezioniamo r righe linearmente indipendenti di M, e sia M' la sottomatrice di M contenente tali righe. $M' = (A'|\underline{b}')$, con $A' \in M(r, n, \mathbb{K})$, $\underline{b}' \in \mathbb{K}^r$ è la matrice completa del sistema lineare $A'\underline{X} = \underline{b}'$ che è equivalente al sistema iniziale, in quanto le righe che abbiamo tolto sono combinazione lineare delle righe di M' e corrispondono a equazioni del sistema iniziale che sono combinazioni lineari delle equazioni del sistema $A'\underline{X} = \underline{b}'$.
Osserviamo che $\operatorname{rnk} A' = r$, quindi selezioniamo r colonne linearmente indipendenti. Per convenienza notazionale, supponiamo siano le prime r colonne.
Scriviamo allora $A = (A_1|A_2)$, $\underline{X} = \left(\frac{\underline{X}_1}{\underline{X}_2}\right)$, dove $A_1 \in M(r, \mathbb{K})$, $X_1 \in \mathbb{K}^r$ e A_1 è invertibile poiché ha le colonne linearmente indipendenti. Scriviamo allora $A_1\underline{X}_1 = \underline{b} - A_2\underline{X}_2$. Fissato $\underline{X}_2 \in \mathbb{K}^{n-r}$ abbiamo un sistema lineare quadrato con matrice invertibile che è quindi risolubile tramite la formula di Cramer. Otteniamo $(\underline{X}_1)_j = \frac{\det A_1(\underline{b}-A_2\underline{X}_2,j)}{\det A_1}$ che esprime le r variabili contenute in $\underline{X}_1$ in funzione delle altre $n - r$ che sono libere di variare su tutto $\mathbb{K}$.

Se $A \in GL(n, \mathbb{K})$, allora se B^j è l'j-ma colonna di A^{-1}, allora B^j è (l'unica) soluzione di $A\underline{X} = \underline{e}_j$. Se x_{ij} è il coefficiente di posto (i, j) di A^{-1}, abbiamo allora

$$x_{i,j} = \frac{\det A(\underline{e}_j, i)}{\det A} = (-1)^{i+j} \frac{\det \overset{i}{\underset{j}{A^|}}}{\det A}$$

(notare l'inversione degli indici: per determinare il coefficiente di posto (i, j) si cancellano la riga j e la colonna i).

In generale, data $A \in M(n, \mathbb{K})$ la matrice $n \times n$ con coefficiente (i, j) uguale a $(-1)^{i+j} \det \overset{i}{\underset{j}{A^|}}$ è detta *matrice aggiunta classica di* A e si indica con $adj(A)$.

Abbiamo $A(adj(A)) = (adj(A))A = (\det A)I_n$.
Infatti, se $i \neq j$;

$$\underset{i}{\underline{\quad}}A(adj(A))\overset{j}{|} = \sum_{k=1}^{n} \underset{i}{A}\overset{k}{|}\underset{k}{\underline{\quad}}(adj(A))\overset{j}{|} = \sum_{k=1}^{n}(-1)^{k+j}\underset{i}{\overset{k}{A}}| \det \underset{j}{\overset{k}{A}}| = 0$$

poiché questo è lo sviluppo rispetto alla riga j della matrice ottenuta da A rimpiazzando la riga j con la riga i, che ha quindi due righe uguali.
Se invece $i = j$,

$$\underset{i}{\underline{\quad}}A(adj(A))\overset{i}{|} = \sum_{k=1}^{n} \underset{i}{A}\overset{k}{|}\underset{k}{\underline{\quad}}(adj(A))\overset{i}{|} = \sum_{k=1}^{n}(-1)^{k+i}\underset{i}{\overset{k}{A}}| \det \underset{i}{\overset{k}{A}}| = \det A.$$

Quindi $A(adj(A)) = (\det A)I_n$.
Allo stesso modo si ragiona per $(adj(A))A$ usando gli sviluppi per colonna.

Osservazioni:
▶ Se A è invertibile, allora $A^{-1} = \dfrac{1}{\det A} adj(A)$.
▶ Se esiste $B \in M(n, \mathbb{K})$ tale che $AB = I_n$, allora $\det A \det B = 1$, per cui $\det A \neq 0$. Inoltre $(\det A)B = adj(A)$, da cui $BA = I_n$ e quindi A è invertibile e B è l'inversa. Lo stesso vale se esiste $C \in M(n, \mathbb{K})$ tale che $CA = I_n$.
▶ Dato $\underline{b} \in \mathbb{K}^n$,

$$adj(A)\underline{b} = \begin{pmatrix} \det A(\underline{b},1) \\ \det A(\underline{b},2) \\ \vdots \\ \det A(\underline{b},n) \end{pmatrix},$$

$$\underline{b}^\top adj(A) = \left(\det A(\underline{b}^\top,1)| \det A(\underline{b}^\top,2)| \cdots | \det A(\underline{b}^\top,n)\right),$$

dove $A(\underline{b}^\top, i)$ è la matrice ottenuta da A rimpiazzando la riga i-ma con $\underline{b}^\top$.

Teoria spettrale per endomorfismi

Per studiare le relazioni di coniugazione e di similitudine, vogliamo trovare altri invarianti oltre la dimensione dell'immagine e il rango. Un primo esempio è il determinante: poiché $\det(AB) = \det(BA)$ abbiamo $\det(PAP^{-1}) = \det(AP^{-1}P) = \det(A)$, ovvero il determinante è un invariante per similitudine; quindi possiamo definire il determinante di un endomorfismo $f \in \mathrm{End}(V)$ come $\det f = \det M_{\mathcal{B}}^{\mathcal{B}}(f)$, dove $\mathcal{B}$ è un'arbitraria base di V (ben definito poiché cambiando base otteniamo matrici simili).
La costruzione di altri invarianti ha bisogno di nuovi concetti e strumenti.

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita n e sia $f \in \mathrm{End}(V)$. Per $\lambda \in \mathbb{K}$ poniamo $V_\lambda(f) = \mathrm{Ker}(\lambda id_V - f)$. Osserviamo che se $\underline{v} \in V_\lambda(f)$, allora $f(\underline{v}) = \lambda\underline{v}$, quindi $V_\lambda(f)$ è f-invariante e $f_{|_{V_\lambda(f)}} = \lambda id_{V_\lambda(f)}$.

Definizione:
$\lambda \in \mathbb{K}$ si dice *autovalore* di f se $V_\lambda(f) \neq \{\underline{0}\}$, ovvero se esiste $\underline{v} \in V$, $\underline{v} \neq \underline{0}$ tale che $f(\underline{v}) = \lambda\underline{v}$. In tal caso, $V_\lambda(f)$ si dice l'*autospazio* di f relativo a λ e i suoi elementi non nulli si dicono *autovettori* di f relativi a λ. L'insieme degli autovalori di f è detto *spettro* di f e si indica con $sp(f) \subset \mathbb{K}$.

Osservazioni:
▶ Se f e g sono endomorfismi coniugati, allora $sp(f) = sp(g)$.
Infatti, se $g = hfh^{-1}$, allora $gh = hf$ e se $\underline{v}$ è un autovettore di f relativo a λ, $g(h(\underline{v})) = h(f(\underline{v})) = h(\lambda\underline{v}) = \lambda h(\underline{v})$. Poiché $h(\underline{v}) \neq 0$, $h(\underline{v})$ è un autovettore di g relativo a λ. Quindi $sp(f) \subset sp(g)$. Usando $fh^{-1} = h^{-1}g$, abbiamo anche $sp(g) \subset sp(f)$, da cui l'uguaglianza.
▶ Più precisamente abbiamo dimostrato che $h_{|_{V_\lambda(f)}} : V_\lambda(f) \to V_\lambda(g)$ è un isomorfismo, quindi $\dim V_\lambda(f) = \dim V_\lambda(g)$ per ogni autovalore $\lambda \in sp(f) = sp(g)$.

Chiamiamo $mg(\lambda) = \dim V_\lambda(f)$ la *molteplicità geometrica* dell'autovalore λ.
Lo spettro e le molteplicità geometriche degli autovalori sono quindi invarianti per coniugazione.

Possiamo dare le stesse definizioni per una matrice $A \in M(n, \mathbb{K})$ pensandola come endomorfismo di $\mathbb{K}^n$:
$V_\lambda(A) = \mathrm{Ker}(\lambda I_n - A) \subset \mathbb{K}^n$ è un autospazio di A se è non nullo; in tal caso λ è detto autovalore di A e i suoi elementi non nulli autovettori di A relativi all'autovalore λ. L'insieme degli autovalori di A, $sp(A)$, si dice spettro di A. Un autovettore di A è dunque un $\underline{x} \in \mathbb{K}^n$ non nullo tale che esiste $\lambda \in \mathbb{K}$ tale che $A\underline{x} = \lambda\underline{x}$. La dimesione dell'autospazio $V_\lambda(A)$ è la molteplicità geometrica dell'autovalore λ. Lo spettro e le molteplicità geometriche degli autovalori sono invarianti per similitudine.

Osservazioni:
▶ $V_0(f) = \mathrm{Ker}\, f$ ed è un autospazio (e quindi $0 \in sp(f)$, cioè è un autovalore di

f) se e solo se f non è un isomorfismo. Analogamente, $V_0(A) = \operatorname{Ker} A$ ed è un autospazio (e quindi $0 \in sp(A)$, cioè è un autovalore di A) se e solo se A non è invertibile.
▶ Fissata $\mathcal{B}$ una base di V, sia $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$. Allora $sp(f) = sp(A)$.
infatti, se $f(\underline{v}) = \lambda\underline{v}$ con $\underline{v} \in V$, $\underline{v} \neq 0$, allora passando alle coordinate nella base $\mathcal{B}$, $A[\underline{v}]_{\mathcal{B}} = [f(\underline{v})]_{\mathcal{B}} = [\lambda\underline{v}]_{\mathcal{B}} = \lambda[\underline{v}]_{\mathcal{B}}$ e $[\underline{v}]_{\mathcal{B}} \neq \underline{0}$. Viceversa, se $A\underline{x} = \lambda\underline{x}$ con $\underline{x} \in \mathbb{K}^n$, $\underline{x} \neq \underline{0}$, allora posto $\underline{v} = ([\ ]_{\mathcal{B}})^{-1}(\underline{x})$, $[f(\underline{v})]_{\mathcal{B}} = A\underline{x} = \lambda\underline{x} = [\lambda\underline{v}]_{\mathcal{B}}$, quindi $\underline{v} \neq \underline{0}$ e $f(\underline{v}) = \lambda\underline{v}$.
▶ Abbiamo in effetti dimostrato che $[\ ]_{\mathcal{B}}(V_\lambda(f)) = V_\lambda(A)$, per cui la molteplicità algebrica dell'autovalore λ è la stessa sia per f che per A.

Per $A \in M(n, \mathbb{K})$, $\lambda \in sp(A)$ se e solo se $\det(\lambda I_n - A) = 0$.
Infatti, $\operatorname{Ker}(\lambda I_n - A) \neq \{\underline{0}\} \iff \dim \operatorname{Ker}(\lambda I_n - A) \geq 1 \iff \operatorname{rnk} A \leq n-1 \iff \det(\lambda I_n - A) = 0$.
Definiamo allora il *polinomio caratteristico* di $A \in M(n, \mathbb{K})$ come

$$p_A(t) = \det(tI_n - A) \in \mathbb{K}[t].$$

È chiaro che sia un polinomio, in quanto il determinante è una somma di prodotti di coefficienti della matrice. Inoltre, dalla formula con le permutazioni che prevede il prodotto di n coefficienti, uno in ogni riga e colonna, il termine di grado più alto si ottiene moltiplicando tra di loro i termini $t - \underset{i}{\underset{—}{A^{i}_{|}}}$ sulla diagonale. Quindi p_A è monico e $\deg p_A = n$.
Analogamente, il temine di grado $n-1$ si ottiene sempre moltiplicando i termini sulla diagonale, e si ottiene $-(\underset{1}{\underset{—}{A^{1}_{|}}} + \underset{2}{\underset{—}{A^{2}_{|}}} + \cdots + \underset{n}{\underset{—}{A^{n}_{|}}})$. La somma degli elementi sulla diagonale di una matrice A si chiama la *traccia* di A, $\operatorname{tr}(A)$; è immediato vedere che $\operatorname{tr} : M(n, \mathbb{K}) \to \mathbb{K}$ è lineare, ovvero $\operatorname{tr} \in (M(n, \mathbb{K}))^*$.
Il termine noto di p_A è $p_A(0) = (-1)^n \det A$.

Abbiamo quindi $p_A(t) = t^n - \operatorname{tr}(A)t^{n-1} + \cdots + (-1)^n \det A$.

Abbiamo quindi $\lambda \in sp(A)$ se e solo se λ è una radice in $\mathbb{K}$ del polinomio caratteristico p_A.
Poiché le radici di un polinomio di grado n sono al più n, $sp(A)$ è finito e contiene al più n elementi.

Osservazioni:
▶ Lo spettro può effettivametne contenere n elementi: se il campo $\mathbb{K}$ contiene almeno n elementi distinti $\lambda_1, \ldots, \lambda_n$, allora la matrice diagonale $diag(\lambda_1, \ldots, \lambda_n)$ ha spettro $\{\lambda_1, \ldots, \lambda_n\}$.
▶ D'altra parte lo spettro può essere vuoto: un esempio è dato dalla matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ che ha polinomio caratteristico $p_A(t) = t^2 + 1$ e se $A \in M(2, \mathbb{R})$ allora $sp(A) = \emptyset$ (se invece $A \in M(2, \mathbb{C})$, $sp(A) = \{i, -i\}$).

In quanto radice di p_A, ogni autovalore λ ha una molteplicità, detta *molteplicità*

algebrica dell'autovalore e indicata con $ma(\lambda)$.

Mostriamo che se A e B sono matrici simili, $A \sim B$, allora $p_A = p_B$. Infatti, scriviamo $B = PAP^{-1}$ con $P \in GL(n, \mathbb{K})$.
Mostriamo che per ogni $x \in \mathbb{K}$, $\det(xI_n - B) = \det(xI_n - A)$:
$\det(xI_n - B) = \det(xI_n - PAP^{-1}) = \det(P(xI_n - A)P^{-1}) = \det(P)\det(xI_n - A)\det(P^{-1}) = \det(xI_n - A)$.
Abbiamo quindi dimostrato che le funzioni polinomiali indotte da p_A e p_B, $p_A, p_B : \mathbb{K} \to \mathbb{K}$ sono uguali. Se $\mathbb{K}$ è infinito, se due funzioni polinomiali coincidono, i polinomi che le inducono sono uguali (la differenza ha un numero infinito di radici), e quindi abbiamo $p_A = p_B$, ma questo è falso per i campi finiti.
Possiamo però considerare l'inclusione dell'anello polinomiale nel campo delle funzioni razionali, $\mathbb{K}[t] \subset \mathbb{K}(t)$, e allora $\det(tI_n - A)$ è un usuale determinante per matrici in $M(n, \mathbb{K}(t))$, per cui la dimostrazione sopra (in cui si è usato la relazione di Binet, valida su tutti i campi) sostituendo x con l'indeterminata t si può ripetere ottenendo $p_A = p_B$ in $\mathbb{K}(t)$, e quindi anche in $\mathbb{K}[t]$.

Osserviamo che il polinomio caratteristico non è un invariante completo (ad esempio non contiene informazioni sul rango o sulle molteplicità geometriche), e che gli invarianti trovati fino ad ora non danno invarianti completi. Ad esempio se $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $p_A(t) = p_B(t) = t^4$, da cui $sp(A) = sp(B) = \{0\}$, $ma(0) = 4$, $mg(0) = 2$ per entrambe, $\operatorname{rnk} A = \operatorname{rnk} B = 2$, ma A e B non sono simili in quanto $A^2 = 0 \neq B^2$ (se fosse $B = PAP^{-1}$, $B^2 = PAP^{-1}PAP^{-1} = PA^2P^{-1} = 0$).

Osserviamo che se $B = PAP^{-1}$, allora $B^m = (PAP^{-1})^m = PA^mP^{-1}$ e quindi $B^m \sim A^m$ per ogni $m > 0$. Allo stesso modo, se due endomorfismi sono coniugati allora anche le loro potenze lo sono: $f \sim g \Rightarrow f^m \sim g^m$ per ogni $m > 0$.

Possiamo definire il polinomio caratteristico p_f di un endomorfismo $f \in \operatorname{End}(V)$ in modo analogo a come abbiamo definito il determinante di un endomorfismo, scegliendo una arbitraria base $\mathcal{B}$ di V e ponendo $p_f = p_{M_{\mathcal{B}}^{\mathcal{B}}(f)}$ (che non dipende dalla scelta della base in quanto cambiando base otteniamo matrici simili).
Inoltre, $f \sim g \Rightarrow p_f = p_g$, il polinomio caratteristico è un invariante per coniugazione.
Ad esempio, $p_{\lambda id_V}(t) = p_{\lambda I_n}(t) = (t - \lambda)^n$.

Notiamo che l'invarianza del polinomio caratteristico (per coniugazione e per similitudine) implica l'invarianza dello spettro e delle molteplicità algebriche degli autovalori. Inoltre, tutti i coefficienti del polinomio caratteristico sono invarianti, in particolare l'invarianza della traccia e l'invarianza (già nota) del determinante.

Osserviamo che $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, per cui $\mathrm{tr}(PAP^{-1}) = \mathrm{tr}(AP^{-1}P) = \mathrm{tr}(A)$ mostra direttamente l'invarianza per similitudine della traccia e permette di definire (in modo analogo al determinante) la traccia di un endomorfismo.
Infatti, se $A \in M(m,n,\mathbb{K})$, $B \in M(n,m,\mathbb{K})$, allora

$$\mathrm{tr}(AB) = \sum_{i=1}^{m} \underline{AB|}^{i}_{i} = \sum_{i=1}^{m} \underline{AB}^{i}_{i|} = \sum_{i=1}^{m}\sum_{j=1}^{n} \underline{A}^{j}_{i|}\,\underline{B}^{i}_{j|} =$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{m} \underline{B}^{i}_{j|}\,\underline{A}^{j}_{i|} = \sum_{j=1}^{n} \underline{BA}^{j}_{j|} = \sum_{j=1}^{n} \underline{BA|}^{j}_{j} = \mathrm{tr}(BA).$$

Proposizione
Dato $f \in \mathrm{End}(V)$, per ogni autovalore $\lambda \in sp(f)$, $ma(\lambda) \geq mg(\lambda)$.

Dimostrazione
Poniamo $m = mg(\lambda)$ e scegliamo una base $\{\underline{v}_1, \ldots, \underline{v}_m\}$ dell'autospazio $V_\lambda(f)$ e completiamola ad una base $\mathcal{B}$ di V. La matrice di f in questa base si suddivide in blocchi $A = M^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda I_m & B \\ 0 & C \end{pmatrix}$. Poiché il determinante di una matrice triangolare a blocchi è il prodotto dei determinanti dei blocchi, abbiamo

$$p_f(t) = p_A(t) = p_{\lambda I_m}(t)p_C(t) = (t-\lambda)^m p_C(t),$$

per cui $ma(\lambda) \geq m$ (strettamente maggiore se λ è radice di p_C). ◼

Dati $W_1, \ldots, W_k \subset V$ sottospazi, la loro somma è definita nel modo usuale:
$W_1 + \cdots + W_k = \mathrm{Span}(W_1 \cup \cdots \cup W_k) = \{\underline{w}_1 + \cdots + \underline{w}_k |\ \underline{w}_i \in W_i,\ i = 1\_k\}$

Proposizione
I seguenti fatti sono equivalenti:

1. Ogni $\underline{w} \in W_1 + \cdots + W_k$ si scrive in modo unico come $\underline{w} = \underline{w}_1 + \cdots + \underline{w}_k$ con $\underline{w}_i \in W_i$, $i = 1\_k$.
2. Dati $\underline{w}_i \in W_i$, $i = 1\_k$, tali che $\underline{w}_1 + \cdots + \underline{w}_k = \underline{0}$ allora $\underline{w}_1 = \cdots = \underline{w}_k = \underline{0}$.
3. Se $\mathcal{B}_i$ è una base di W_i, $i = 1, \ldots, k$, allora $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ (intendiamo la lista dei vettori, non l'unione (a posteriori va bene anche l'unione)) è una base di $W_1 + \cdots + W_k$.
4. $\dim(W_1 + \cdots + W_k) = \dim W_1 + \cdots + \dim W_k$.

Se una di queste condizioni è verificata (e quindi tutte lo sono), diciamo che i sottospazi sono in *somma diretta* (o che la somma dei sottospazi è diretta, o che esiste la somma diretta dei sottospazi) e scriviamo $W_1 \oplus \cdots \oplus W_k$ al posto di $W_1 + \cdots + W_k$.

Dimostrazione
$1 \Rightarrow 2$. Se qualche $\underline{w}_i$ fosse non nullo, $\underline{w}_1 + \cdots + \underline{w}_k = \underline{0} = \underline{0} + \cdots + \underline{0}$ (dove pensiamo ogni addendo $\underline{0}$ come appartenente ad un diverso W_i) sarebbero due modi diversi di scrivere $\underline{0} \in W_1 + \cdots + W_k$ ↯

$2 \Rightarrow 1$. Se un elemento $\underline{w}$ di $W_1 + \cdots + W_k$ si scrivesse in due modi diversi $\underline{w} = \underline{w}_1 + \cdots + \underline{w}_k = \underline{w}'_1 + \cdots + \underline{w}'_k$, con $\underline{w}_i, \underline{w}'_i \in W_i$, $i = 1\_k$, allora $\underline{0} = (\underline{w}_1 - \underline{w}'_1) + \cdots + (\underline{w}_k - \underline{w}'_k)$ con addendi non tutti nulli ↯.
$2 \Rightarrow 3$. Poiché gli elementi di $W_1 + \cdots + W_k$ si scrivono come somme di elementi dei W_i e ogni elemento di W_i è combinazione lineare di $\mathcal{B}_i$, $\mathcal{B}$ genera $W_1 + \cdots + W_k$. Inoltre, ogni combinazione lineare C di elementi di $\mathcal{B}$ si scrive come $C = C_1 + \cdots + C_k$ dove C_i è una combinazione lineare degli elementi di $\mathcal{B}_i$. Se quindi $C = \underline{0}$, per ipotesi ogni $C_i = \underline{0}$, e quindi tutti i coefficienti di ogni C_i sono nulli, e lo sono anche quelli di C. Quindi $\mathcal{B}$ è linearmente indipendente.
$3 \Rightarrow 2$. Se $\underline{w}_1 + \cdots + \underline{w}_k = \underline{0}$, scriviamo ogni $\underline{w}_i$ come combinazione lineare $\underline{w}_i = C_i$ degli elementi di $\mathcal{B}_i$. Allora $C_1 + \cdots + C_k$ è una combinazione lineare nulla di elementi di $\mathcal{B}$ e quindi tutti i coefficienti sono nulli, per cui ogni $C_i = \underline{0}$
$3 \Rightarrow 4$. Evidente per definizione di dimensione.
$4 \Rightarrow 3$ Se $\mathcal{B}$ non fosse linearmente indipendente, si potrebbe estrarre da $\mathcal{B}$ una base di $W_1 + \cdots + W_k$ che avrebbe un numero di elementi strettamente minore di $\dim W_1 + \cdots + \dim W_k$ ↯.

❑

Una base di $W_1 \oplus \cdots \oplus W_k$ come in 3 si dice una *base adattata alla somma diretta*.

Proposizione
Siano $\lambda_1, \ldots, \lambda_k$ autovalori distinti dell'endomorfismo f. Allora la somma dei corrispettivi autospazi è diretta.

Dimostrazione
Per induzione su k, verifichiamo la condizione 2, dove per $k = 1$ non c'è niente da dimostrare.
Sia $k > 1$ e, per $i = 1\_k$ siano $\underline{v}_i \in V_{\lambda_i}(f)$ tali che $\underline{v}_1 + \cdots + \underline{v}_k = \underline{0}$. Applicando f otteniamo $\lambda_1\underline{v}_1 + \cdots + \lambda_k\underline{v}_k = \underline{0}$. Ricavando $\underline{v}_k$ dalla prima equazione e sostituendo nella seconda abbiamo $(\lambda_1 - \lambda_k)\underline{v}_1 + \cdots + (\lambda_{k-1} - \lambda_k)\underline{v}_{k-1} = \underline{0}$. Osservando che $(\lambda_i - \lambda_k)\underline{v}_i \in V_{\lambda_i}(f)$, per l'ipotesi induttiva abbiamo $(\lambda_i - \lambda_k)\underline{v}_i = \underline{0}$ per ogni $i = 1\_k - 1$, e poiché $\lambda_k \neq \lambda_i$, $\underline{v}_i = \underline{0}$ per ogni $i = 1\_k - 1$. Dalla prima equazione allora otteniamo $\underline{v}_k = \underline{0}$. ❑

Osservazioni:
► Vale la stesa cosa per una matrice quadrata: $\lambda_1, \ldots, \lambda_k$ autovalori distinti di $A \in M(n, \mathbb{K})$, allora la somma dei corrispettivi autospazi è diretta.
► In generale avremo, se $sp(f) = \{\lambda_1, \ldots, \lambda_k\}$, $V = V_{\lambda_1}(f) \oplus \cdots \oplus V_{\lambda_k}(f) \oplus U$ e sui primi k fattori f è completamente determinata: se $\underline{v} \in V_{\lambda_i}(f)$, $f(\underline{v}_i) = \lambda_i\underline{v}_i$. Resta da scopire come trattare $f_{|_U}$, e in generale U non è f-invariante.

Endomorfismi diagonalizzabili e triangolabili

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita $\dim V = n$, $f \in \mathrm{End}(V)$ con $sp(f) = \{\lambda_1, \ldots, \lambda_k\}$.

Proposizione
I seguenti fatti sono equivalenti:

1. Esiste una base $\mathcal{B}$ di V composta da autovettori per f.
2. Esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale (si dice che $\mathcal{B}$ *diagonalizza* f).
3. V è somma diretta degli autospazi di f, $V = V_{\lambda_1}(f) \oplus \cdots \oplus V_{\lambda_k}(f)$.

Se una di queste condizioni è verificata (e quindi tutte lo sono), diciamo che f è *diagonalizzabile*.
L'insieme degli endomorfismi di V diagonalizzabili si indica con $\mathcal{D}(V)$

Dimostrazione
$1 \iff 2$.
L'equivalenza dei primi due punti deriva dal fatto che, posto $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$, allora $[\underline{v}_i]_{\mathcal{B}} = \underline{e}_i$ e quindi $f(\underline{v}_i) = \lambda \underline{v}_i \iff M_{\mathcal{B}}^{\mathcal{B}}(f)(\underline{e}_i) = \lambda \underline{e}_i$.
$1 \Rightarrow 3$.
La somma diretta degli autospazi esiste sempre. $V_{\lambda_1}(f) + \cdots + V_{\lambda_k}(f) \subset V$, ma se esiste una base $\mathcal{B}$ di V composta da autovettori per f, poiché $\mathcal{B} \subset V_{\lambda_1}(f) \cup \cdots \cup V_{\lambda_k}(f)$, allora $V = \mathrm{Span}(\mathcal{B}) \subset V_{\lambda_1}(f) + \cdots + V_{\lambda_k}(f)$, da cui l'uguaglianza.
$3 \Rightarrow 1$.
Se $V = V_{\lambda_1}(f) \oplus \cdots \oplus V_{\lambda_k}(f)$, allora per la proprietà 3 delle somme dirette, una base adattata alla somma diretta (data da una base in ogni autospazio, quindi formata da autovettori per f) è una base di V di autovettori per f. ❑

Osservazioni:
▶ Essendo espressa tramite matrici associate, la condizione di diagonalizzabilità è invariante per coniugazione:
$f, g \in \mathrm{End}(V)$, $f \sim g$ allora $f \in \mathcal{D}(V) \iff g \in \mathcal{D}(V)$.
▶ Ha senso quindi restringere la relazione di coniugazione all'insieme $\mathcal{D}(V)$ e studiarne il quoziente $\mathcal{D}(V)/\sim$.
▶ Possiamo dare una definizione analoga per le matrici quadrate:
$A \in M(n, \mathbb{K})$ si dice diagonalizzabile se è simile ad una matrice diagonale, ovvero se esiste $P \in GL(n, \mathbb{K})$ tale che PAP^{-1} è diagonale.
▶ $A \in M(n, \mathbb{K})$ è diagonalizzabile se e solo se esiste una base di $\mathbb{K}^n$ di autovettori per A.
▶ $A \in M(n, \mathbb{K})$ è diagonalizzabile se e solo se la somma diretta degli autospazi di A dà tutto $\mathbb{K}^n$.
▶ Se $A, B \in M(n, \mathbb{K})$ sono simili, A è diagonalizzabile se e solo se B lo è.

Vediamo adesso un criterio di diagonalizzabilità:

Proposizione
$f \in \mathrm{End}(V)$ è diagonalizzabile se e solo se sono verificate le seguenti:

1. il polinomimo caratteristico di f è *completamente fattorizzabile* in $\mathbb{K}[t]$ (la sua fattorizzazione in polinomi irriducibili in $\mathbb{K}[x]$ ha solo polinomi irriducibili di grado 1), ovvero $p_f(t) = \prod_{\lambda \in sp(f)} (t-\lambda)^{ma(\lambda)}$.

2. per ogni $\lambda \in sp(f)$, $mg(\lambda) = ma(\lambda)$.

Un enunciato simile vale per le matrici quadrate.

Dimostrazione
Sia $sp(f) = \{\lambda_1, \dots, \lambda_k\}$.
Se f è diagonalizzabile, sia $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ una base di V adattata alla somma diretta $V = V_{\lambda_1}(f) \oplus \cdots \oplus V_{\lambda_k}(f)$. La matrice associata ad f in tale base è diagonale a blocchi, $M_{\mathcal{B}}^{\mathcal{B}}(f) = diag(\lambda_1 I_{mg(\lambda_1)}, \dots, \lambda_k I_{mg(\lambda_k)})$, per cui

$$p_f(t) = p_{\lambda_1 I_{mg(\lambda_1)}}(t) \cdots p_{\lambda_k I_{mg(\lambda_k)}}(t) = (t-\lambda_1)^{mg(\lambda_1)} \cdots (t-\lambda_k)^{mg(\lambda_k)}.$$

Quindi p_f è completamente fattorizzabile in $\mathbb{K}[t]$ e, poiché i λ_i sono distinti, $ma(\lambda_i) = mg(\lambda_i)$ per ogni i.
Viceversa, supponiamo valgano 1 e 2. Consideriamo la somma diretta degli autospazi di f, $W = V_{\lambda_1}(f) \oplus \cdots \oplus V_{\lambda_k}(f) \subset V$, e calcoliamone la dimensione:

$$\begin{aligned}\dim W &= \dim V_{\lambda_1}(f) + \cdots + \dim V_{\lambda_k}(f) = mg(\lambda_1) + \cdots + mg(\lambda_k) = \\ &= ma(\lambda_1) + \cdots + ma(\lambda_k).\end{aligned}$$

Essendo p_f completamente fattorizzabile in $\mathbb{K}[t]$, la somma delle molteplicità delle sue radici dà il grado di p_f, che vale $\dim V$, per cui $\dim W = \dim V$ e quindi $W = V$. ∎

Osserviamo che il polinomio caratteristico di un endomorfismo diagonalizzabile ne individua (in modo unico a meno di riordinare i blocchi) una matrice associata.
Come corollario otteniamo che il polinomio caratteristico è un invariante completo per la relazione di coniugazione su $\mathcal{D}(V)$:
se $f, g \in \mathcal{D}(V)$, $f \sim g$ se e solo se $p_f = p_g$.
Il quoziente $\mathcal{D}(V)/\sim$ è quindi in bigezione con l'insieme dei polinomi monici di grado n completamente fattorizzabili in $\mathbb{K}[t]$.
Lo stesso vale per matrici diagonalizzabili e la relazione di similitudine.

Sia $W \subset V$ un sottospazio f-invariante. Allora possiamo pensare $f_{|_W}$ come endomorfismo di W e sappiamo che $p_{f_{|_W}} \mid p_f$ e quindi $sp(f_{|_W}) \subset sp(f)$. Inoltre, per ogni $\lambda \in sp(f_{|_W})$, $V_\lambda(f_{|_W}) = \{\underline{w} \in W \mid f(\underline{w}) = \lambda \underline{w}\} = W \cap V_\lambda(f)$.

Supponiamo $V = W \oplus U$ con W, U sottospazi f-invarianti.
Allora $f \in \mathcal{D}(V) \iff f_{|_W} \in \mathcal{D}(W)$, $f_{|_U} \in \mathcal{D}(U)$.
Infatti, se $f_{|_W}$ e $f_{|_U}$ sono diagonalizzabili, allora esistono basi $\mathcal{C}$, $\mathcal{D}$ di W e U rispettivamente composte da autovettori di f e la loro unione dà una base di V complosta da autovettori di f.

Viceversa, consideriamo $\underline{v} \in V$ autovettore di f, $f(\underline{v}) = \lambda\underline{v}$, $\lambda \in \mathbb{K}$. Scriviamo $\underline{v} = \underline{w} + \underline{u}$ con $\underline{w} \in W, \underline{u} \in U$. Abbiamo $\lambda\underline{u} + \lambda\underline{w} = f(\underline{v}) = f(\underline{u}) + f(\underline{w})$ ed essendo i due sottospazi f-invarianti e in somma diretta, $f(\underline{u}) = \lambda\underline{u}$, $f(\underline{w}) = \lambda\underline{w}$ (se $\underline{u} \neq 0$, $\underline{u}$ è autovettore per f, idem per $\underline{w}$).
Consideriamo quindi una base $\underline{v}_1, \dots, \underline{v}_n$ di V di autovettori di f e scriviamo $\underline{v}_i = \underline{w}_i + \underline{u}_i$ con $\underline{w}_i \in W, \underline{u}_i \in U$, $i = 1\_n$. Usando il fatto che le due proiezioni p_U, p_W date dalla somma diretta sono surgettive, $\underline{u}_1, \dots, \underline{u}_n$ generano U e $\underline{w}_1, \dots, \underline{w}_n$ generano W e da questi insiemi di generatori si possono estrarre una base di U e una di W composte da autovettori di f.

Proposizione
Sia $W \subset V$ un sottospazio f-invariante. Allora $f \in \mathcal{D}(V) \Rightarrow f_{|_W} \in \mathcal{D}(W)$.
Si dice che la proprietà di essere diagonalizzabile è *ereditaria* per i sottospazi f-invarianti.

Dimostrazione
Basta mostrare che esiste $U \subset V$ sottospazio f-invariante supplementare di W. Sia $\underline{v}_1, \dots, \underline{v}_n$ una base di V di autovettori di f e sia $\underline{w}_1, \dots, \underline{w}_m$ una base di W. Estraendo una base di V dall'insieme di generatori $\underline{w}_1, \dots, \underline{w}_m, \underline{v}_1, \dots, \underline{v}_n$ si ottiene $\underline{w}_1, \dots, \underline{w}_m, \underline{v}_{i_{m+1}}, \dots, \underline{v}_{i_n}$ per cui $U = \mathrm{Span}(\underline{v}_{i_{m+1}}, \dots, \underline{v}_{i_n})$ è un supplementare di W generato da autovettori di f e quindi è f-invariante. ∎

Osserviamo che $V = \bigoplus_{\lambda \in sp(f)} V_\lambda(f)$ implica $W = \bigoplus_{\lambda \in sp(f)} (W \cap V_\lambda(f))$, cosa falsa in generale: $V = W_1 \oplus \cdots \oplus W_k$ non implica $W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k)$ (ad esempio, tutte le intersezioni a destra potrebbero essere nulle!).

Definizione:
Una *bandiera di sottospazi* è data da n sottospazi non nulli di V, $W_1, W_2 \dots, W_n$, tali che $W_1 \subsetneqq W_2 \cdots \subsetneqq W_n$.
Osserviamo che necessariamente $\dim W_i = i$ per ogni $i = 1\_n$, in particolare $W_n = V$; a volte si pone anche $W_0 = \{\underline{0}\}$.

Ogni base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ di V dà una bandiera ponendo, per $i = 1\_n$, $W_i = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_i)$ (notare che permutando i vettori della base, la bandiera cambia).

Sia $f \in \mathrm{End}(V)$. Diciamo che una bandiera di sottospazi è f-invariante se ogni sottospazio della bandiera lo è.
Osserviamo che la bandiera data dalla base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ è f-invariante se e solo se $f(\underline{v}_i) \in W_i = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_i)$ per ogni $i = 1\_n$. Infatti, se la bandiera è f-invariante, poiché $\underline{v}_i \in W_i$, $f(\underline{v}_i) \in W_i$. Viceversa dato $\underline{w} \in W_i$, scriviamo $\underline{w} = a_1\underline{v}_1 + \cdots + a_i\underline{v}_i$, quindi $f(\underline{w}) = a_1 f(\underline{v}_1) + \cdots + a_i f(\underline{v}_i)$ e tutti i termini della somma appartengono a W_i.

Proposizione
I seguenti fatti sono equivalenti:

1. Esiste una base $\mathcal{B}$ di V la cui bandiera è f-invariante.
2. Esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è triangolare superiore (si dice che $\mathcal{B}$ *triangola* f).

Se una di queste condizioni è verificata (e quindi entrambe lo sono), diciamo che f è *triangolabile*.
L'insieme degli endomorfismi di V triangolabili si indica con $\mathcal{T}(V)$.

Dimostrazione
Osserviamo che se $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$, allora l'immagine di $W_i = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_i)$ tramite l'isomorfismo delle coordinate nella base $\mathcal{B}$ è $Z_i = \mathrm{Span}(\underline{e}_1, \dots, \underline{e}_i)$ (ovvero, $[\ ]_{\mathcal{B}}$ manda la bandiera data da $\mathcal{B}$ nella bandiera data dalla base canonica di $\mathbb{K}^n$).
L'equivalenza dei due punti deriva quindi dal fatto che, $f(\underline{v}_i) \in W_i$ per ogni $i = 1\_n \iff M_{\mathcal{B}}^{\mathcal{B}}(f)(\underline{e}_i) \in Z_i$ per ogni $i = 1\_n \iff M_{\mathcal{B}}^{\mathcal{B}}(f)$ è triangolare superiore. ◻

Osservazioni:
▶ Essendo espressa tramite matrici associate, la condizione di triangolabilità è invariante per coniugazione:
$f, g \in \mathrm{End}(V)$, $f \sim g$ allora $f \in \mathcal{T}(V) \iff g \in \mathcal{T}(V)$.
▶ Ha senso quindi restringere la relazione di coniugazione all'insieme $\mathcal{T}(V)$ e studiarne il quoziente $\mathcal{T}(V)/_{\sim}$.
▶ Possiamo dare una definizione analoga per le matrici quadrate:
$A \in M(n, \mathbb{K})$ si dice triangolabile se è simile ad una matrice triangolare superiore, ovvero se esiste $P \in GL(n, \mathbb{K})$ tale che PAP^{-1} è triangolare superiore.
▶ $A \in M(n, \mathbb{K})$ è triangolabile se e solo se esiste una base di $\mathbb{K}^n$ la cui bandiera è A-invariante.
▶ Se $A, B \in M(n, \mathbb{K})$ sono simili, A è triangolabile se e solo se B lo è.

Vediamo adesso un criterio di triagolabilità:

Proposizione
Un endomorfismo $f \in End(V)$ è triangolabile se e solo se il polinomio caratteristico p_f è completamente fattorizzabile in $\mathbb{K}[t]$.

Dimostrazione
Per una matrice $T \in M(n, \mathbb{K})$ triangolare superiore il polinomio caratteristico è $p_T(t) = \prod_{i=1}^{n}(t - \underset{i}{T}_{i})$, quindi un endomorfismo triangolabile ha polinomio caratteristico completamente fattorizzabile in $\mathbb{K}[t]$.
Viceversa, supponiamo che p_f sia completamente fattorizzabile in $\mathbb{K}[t]$ e dimostriamo che f è triangolabile per induzione su $n = \dim V$ (per $n = 1$ non c'è niente da dimostrare).
Sia quindi $n \geq 2$. Poiché p_f è completamente fattorizzabile in $\mathbb{K}[t]$, allora esiste $\lambda \in \mathbb{K}$ radice di p_f, ovvero $\lambda \in sp(f)$ è un autovalore per f. Sia allora $\underline{v}_1$ un

autovettore per f relativo a λ. Estendiamo $\underline{v}_1$ a base $\mathcal{D} = \{\underline{v}_1, \dots, \underline{v}_n\}$ di V e poniamo $\mathcal{D}' = \{\underline{v}_2, \dots, \underline{v}_n\}$, $W = \mathrm{Span}(\mathcal{D}')$, $U = \mathrm{Span}(\underline{v}_1)$. Osserviamo che $\dim W = n-1$ e che $V = U \oplus W$. Siano $p_U : V \to U$ e $p_W : V \to W$ le proiezioni su U e W data da tale somma diretta e sia $g \in \mathrm{End}(W)$ data da $g = p_W \circ f_{|W}$.
La matrice associata a f nella base $\mathcal{D}$ è una matrice a blocchi del tipo

$$M = M_{\mathcal{D}}^{\mathcal{D}}(f) = \begin{pmatrix} \lambda & A \\ 0 & B \end{pmatrix} \text{ con } A \in M(1, n-1, \mathbb{K}),\ B \in M(n-1, \mathbb{K}).$$

Osserviamo che $M_{\mathcal{D}'}^{\mathcal{D}'}(g) = B$.

Infatti abbiamo che, se $j = 2 \_ n$, $f(\underline{v}_j) = \underset{1}{\underline{M}}^{j} \underline{v}_1 + \underset{2}{\underline{M}}^{j} \underline{v}_2 + \cdots + \underset{n}{\underline{M}}^{j} \underline{v}_n$, e allora $g(\underline{v}_j) = p_W(f(\underline{v}_j)) = \underset{2}{\underline{M}}^{j} \underline{v}_2 + \cdots + \underset{n}{\underline{M}}^{j} \underline{v}_n$, dove compaiono i coefficienti della $(j-1)$-ma colonna di B ($\underset{i}{\underline{M}}^{j} = \underset{i-1}{\underline{B}}^{j-1}$, se $i, j \geq 2$).
Inoltre, $p_f(t) = p_M(t) = (t - \lambda) p_B(t)$, per cui $p_g = p_B$ è completamente fattorizzabile in $\mathbb{K}[t]$.
Per l'ipotesi induttiva, esiste una base $\mathcal{B}' = \{\underline{w}_2, \dots, \underline{w}_n\}$ di W che triangola g, e allora la base $\mathcal{B} = \{\underline{v}_1, \underline{w}_2, \dots, \underline{w}_n\}$ di V triangola f (è una base di V essendo una base adattata alla somma diretta $V = U \oplus W$).
Infatti, da $p_U + p_W = id_V$ abbiamo, per $i = 2 \_ n$, $f(\underline{w}_i) = p_U(f(\underline{w}_i)) + g(\underline{w}_i)$ e poiché $p_U(f(\underline{w}_i)) \in \mathrm{Span}(\underline{v}_1)$, $g(\underline{w}_i) \in \mathrm{Span}(\underline{w}_2, \dots, \underline{w}_i)$, $f(\underline{w}_i) \in \mathrm{Span}(\underline{v}_1, \underline{w}_2, \dots, \underline{w}_i)$. ■

Segue immediatamente dal criterio che la proprietà di essere triangolabile è ereditaria per sottospazi f-invarianti:
$f \in \mathcal{T}(V)$, $W \subset V$ sottospazio f-invariante, $\Rightarrow f_{|W} \in \mathcal{T}(W)$
infatti, $p_{f_{|W}} \mid p_f$ e se p_f è completamente fattorizzabile in $\mathbb{K}[t]$, anche $p_{f_{|W}}$ lo è.

Come ulteriore corollario otteniamo che se $\mathbb{K}$ è algebricamente chiuso (ad esempio per $\mathbb{K} = \mathbb{C}$), allora $\mathcal{T}(V) = \mathrm{End}(V)$.
In generale $\mathcal{T}(V) \neq \mathrm{End}(V)$ e la ricerca di invarianti completi è più complessa e non verrà trattata (eccetto nel caso $\mathbb{K} = \mathbb{R}$, dove gli invarianti completi si possono dedurre dagli invarianti completi su $\mathbb{C}$).

Dati $f, g \in \mathrm{End}(V)$ tali che $fg = gf$, allora $\mathrm{Ker}\, g$ e $\mathrm{Im}\, g$ sono f-invarianti.
Infatti, se $\underline{v} \in \mathrm{Ker}\, g$, $g(f(\underline{v})) = f(g(\underline{v})) = f(\underline{0}) = \underline{0}$, per cui $f(\underline{v}) \in \mathrm{Ker}\, g$. Se poi $\underline{v} \in V$, $\underline{w} = g(\underline{v}) \in \mathrm{Im}\, g$ e $f(\underline{w}) = f(g(\underline{v})) = g(f(\underline{v})) \in \mathrm{Im}\, g$.
In particolare, se f commuta con g, allora f commuta con $\lambda id_V - g$ per ogni $\lambda \in \mathbb{K}$, quindi gli autospazi di g sono f-invarianti.

Proposizione
Dati $f, g \in \mathcal{D}(V)$. Esiste una base di V di autovettori sia di f che di g se e solo se $fg = gf$.
Si dice che f e g sono *simultaneamente diagonalizzabili*.

Dimostrazione
Sia $\underline{v}_1, \dots, \underline{v}_n$ una base di V di autovettori sia per f che per g, $f(\underline{v}_i) = \lambda_i \underline{v}_i$, $g(\underline{v}_i) = \mu_i \underline{v}_i$, $\lambda_i, \mu_i \in \mathbb{K}$, $i = 1 \_\_ n$. Allora $f(g(\underline{v}_i)) = \lambda_i \mu_i \underline{v}_i = g(f(\underline{v}_i))$ e quindi $fg = gf$ poiché coincidono su una base di V.
Viceversa, se $fg = gf$, per ogni $\lambda \in sp(f)$, $V_\lambda(f)$ è g-invariante e quindi, essendo g diagonalizzabile, $g_{|_{V_\lambda(f)}}$ è diagonalizzabile. Esiste quindi una base di $V_\lambda(f)$ di autovettori di g (e di f). Adesso, $f \in \mathcal{D}(V) \Rightarrow V = \bigoplus_{\lambda \in sp(f)} V_\lambda(f)$ e in ogni autospazio di f troviamo una base di autovettori di g (e di f). La loro unione è una base di V di autovettori di g e di f come voluto. ■

Proposizione
Dati $f, g \in \mathcal{T}(V)$. Se $fg = gf$ allora esiste una base di V con bandiera sia f-invariante che g-invariante.
Si dice che f e g sono *simultaneamente triangolabili*.

Dimostrazione
Procediamo, come nel criterio di triangolabilità, per induzione su $n = \dim V$.
Per prima cosa troviamo un autovettore comune a f e g. p_f è completamente fattorizzabile in $\mathbb{K}[t]$, per cui esiste una radice $\lambda \in \mathbb{K}$ di p_f, ovvero $\lambda \in sp(f)$. L'autospazio $V_\lambda(f)$ è g invariante ed essendo g triangolabile, $g_{|_{V_\lambda(f)}}$ è triangolabile, e quindi ammette un autovettore $\underline{v}_1$ che è dunque autovettore sia per f che per g.
Sia W un supplementare di $U = \mathrm{Span}(\underline{v}_1)$ e siano $h = p_W \circ f_{|_W}$, $k = p_W \circ g_{|_W}$. $h, k \in \mathcal{T}(W)$ ed inoltre commutano tra di loro.
Infatti, se $\underline{w} \in W$, $g(f(\underline{w})) = g(p_U(f(\underline{w})) + p_W(f(\underline{w}))) = \alpha v_1 + g(h(\underline{w}))$ per qualche $\alpha \in \mathbb{K}$. Proiettando su W si ha $p_W(g(f(\underline{w}))) = k(h(\underline{w}))$. In modo analogo, $p_W(f(g(\underline{w}))) = h(k(\underline{w}))$. Poiché $gf = fg$, $hk = kh$.
Possiamo allora applicare l'ipotesi induttiva a h e k, per cui esiste una base $\underline{w}_2, \dots, \underline{w}_n$ di W con bandiera sia h-invariante che k-invariante. Si verifica, come nel criterio di triangolabilità, che la base di V $\underline{v}_1, \underline{w}_2, \dots, \underline{w}_n$ ha bandiera sia f-invariante che g-invariante. ■.

Il polinomio minimo di un endomorfismo

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita, $\dim V = n$, e sia f un endomorfismo di V.

Dato $p \in \mathbb{K}[t]$, $p(t) = a_d t^d + \cdots + a_0 t^0$, definiamo

$$p(f) = a_d f^d + \cdots + a_0 f^0,$$

dove conveniamo che $f^0 = id_V$.
Notare che $p(f)$ è ancora un endomorfismo di V. Si dice che $p(f)$ si ottiene *valutando* p su f e che $p(f)$ è *polinomiale* in f.

Abbiamo quindi una applicazione

$$val_f : \mathbb{K}[t] \to \mathrm{End}(V), \quad p \mapsto p(f),$$

detta *valutazione su* f, con immagine gli endomorfismi polinomiali in f.

È immediato vedere che $(p_1 + p_2)(f) = p_1(f) + p_2(f)$ e $(p_1 p_2)(f) = p_1(f)p_2(f)$ per ogni $p_1, p_2 \in \mathbb{K}[t]$, e quindi che val_f è un omomorfismo di anelli (in particolare è lineare).

Dunque, l'immagine di val_f, denotata con $\mathbb{K}[f]$, è un sottospazio di $\mathrm{End}(V)$ che è anche un sottoanello commutativo, mentre il nucleo di val_f, denotato con $I(f)$, è un ideale di $\mathbb{K}[t]$:

$$\mathbb{K}[f] = \mathrm{Im}\, val_f = \mathrm{Span}(id_V, f, f^2, f^3, \ldots),$$

$$I(f) = \mathrm{Ker}\, val_f = \{p \in \mathbb{K}[t] |\ p(f) = 0 \in \mathrm{End}(V)\}.$$

Mostriamo direttamente che $I(f)$ è un ideale:
- $0(f) = 0 \in \mathrm{End}(V)$, per cui $0 \in I(f)$;
- se $p, q \in I(f)$, allora $(p+q)(f) = p(f) + q(f) = 0 + 0 = 0 \in \mathrm{End}(V)$, per cui $p + q \in I(f)$;
- se $p \in I(f)$, $q \in \mathbb{K}[t]$, allora $(pq)(f) = p(f)q(f) = 0q(f) = 0 \in \mathrm{End}(V)$, per cui $pq \in I(f)$.

Notiamo che $I(f)$ è un ideale proprio, in quanto $1(f) = id_V \neq 0$.
Inoltre, $I(f) \neq \{0\}$.
Infatti, $\dim \mathrm{End}(V) = n^2$ e quindi $id_V, f, f^2, \ldots, f^{n^2}$ ($n^2 + 1$ endomorfismi) non sono linearmente indipendenti. Esistono quindi $a_0, a_1, \ldots, a_{n^2} \in \mathbb{K}$ non tutti nulli tali che $a_0 id_V + a_1 f + \cdots + a_{n^2} f^{n^2} = 0 \in \mathrm{End}(V)$ per cui il polinomio $a_0 + a_1 t + \cdots + a_{n^2} t^{n^2}$ è non nullo ed appartiene a $I(f)$.
Poiché che gli ideali di $\mathbb{K}[t]$ sono principali, otteniamo che, esiste un unico polinomio monico μ_f, di grado almeno 1, tale che $I(f) = (\mu_f)$.
Ricordiamo che μ_f è l'unico polinomio monico tra i polinomi di grado minimo e positivo in $I(f)$ e che divide tutti i polinomi dell'ideale.

Definizione:

Con le notazioni di sopra, $I(f)$ si dice l'*ideale di* f, μ_f si dice il *polinomio minimo* di f.

Osserviamo che la definizione di ideale di un endomorfismo ha senso anche nel caso V abbia dimensione infinita. È però possibile che tale ideale sia nullo e che quindi non si possa definire un polinomio minimo.

Possiamo dare le stesse definizioni per una matrice $A \in M(n, \mathbb{K})$:
dato $p \in \mathbb{K}[t]$, $p(t) = a_d t^d + \cdots + a_0 t^0$, definiamo $p(A) = a_d A^d + \cdots + a_0 A^0$ (dove $A^0 = I_n$); otteniamo l'omomorfismo di anelli $val_A : \mathbb{K}[t] \to M(n, \mathbb{K})$, $A \mapsto p(A)$, la cui immagine è $\mathbb{K}[A] \subset M(n, \mathbb{K})$, data dalle matrici polinomiali in A, e il cui nucleo è $I(A)$, l'ideale di A; il generatore monico di $I(A)$, μ_A, è il polinomio minimo di A.

Fissata $\mathcal{B}$ una base di V, $M_{\mathcal{B}}^{\mathcal{B}} : \mathrm{End}(V) \to M(n, \mathbb{K})$ è un isomorfismo di anelli, per cui $M_{\mathcal{B}}^{\mathcal{B}}(p(f)) = p(M_{\mathcal{B}}^{\mathcal{B}}(f))$ per ogni $f \in \mathrm{End}(V)$ e per ogni $p \in \mathbb{K}[t]$.
Abbiamo il diagramma commutativo

$$\begin{array}{ccc} \mathbb{K}[t] & \xrightarrow{id_{\mathbb{K}[t]}} & \mathbb{K}[t] \\ {\scriptstyle val_f}\downarrow & \circlearrowright & \downarrow{\scriptstyle val_f} \\ \mathrm{End}(V) & \xrightarrow{M_{\mathcal{B}}^{\mathcal{B}}} & M(n, \mathbb{K}) \end{array}$$

per cui, per ogni $f \in \mathrm{End}(V)$, posta $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$, $\mathbb{K}[f]$ e $\mathbb{K}[A]$ sono isomorfi (tramite $M_{\mathcal{B}}^{\mathcal{B}}$), e $I(f) = I(A)$. In particolare $\mu_f = \mu_A$.

Osserviamo che dato un endomorfismo invertibile $h \in GL(V)$, allora per ogni $f, g \in \mathrm{End}(V)$ si ha $h(f+g)h^{-1} = hfh^{-1} + hgh^{-1}$ e $hfgh^{-1} = (hfh^{-1})(hgh^{-1})$.
La coniugazione per h definisce un isomorfismo di anelli $C_h : \mathrm{End}(V) \to \mathrm{End}(V)$, $f \mapsto hfh^{-1}$. In particolare, $hf^k h^{-1} = (hfh^{-1})^k$ per ogni k, e dunque, per ogni polinomio $p \in \mathbb{K}[t]$, $p(hfh^{-1}) = hp(f)h^{-1}$.
Abbiamo il diagramma commutativo

$$\begin{array}{ccc} \mathbb{K}[t] & \xrightarrow{id_{\mathbb{K}[t]}} & \mathbb{K}[t] \\ {\scriptstyle val_f}\downarrow & \circlearrowright & \downarrow{\scriptstyle val_{hfh^{-1}}} \\ \mathrm{End}(V) & \xrightarrow{C_h} & \mathrm{End}(V) \end{array}$$

da cui otteniamo che, se due endomorfismi $f, g \in \mathrm{End}(V)$ sono coniugati, $g \sim f$, allora $\mathbb{K}[f]$ e $\mathbb{K}[g]$ sono isomorfi e $I(f) = I(g)$. In particolare $\mu_f = \mu_g$.

Lo stesso vale per matrici simili: $A, B \in M(n, \mathbb{K})$, $A \sim B$, allora $\mathbb{K}[A]$ e $\mathbb{K}[B]$ sono isomorfi e $I(A) = I(B)$, $\mu_A = \mu_B$.

Abbiamo ottenuto un nuovo invariante per coniugazione (o similitudine) indipendente da quelli trovati fino ad ora: il polinomio minimo (ovvero l'ideale) di un endomorfismo (o di una matrice quadrata).

Vediamo il legame tra polinomio caratteristico e polinomio minimo:

Teorema (Hamilton-Cayley)
Sia V uno spazio vettoriale di dimensione finita e $f \in \mathrm{End}(V)$ un endomorfismo. Allora $p_f \in I(f)$, ovvero $\mu_f | \, p_f$.

Dimostrazione
Vogliamo dimostrare che $p_f(f)$ è l'endomorfismo nullo, ovvero $\operatorname{Ker} p_f(f) = V$.

Supponiamo dapprima che f sia triangolabile.
Sia $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ una base di V con bandiera f-invariante. Notiamo che i sottospazi della bandiera, $W_i = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_i)$, sono invarianti per gli endomorfismi del tipo $f - \lambda id_V$, $\lambda \in \mathbb{K}$.
Sia $T = M_{\mathcal{B}}^{\mathcal{B}}(f)$ la matrice di f nella base $\mathcal{B}$, che è triangolare superiore, e siano $\lambda_j = \underset{j-}{\overset{j}{T}}|$, $j = 1\_\_n$, gli elementi sulla diagonale di T (che danno lo spettro di f).
Il polinomio carateristico di f è $p_f(t) = p_T(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ e quindi $p_f(f) = (f - \lambda_1 id_V) \cdots (f - \lambda_n id_V)$.
Per $i = 1\_\_n$, poniamo $g_i = (f - \lambda_1 id_V) \cdots (f - \lambda_i id_V)$ e mostriamo per induzione su i che $W_i \subset \operatorname{Ker} g_i$; per $i = n$ avremo quindi $V = W_n \subset \operatorname{Ker} g_n = \operatorname{Ker} p_f(f)$ come voluto.
Per $i = 1$ abbiamo $f(\underline{v}_1) = \lambda_1 \underline{v}_1$, ovvero $g_1(\underline{v}_1) = \underline{0}$ e quindi si ha $W_1 \subset \operatorname{Ker} g_1$.
Se $i > 1$, supponiamo che, $W_{i-1} \subset \operatorname{Ker} g_{i-1}$.
Abbiamo $f(\underline{v}_i) = \lambda_i \underline{v}_i + \underline{w}_{i-1}$ con $\underline{w}_{i-1} \in W_{i-1}$, per cui $(f - \lambda_i id_V)(\underline{v}_i) = \underline{w}_{i-1}$.
Ne segue che $g_i(\underline{v}_i) = (g_{i-1}(f - \lambda_i id_V))(\underline{v}_i) = g_{i-1}(\underline{w}_{i-1}) = \underline{0}$.
Inoltre, W_{i-1} è invariante per $f - \lambda_i id_V$, per cui

$$g_i(W_{i-1}) = (g_{i-1}(f - \lambda_i id_V))(W_{i-1}) \subset g_{i-1}(W_{i-1}) = \{\underline{0}\},$$

quindi $W_{i-1} \subset \operatorname{Ker} g_i$. Dunque $W_i = W_{i-1} \oplus \mathrm{Span}(\underline{v}_i) \subset \operatorname{Ker} g_i$.

Supponiamo adesso che f non sia triangolabile.
Allora il suo polinomio caratteristico non è completamente fattorizzabile in $\mathbb{K}[t]$, ovvero p_f ammette un fattore irriducibile q di grado almeno 2. Scegliendo un tale fattore iriducibile q di grado massimo e passando al campo $K = \mathbb{K}[t]/(q)$ che estende $\mathbb{K}$, il polinomio q come elemento di $K[t]$ non è più irriducibile e la fattorizzazione di p_f come elemento di $K[t]$ ha un fattore di grado massimo in meno.
Iterando un numero finito di volte otteniamo un campo $\mathcal{K}$, detto *campo di spezzamento di* p_f, per cui p_f considerato come elemento di $\mathcal{K}[t]$ è completamente

fattorizzabile.
Fissata $\mathcal{B}$ una base di V e posta $M = M_{\mathcal{B}}^{\mathcal{B}}(f)$, notiamo che $p_f = p_M$ e $p_f(f) = 0 \iff P_M(M) = 0$.
Poiché $\mathbb{K} \subset \mathcal{K}$, $M(n, \mathbb{K}) \subset M(n, \mathcal{K})$, per cui possiamo pensare M come matrice a coefficienti in $\mathcal{K}$. Notiamo che il polinomio caratteristico di M non cambia se estendiamo il campo, e che come endomorfismo su $\mathcal{K}^n$, M è triangolabile.
Quindi, $P_M(M) = 0 \in M(n, \mathcal{K})$, ma $P_M(M) \in M(n, \mathbb{K})$, e quindi $P_M(M) = 0$ anche in $M(n, \mathbb{K})$. ❑

Dato $\lambda \in sp(f)$, sia $\underline{v} \in V$ un autovettore di f relativo a λ.
Da $f^k(\underline{v}) = \lambda^k \underline{v}$ per ogni $k \geq 0$, otteniamo $p(f)(\underline{v}) = p(\lambda)\underline{v}$ per ogni polinomio $p \in \mathbb{K}[t]$.
In particolare, se $p \in I(f)$ (ricordando che un autovettore è un vettore non nullo), $p(\lambda) = 0$, ovvero $sp(f) \subset \{\text{radici in } \mathbb{K} \text{ di } p\}$ per ogni $p \in I(f)$.

Per il polinomio minimo abbiamo (usando H-C):

$$sp(f) \subset \{\text{radici in } \mathbb{K} \text{ di } \mu_f\} \subset \{\text{radici in } \mathbb{K} \text{ di } p_f\} = sp(f),$$

da cui $sp(f) = \{\text{radici in } \mathbb{K} \text{ di } \mu_f\}$.
Ad esempio, se f è triangolabile, scrivendo $p_f(t) = \prod_{\lambda \in sp(f)} (t - \lambda)^{m_\lambda}$ (dove m_λ è la molteplicità algebrica dell'autovalore λ), allora $\mu_f(t) = \prod_{\lambda \in sp(f)} (t - \lambda)^{r_\lambda}$ con $1 \leq r_\lambda \leq m_\lambda$ (ovvero, non si perdono fattori irriducibili).
Questo è vero in generale: il polinomio minimo e il polinomo caratteristico hanno gli stessi fattori irriducibili.
Per dimostrarlo ci servono alcuni risultati sui polinomi minimi di alcune restrizioni di f.

Sia $W \subset V$ un sottospazio f-invariante.
Osserviamo che, in generale, se $g \in \mathrm{End}(V)$

$$(g \circ f)_{|W} = g \circ f_{|W} = g_{|f(W)} \circ f_{|W}$$

Essendo W f-invariante, $f(W) \subset W$ e allora $(g \circ f)_{|W} = g_{|W} \circ f_{|W}$.
Quindi, per ogni intero $k \geq 0$, $(f^k)_{|W} = (f_{|W})^k$ (inoltre $(f+g)_{|W} = f_{|W} + g_{|W}$ e se $\lambda \in \mathbb{K}$, $(\lambda f)_{|W} = \lambda f_{|W}$).
Otteniamo che se $p \in \mathbb{K}[t]$, allora $p(f_{|W}) = p(f)_{|W}$ e quindi

$$\begin{aligned} I(f_{|W}) &= \left\{ p \in \mathbb{K}[t] \mid p(f_{|W}) = 0 \right\} \\ &= \left\{ p \in \mathbb{K}[t] \mid p(f)_{|W} = 0 \right\} \\ &= \{ p \in \mathbb{K}[t] \mid p(f)(W) = \{\underline{0}\} \} \end{aligned}$$

Otteniamo $I(f) \subset I(f_{|W})$ per cui $\mu_{f_{|W}} \mid \mu_f$.

Come caso particolare, per $\underline{v} \in V$, $\underline{v} \neq \underline{0}$, il sottospazio $W = \mathbb{K}[f](\underline{v})$ generato dalle potenze di f applicate a $\underline{v}$,

$$\mathbb{K}[f](\underline{v}) = \mathrm{Span}(\underline{v}, f(\underline{v}), f^2(\underline{v}), \ldots)$$

è un sottospazio f-invariante.
Mostriamo che W ammette una base del tipo $\underline{v}, f(\underline{v}), \ldots, f^k(\underline{v})$.
Infatti, sia k il massimo degli indici h tali che $\underline{v}, f(\underline{v}), \ldots, f^h(\underline{v})$ sono linearmente indipendenti (notiamo che $k < n$). Allora $f^{k+1}(\underline{v}) = a_0\underline{v} + a_1 f(\underline{v}) + \cdots + a_k f^k(\underline{v})$ per opportuni $a_0, \ldots, a_k \in \mathbb{K}$.
Si conclude se mostriamo che $f^{k+j} \in \mathrm{Span}(\underline{v}, f(\underline{v}), \ldots, f^k(\underline{v}))$ per $j \geq 1$.
Per induzione (per $j = 1$ segue dalla definizione di k), supponiamo che per ogni $i = 1\_j - 1$, $f^{k+i}(\underline{v}) \in \mathrm{Span}(\underline{v}, f(\underline{v}), \ldots, f^k(\underline{v}))$. Allora applicando f^{j-1} alla relazione sopra otteniamo:

$$f^{k+j}(\underline{v}) = a_0 f^{j-1}(\underline{v}) + a_1 f^j(\underline{v}) + \cdots + a_k f^{k+j-1}(\underline{v}) \in \mathrm{Span}(\underline{v}, f(\underline{v}), \ldots, f^k(\underline{v})).$$

Quindi, con le notazioni sopra, $\dim W = k + 1$.

Notiamo che $p \in I(f_{|_W})$ se e solo se $p(f)(\underline{v}) = \underline{0}$.
Infatti, se $p \in I(f_{|_W})$, allora $p(f)(W) = \{\underline{0}\}$ e quindi $p(f)(\underline{v}) = \underline{0}$.
Viceversa, se $p(f)(\underline{v}) = \underline{0}$, poiché $p(f)(f^j(\underline{v})) = f^j(p(f)(\underline{v}))$ per ogni j, $p(f)$ annulla tutti i generatori di W, e quindi $p(f)(W) = \{\underline{0}\}$, cioè $p \in I(f_{|_W})$.

Essendo $\underline{v}, f(\underline{v}), \ldots, f^k(\underline{v})$ linearmente indipendenti, nessun polinomio non nullo di grado minore a $k + 1$ può appartenere a $I(f_{|_W})$: se $p \in I(f_{|_W})$ ha grado $\deg p = h \leq k$, scrivendo $p(t) = b_0 + \cdots + b_h t^h$, $b_0, \ldots, b_h \in \mathbb{K}$, allora $\underline{0} = p(f)(\underline{v}) = b_0\underline{v} + \cdots + b_h f^h(\underline{v})$ da cui essendo $\underline{v}, \ldots, f^h(\underline{v})$ linearmente indipendenti, $b_0 = \cdots = b_h = 0$, cioè $p = 0$.
Inoltre $t^{k+1} - a_k t^k - \cdots - a_1 t - a_0 \in I(f_{|_W})$, per cui il polinomio minimo di $f_{|_W}$ è proprio $t^{k+1} - a_k t^k - \cdots - a_1 t - a_0$. In particolare, $\deg \mu_{f_{|_W}} = k+1 = \deg p_{f_{|_W}}$.
Essendo i due polinomi monici, e poiché $\mu_{f_{|_W}} \mid p_{f_{|_W}}$, $\mu_{f_{|_W}} = p_{f_{|_W}}$.

Sia adesso $U \subset V$ un sottospazio tale che $V = W \oplus U$.
Sia $\pi_U \in \mathrm{End}(V)$ la proiezione su U data dalla decomposizione in somma diretta e sia $g = \pi_U \circ f_{|_U} \in \mathrm{End}(U)$.
Osserviamo che

$$\pi_U \circ f \circ \pi_U = \pi_U \circ f$$

Infatti, dato $\underline{v} \in V$ esistono $\underline{w} \in W$ e $\underline{u} \in U$ tali che $\underline{v} = \underline{w} + \underline{u}$, per cui

$$(\pi_U \circ f \circ \pi_U)(\underline{v}) = \pi_U(f(\underline{u}))$$

$$(\pi_U \circ f)(\underline{v}) = \pi_U(f(\underline{w}) + f(\underline{u})) = \pi_U(f(\underline{u}))$$

poiché $f(\underline{w}) \in W$.

Allora, per ogni intero $k \geq 0$,

$$\begin{aligned} g^k &= (\pi_U \circ f_{|_U})^k \\ &= \underbrace{\pi_U \circ f \circ \cdots \circ \pi_U \circ f}_{k-1} \circ \pi_U \circ f_{|_U} \\ &= \underbrace{\pi_U \circ f \circ \cdots \circ \pi_U \circ f}_{k-2} \circ \pi_U \circ f \circ f_{|_U} \\ &= \underbrace{\pi_U \circ f \circ \cdots \circ \pi_U \circ f}_{k-3} \circ \pi_U \circ f^2 \circ f_{|_U} \\ &= \cdots \\ &= \pi_U \circ f^{k-1} \circ f_{|_U} = \pi_U \circ (f^k)_{|_U} \end{aligned}$$

Quindi, se $p \in \mathbb{K}[t]$, $p(g) = \pi_U \circ p(f)_{|_U}$, per cui

$$\begin{aligned} I(g) &= \{p \in \mathbb{K}[t] \mid p(g) = 0\} \\ &= \left\{p \in \mathbb{K}[t] \mid \pi_U \circ p(f)_{|_U} = 0\right\} \\ &= \{p \in \mathbb{K}[t] \mid \pi_U(p(f)(U)) = \{0\}\} \\ &= \{p \in \mathbb{K}[t] \mid p(f)(U) \subset \operatorname{Ker} \pi_U = W\} \end{aligned}$$

Otteniamo $I(f) \subset I(g)$ per cui $\mu_g \mid \mu_f$.

Proposizione
I fattori irriducibili del polinomio caratteristico e del polinomio minimo di un endomorfismo $f \in \operatorname{End}(V)$ coincidono.

Dimostrazione
Sia $q \in \mathbb{K}[t]$ irriducibile.
Se $q \mid \mu_f$, poiché $\mu_f \mid p_f$, allora $q \mid p_f$.
Dimostriamo il viceversa per induzione su $\dim V = n$ (vero per $n = 1$).
Supponiamo quindi $n \geq 2$ e $q \mid p_f$. Sia $\underline{v} \in V$, $\underline{v} \neq 0$, e, con le notazioni precedenti, sia U un supplementare del sottospazio f-invariante $W = \mathbb{K}[f](\underline{v})$ e sia $g = \pi_U \circ f_{|_U}$.
Fissiamo $\mathcal{C}$ una base di W e $\mathcal{D}$ una base di U, otteniamo una base $\mathcal{B}$ di V tale che la matrice di f in tale base è del tipo:

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

dove A è la matrice di $f_{|_W}$ nella base $\mathcal{C}$, e D è la matrice di g nella base $\mathcal{D}$.
$p_f = p_A p_D = p_{f_{|_W}} p_g$, quindi $q \mid p_{f_{|_W}}$ oppure $q \mid p_g$.
Nel primo caso, $q \mid p_{f_{|_W}} = \mu_{f_{|_W}} \mid \mu_f$. Nel secondo caso, per l'ipotesi induttiva, $q \mid \mu_g \mid \mu_f$ ■

Osservazioni:
▶ Dati $p_1, \dots, p_k \in \mathbb{K}[t]$ non nulli, il loro *minimo comune multiplo*, denotato con $mcm(p_1, \cdots, p_k)$, è il generatore monico dell'ideale $(p_1) \cap \cdots \cap (p_k)$. Osserviamo che il minimo comune multiplo di $p_1, \dots, p_k$ è un multiplo di ogni p_i e divide tutti i multipli comuni di $p_1, \dots, p_k$ (e in effetti, queste due proprietà insieme al fatto di essere monico lo caratterizzano completamente). In termini delle fattorizzazioni in fattori irriducibili, un irriducibile q compare nella fattorizzazione del minimo comune multiplo se e solo se q compare nella fattorizzazione di qualche p_i e, in tal caso, q compare elevato alla potenza massima con cui compare nelle fattorizzazioni dei p_i.
▶ Se $V = W + U$ con W, U sottospazi f-invarianti, allora se $p \in I(f_{|_W})$, $q \in I(f_{|_U})$ allora $mcm(p, q) \in I(f)$.
Infatti, poniamo $m = mcm(p, q) \in \mathbb{K}[t]$ per cui si ha $m = p_1 p$, $m = q_1 q$ con $p_1, q_1 \in \mathbb{K}[t]$. Dato $\underline{v} \in V$, scriviamo $\underline{v} = \underline{w} + \underline{u}$, con $\underline{w} \in W$ e $\underline{u} \in U$, e allora

$$\begin{aligned} m(f)(\underline{v}) &= m(f)(\underline{w}) + m(f)(\underline{u}) = \\ &= p_1(f)(p(f)(\underline{w})) + q_1(f)(q(f)(\underline{u})) = \\ &= \underline{0} + \underline{0} = \underline{0}. \end{aligned}$$

▶ Se $V = W + U$ con W, U sottospazi f-invarianti, allora il polinomio minimo di f è il minimo comune multiplo dei polinomi minimi delle restrizioni $f_{|_W}$, $f_{|_U}$.
Infatti, per l'osservazione precedente, $\mu_f | \, mcm(\mu_{f_{|_W}}, \mu_{f_{|_U}})$ e poiché $\mu_{f_{|_W}} | \, \mu_f$ e $\mu_{f_{|_U}} | \, \mu_f$, allora $mcm(\mu_{f_{|_W}}, \mu_{f_{|_U}}) | \, \mu_f$. Dunque, esendo entrambi monici, $\mu_f = mcm(\mu_{f_{|_W}}, \mu_{f_{|_U}})$.
▶ Le due osservazioni precedenti si estendono al caso in cui V è somma di un numero finito di sottospazi f-invarianti.
▶ Se $q \in \mathbb{K}[t]$ è un divisore proprio non costante del polinomio minimo di f, $q | \mu_f$, $0 < \deg q < \deg \mu_f$, allora $\operatorname{Ker} q(f) \neq \{\underline{0}\}$.
Infatti, altrimenti $q(f)$ sarebbe invertibile e scrivendo $\mu_f = pq$ e valutando su f avremmo $0 = \mu_f(f) = p(f)q(f)$, da cui $p(f) = 0$. Quindi $p \in I(f)$, ma $0 < \deg p < \deg \mu_f$ ↯.
▶ In generale, dato $q \in \mathbb{K}[t]$, $q(f)$ è invertibile se e solo se $(q, \mu_f) = 1$.
Infatti, se $(q, \mu_f) = 1$, scriviamo $1 = qp_1 + \mu_f p_2$ per opportuni $p_1, p_2 \in \mathbb{K}[t]$. Valutando su f abbiamo $id_V = q(f)p_1(f)$, per cui $p_1(f) = q(f)^{-1}$. Viceversa, se $q(f)$ è invertibile, sia $m = (q, \mu_f)$. Poiché $m | q$, $m(f)$ è invertibile (se $q = q_1 m$ con $q_1 \in \mathbb{K}[t]$, $q_1(f)m(f) = q(f)$, per cui $\operatorname{Ker} m(f) \subset \operatorname{Ker} q(f)$), e poiché $m | \mu_f$, $m = 1$ (se $\mu_f = q_2 m$ con $q_2 \in \mathbb{K}[t]$, come prima $q_2 \in I(f)$).

Dato $\underline{v} \in V$, $\underline{v} \neq \underline{0}$, possiamo comporre la valutazione su f con la valutazione su $\underline{v}$ ed ottenere: $val_{f,\underline{v}} : \mathbb{K}[t] \to V$, $p \mapsto p(f)(\underline{v})$ lineare.
L'immagine di $val_{f,\underline{v}}$ è il sottospazio di V f-invariante $\mathbb{K}[f](\underline{v})$.
Il nucleo di $val_{f,\underline{v}}$ è un ideale di $\mathbb{K}[t]$, detto *ideale relativo* di f e $\underline{v}$ indicato con $I(f, \underline{v})$. Osserviamo che $I(f) \subset I(f, \underline{v})$, per cui l'ideale relativo è un ideale non nullo e proprio ($val_{f,\underline{v}}(1) = \underline{v} \neq 0$) che quindi è generato da un polinomio monico $\mu_{f,\underline{v}}$ detto *polinomio minimo relativo* di f e $\underline{v}$. Si ha $\mu_{f,\underline{v}} | \, \mu_f$.

Per quanto visto prima, l'ideale relativo coincide con l'ideale della restrizione di f a $\mathbb{K}[f](\underline{v})$, $I(f,\underline{v}) = I(f_{|_{\mathbb{K}[f](\underline{v})}})$, per cui $\mu_{f,\underline{v}} = \mu_{f_{|_{\mathbb{K}[f](\underline{v})}}} = p_{f_{|_{\mathbb{K}[f](\underline{v})}}}$ e $\deg(\mu_{f,\underline{v}}) = \dim \mathbb{K}[f](\underline{v})$.
Inoltre, abbiamo visto che $\mathbb{K}[f](\underline{v})$ ammette una base del tipo $\underline{v}, f(\underline{v}), \dots, f^k(\underline{v})$, $k+1 = \deg \mu_{f,\underline{v}}$, e l'espressione di $f^{k+1}(\underline{v})$ in termini di tale base dà esplicitamente il polinomio minimo relativo:
$f^{k+1}(\underline{v}) = a_0\underline{v} + \dots + a_k f^k(\underline{v}) \iff \mu_{f,\underline{v}} = t^{k+1} - a_k t^k - \dots - a_0$.

Proposizione
Dati $\underline{v}_1, \dots, \underline{v}_k$ generatori non nulli di V, allora il polinomio minimo di f è il minimo comune multiplo dei polinomi minimi relativi di $\underline{v}_1, \dots, \underline{v}_k$:

$$\mu_f = mcm(\mu_{f,\underline{v}_1}, \dots, \mu_{f,\underline{v}_k}).$$

Dimostrazione
Chiamiamo m tale minimo comune multiplo (ricordando che è monico). Poiché $\mu_{f,\underline{v}_i} | \, \mu_f$ per ogni i, μ_f è un multiplo comune dei $\mu_{f,\underline{v}_i}$ e quindi $m | \, \mu_f$.
Poiché $\mu_{f,\underline{v}_i} | \, m$, $m \in I(f,\underline{v}_i)$ per ogni i. Dato $\underline{v} \in V$, scriviamo per opportuni $a_i \in \mathbb{K}$, $\underline{v} = a_i\underline{v}_1 + \dots + a_k\underline{v}_k$.
Allora, $m(f)(\underline{v}) = \sum_{i=1}^{k} a_i m(f)(\underline{v}_i) = \underline{0}$ e quindi $m \in I(f)$, da cui $\mu_f | m$.
La doppia divisibilità e il fatto che siano entrambi monici ci dà $m = \mu_f$. ∎

Vogliamo dimostrare che esiste $\underline{v} \in V$ tale che $\mu_{f,\underline{v}} = \mu_f$. Per la dimostrazione, premettiamo due lemmi.

Lemma1
Dato $\underline{v} \in V$, supponiamo $\mu_{f,\underline{v}} = p_1 p_2$ con $p_1, p_2 \in \mathbb{K}[t]$ monici e coprimi. Poniamo $\underline{w}_1 = p_1(f)(\underline{v})$, $\underline{w}_2 = p_2(f)(\underline{v})$. Allora $\mu_{f,\underline{w}_1} = p_2$, $\mu_{f,\underline{w}_2} = p_1$.

Dimostrazione
Vediamolo per $\underline{w}_1$ e mostriamo le due divisibilità $\mu_{f,\underline{w}_1} | \, p_2$, $p_2 | \, \mu_{f,\underline{w}_1}$: si conclude poiché i polinomi sono entrambi monici.
Mostriamo che p_2 appartiene all'ideale relativo $I(f,\underline{w}_1)$ e quindi $\mu_{f,\underline{w}_1} | \, p_2$. Infatti, $p_2(f)(\underline{w}_1) = p_2(f)(p_1(f)(\underline{v})) = (p_2(f)p_1(f))(\underline{v}) = \mu_{f,\underline{v}}(f)(\underline{v}) = \underline{0}$.
Per l'identità di Bezout, esistono $h_1, h_2 \in \mathbb{K}[t]$ tali che $1 = h_1 p_1 + h_2 p_2$ da cui $\underline{v} = h_1(f)(p_1(f)(\underline{v})) + h_2(f)(p_2(f)(\underline{v})) = h_1(f)(\underline{w}_1) + p_2(f)(h_2(f)(\underline{v}))$. Applicando $(\mu_{f,\underline{w}_1} p_1)(f)$ ad entrambi i membri otteniamo $(\mu_{f,\underline{w}_1} p_1)(f)(\underline{v}) = \underline{0}$, ovvero $\mu_{f,\underline{w}_1} p_1 \in I(f,\underline{v})$. Quindi, $p_1 p_2 = \mu_{f,\underline{v}} | \, \mu_{f,\underline{w}_1} p_1$, ovvero $p_2 | \, \mu_{f,\underline{w}_1}$. ∎

Lemma2
Per ogni $\underline{v}, \underline{w} \in V$ esiste $\underline{z} \in V$ tale che $\mu_{f,\underline{z}} = mcm(\mu_{f,\underline{v}}, \mu_{f,\underline{w}})$.

Dimostrazione
Fattorizziamo $\mu_{f,\underline{v}}$ e $\mu_{f,\underline{w}}$ in fattori irriducibili e per convenienza notazionale

scriviamo

$$\mu_{f,\underline{v}} = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}}} p^{m_p}, \qquad \mu_{f,\underline{w}} = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}}} p^{n_p},$$

dove gli m_p e gli n_p sono tutti nulli eccetto un numero finito.
Poniamo

$$P_1 = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}\\ m_p\geq n_p}} p^{m_p}, \qquad Q_1 = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}\\ n_p> m_p}} p^{n_p},$$

$$P_2 = \frac{\mu_{f,\underline{v}}}{P_1} = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}\\ m_p< n_p}} p^{m_p}, \quad Q_2 = \frac{\mu_{f,\underline{w}}}{Q_1} = \prod_{\substack{p\in\mathbb{K}[t]\\ \text{irriducibile}\\ n_p\leq m_p}} p^{n_p}$$

in modo tale che

$$mcm(\mu_{f,\underline{v}}, \mu_{f,\underline{w}}) = P_1Q_1$$

$$(P_1, Q_1) = (P_1, P_2) = (Q_1, Q_2) = (P_2, Q_2) = 1.$$

Poniamo $\underline{z} = P_2(f)(\underline{v}) + Q_2(f)(\underline{w})$.

$$\begin{aligned}(P_1Q_1)(f)(\underline{z}) &= (Q_1(f)P_1(f)P_2(f))(\underline{v}) + (P_1(f)Q_1(f)Q_2(f))(\underline{w}) = \\ &= Q_1(f)(\mu_{f,\underline{v}}(f)(\underline{v})) + P_1(f)(\mu_{f,\underline{w}}(f)(\underline{w})) = \underline{0} + \underline{0} = \underline{0}\end{aligned}$$

per cui $P_1Q_1 \in I(f, \underline{z})$ e dunque $\mu_{f,\underline{z}} | P_1Q_1$.
Si conclude mostrando l'altra divisibilità.
Scrivendo $P_2(f)(\underline{v}) = \underline{z} - Q_2(f)(\underline{w})$, come sopra $\mu_{f,\underline{z}}Q_1 \in I(f, P_2(f)(\underline{v}))$, da cui $\mu_{f,P_2(f)(\underline{v})} | \mu_{f,\underline{z}}Q_1$.
Per il Lemma1, $\mu_{f,P_2(f)(\underline{v})} = P_1$ ed inoltre P_1 e Q_1 sono coprimi, e allora $P_1 | \mu_{f,\underline{z}}$.
Analogamente, usando $Q_2(f)(\underline{w}) = \underline{z} - P_2(f)(\underline{v})$, si ottiene $Q_1 | \mu_{f,\underline{z}}$.
Ma P_1 e Q_1 sono coprimi, e allora $P_1Q_1 | \mu_{f,\underline{z}}$. ■

Siamo pronti per dimostare la seguente

Proposizione
Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita e sia $f \in \mathrm{End}(V)$. Allora eisite $\underline{v} \in V$ tale che $\mu_{f,\underline{v}} = \mu_f$.

Dimostrazione
Siano $\underline{v}_1, \ldots, \underline{v}_m$ generatori di V.
Per induzione e usando il Lemma2 è immediato vedere che per ogni $i = 1\_\_m$ esiste $\underline{z}_i \in V$ tale che $\mu_{f,\underline{z}_i} = mcm(\mu_{f,\underline{v}_1}, \ldots, \mu_{f,\underline{v}_i})$. Allora $\underline{v} = \underline{z}_m$ ha la proprietà voluta, in quanto $\mu_{f,\underline{v}} = mcm(\mu_{f,\underline{v}_1}, \ldots, \mu_{f,\underline{v}_m}) = \mu_f$. ■

Diamo una dimostrazione alternativa più diretta ma valida solo nel caso in cui il campo $\mathbb{K}$ sia infinito.
Consideriamo l'insieme di tutti i possibili polinomi minimi relativi

$$S = \{\mu_{f,\underline{v}} |\ \underline{v} \in V, \underline{v} \neq \underline{0}\}.$$

S è contenuto nell'insieme dei divisori di μ_f, quindi è un insieme finito.
Siano allora $\underline{v}_1, \dots, \underline{v}_k \in V$ tali che $S = \{\mu_{f,\underline{v}_1}, \dots, \mu_{f,\underline{v}_k}\}$ e poniamo $W_i = \mathrm{Ker}(\mu_{f,\underline{v}_i}(f))$, per $i = 1, \dots, k$.
Dato $\underline{v} \in V$, $\underline{v} \neq 0$, esite $1 \leq j \leq k$ tale che $\mu_{f,\underline{v}} = \mu_{f,\underline{v}_j}$ per cui $\mu_{f,\underline{v}_j}(f)(\underline{v}) = \mu_{f,\underline{v}}(f)(\underline{v}) = 0$, ovvero $\underline{v} \in W_j$.
Allora $V = \cup_{i=1}^{k} W_i$ ma, se $\mathbb{K}$ è infinito, V non è unione finita di sottospazi propri. Quindi esiste $1 \leq h \leq k$ tale che $V = W_h$, ma allora $\mu_{f,\underline{v}_h} \in I(f)$ e dunque $\mu_{f,\underline{v}_h} = \mu_f$. ■

Un endomorfismo $f \in \mathrm{End}(V)$ si dice *ciclico* se esiste una base $\mathcal{B}$ di V del tipo $\underline{v}, f(\underline{v}), \dots, f^{n-1}(\underline{v})$. La base $\mathcal{B}$ si dice *ciclica* per f e si dice che $\underline{v}$ *genera* la base ciclica o che è un generatore ciclico.
Ad esempio, per ogni $\underline{v} \in V$ non nullo, la restrizione di f a $\mathbb{K}[f](\underline{v})$ è un endomorfismo ciclico e $\underline{v}$ genera una base ciclica. Inoltre, f è ciclico se e solo se esiste $\underline{v} \in V$ tale che $\mathbb{K}[f](\underline{v}) = V$.
La matrice di f in una base ciclica $\mathcal{B}$ generata da $\underline{v}$ ha la forma

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

dove $f^n(\underline{v}) = -a_0\underline{v} - a_1 f(\underline{v}) \cdots - a_{n-1} f^{n-1}(\underline{v})$.
Osserviamo che il polinomio caratteristico di f è $t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$.
Infatti, per induzione su n, sviluppando secondo la prima riga,

$$\det \begin{pmatrix} t & 0 & \cdots & 0 & a_0 \\ -1 & t & \ddots & \vdots & a_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & t & \vdots \\ 0 & \cdots & 0 & -1 & t + a_{n-1} \end{pmatrix} =$$

$$= t \det \begin{pmatrix} t & & & a_1 \\ -1 & \ddots & & \vdots \\ & \ddots & t & \vdots \\ & & -1 & t + a_{n-1} \end{pmatrix} + a_0(-1)^{n+1} \det \begin{pmatrix} -1 & t & & \\ & -1 & \ddots & \\ & & \ddots & t \\ & & & -1 \end{pmatrix} =$$

$= t(t^{n-1} + a_{n-1}t^{n-2} + \cdots + a_1) + a_0 = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$.
Per questo motivo la matrice sopra si dice *matrice compagna* del polinomio monico q, e si sindica con $C(q)$.

In effetti, abbiamo già osservato che se f è ciclico allora $\mu_f = p_f = \mu_{f,\underline{v}}$ per ogni generatore ciclico $\underline{v}$. Viceversa, se $\mu_f = p_f$, allora ogni $\underline{v} \in V$ tale che $\mu_{f,\underline{v}} = \mu_f$

genera una base ciclica per f.
Abbiamo quindi dimostrato la seguente

Proposizione
$f \in \mathrm{End}(V)$ è ciclico se e solo se $\mu_f = p_f$.

La forma normale di Jordan

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita e sia $f \in \mathrm{End}(V)$.

Osserviamo che per ogni $p \in \mathbb{K}[t]$, $p(f)$ commuta con f, e quindi $\mathrm{Ker}\, p(f)$ e $\mathrm{Im}\, p(f)$ sono sottospazi f-invarianti.

Teorema (Decomposizione Primaria)
Dati $p_1, p_2 \in \mathbb{K}[t]$ coprimi, allora $\mathrm{Ker}(p_1p_2)(f) = \mathrm{Ker}\, p_1(f) \oplus \mathrm{Ker}\, p_2(f)$.

Dimostrazione
Poiché p_1 e p_2 sono coprimi, dall'identità di Bézout abbiamo $1 = q_1p_1 + q_2p_2$ per opportuni $q_1, q_2 \in \mathbb{K}[t]$.
Valutando su f otteniamo $id_V = q_1(f)p_1(f) + q_2(f)p_2(f)$.
Dato $\underline{v} \in \mathrm{Ker}\, p_1(f) \cap \mathrm{Ker}\, p_2(f)$, valutando la precedente relazione su $\underline{v}$ abbiamo $\underline{v} = (q_1(f)p_1(f))(\underline{v}) + (q_2(f)p_2(f))(\underline{v}) = \underline{0} + \underline{0} = \underline{0}$, per cui la somma è diretta.
Le inclusioni $\mathrm{Ker}\, p_1(f), \mathrm{Ker}\, p_2(f) \subset \mathrm{Ker}(p_1p_2)(f)$ sono ovvie in quanto $(p_1p_2)(f)$ è dato dalle composizioni $p_1(f)p_2(f) = p_2(f)p_1(f)$.
Quindi $\mathrm{Ker}(p_1p_2)(f) \supset \mathrm{Ker}\, p_1(f) \oplus \mathrm{Ker}\, p_2(f)$.
Per dimostrare l'altra inclusione, sia $\underline{v} \in \mathrm{Ker}(p_1p_2)(f)$. Come prima abbiamo $\underline{v} = (q_1(f)p_1(f))(\underline{v}) + (q_2(f)p_2(f))(\underline{v}) = \underline{w}_1 + \underline{w}_2$. Si è finito se mostriamo che $\underline{w}_1 \in \mathrm{Ker}\, p_2(f)$, $\underline{w}_2 \in \mathrm{Ker}\, p_1(f)$. Infatti,
$p_2(f)(\underline{w}_1) = (p_2(f)q_1(f)p_1(f))(\underline{v}) = q_1(f)((p_1p_2)(f)(\underline{v})) = q_1(f)(\underline{0}) = \underline{0}$. Allo sesso modo si ragiona per $\underline{w}_2$. ❏

Notiamo che il risultato si estende facilmemente per induzione al caso di $k > 2$ polinomi a due a due coprimi:
dati $p_1, \ldots, p_k \in \mathbb{K}[t]$ a due a due coprimi, allora

$$\mathrm{Ker}(p_1 \cdots p_k)(f) = \mathrm{Ker}\, p_1(f) \oplus \cdots \oplus \mathrm{Ker}\, p_k(f).$$

Infatti, p_1 e il prodotto $p_2 \cdots p_k$ sono coprimi e quindi, per decomposizione primaria, $\mathrm{Ker}(p_1 \cdots p_k)(f) = \mathrm{Ker}\, p_1(f) \oplus \mathrm{Ker}(p_2 \cdots p_k)(f)$ e per ipotesi induttiva il secondo addendo è $\mathrm{Ker}\, p_2(f) \oplus \cdots \oplus \mathrm{Ker}\, p_k(f)$.
Lasciamo al lettore la facile verifica che, in generale, per $W_i \subset V$ sottospazi, $i = 1\_k$, vale $W_1 \oplus (W_2 \oplus \cdots \oplus W_k) = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ (dove la questione non è l'uguaglianza, ma se esitono le varie somme dirette).

In particolare, se il prodotto $p_1 \cdots p_k \in I(f)$, allora $\mathrm{Ker}(p_1 \cdots p_k)(f) = V$ e abbiamo una decomposizione di V, detta *primaria*, in sottospazi f-invarianti:

$$V = \mathrm{Ker}\, p_1(f) \oplus \cdots \oplus \mathrm{Ker}\, p_k(f).$$

Abbiamo individuato due polinomi speciali nell'ideale di f, il polinomio caratteristico e il polinomio minimo di f. Se scriviamo le loro fattorizzazioni in

irriducibili in questo modo $p_f(t) = \prod_{i=1}^{k} p_i^{m_i}$, $\mu_f(t) = \prod_{i=1}^{k} p_i^{r_i}$, dove i p_i sono irriducibili distinti (e sappiamo che $1 \leq r_i \leq m_i$ per ogni i), allora i fattori delle due scritture sono coprimi a due a due e quindi abbiamo le due decomposizioni primarie di V in sottospazi f-invarianti (ed osserviamo che gli addendi sono tutti non nulli):

$$V = \operatorname{Ker} p_1(f)^{m_1} \oplus \cdots \oplus \operatorname{Ker} p_k(f)^{m_k},$$
$$V = \operatorname{Ker} p_1(f)^{r_1} \oplus \cdots \oplus \operatorname{Ker} p_k(f)^{r_k}.$$

Vogliamo vedere che in effetti le due decomposizioni coincidono e danno quella che si chiama la *decomposizione primaria canonica* di V data da f.

In generale, dato $f \in \operatorname{End}(V)$, la successione dei nuclei delle potenze di f è strettamente crescente fino a che si stabilizza.
Cioè, per ogni $k \geq 0$, $\operatorname{Ker} f^k \subset \operatorname{Ker} f^{k+1}$ ed esiste un $k_0 \geq 0$ tale che per ogni $k \geq k_0$, $\operatorname{Ker} f^k = \operatorname{Ker} f^{k_0}$ mentre se $0 \leq k \leq k_0 - 1$, $\operatorname{Ker} f^k \subsetneqq \operatorname{Ker} f^{k+1}$
La prima proprietà è ovvia, in quanto f^{k+1} è la composizione $f f^k$.
Per la seconda, sicuramente esiste un $\bar{k}$ tale che $\operatorname{Ker} f^{\bar{k}} = \operatorname{Ker} f^{\bar{k}+1}$. Infatti, altrimenti avremmo $\operatorname{Ker} f^k \subsetneqq \operatorname{Ker} f^{k+1}$ per ogni k, da cui $\dim \operatorname{Ker} f^k \geq k$, ma $\dim \operatorname{Ker} f^k \leq \dim V = n$ per ogni k ↯.
Sia allora k_0 il minimo tra i $\bar{k}$ di cui sopra.
Per definizione di k_0 abbiamo

$$\{\underline{0}\} = \operatorname{Ker} f^0 \subsetneqq \operatorname{Ker} f \subsetneqq \cdots \subsetneqq \operatorname{Ker} f^{k_0-1} \subsetneqq \operatorname{Ker} f^{k_0} = \operatorname{Ker} f^{k_0+1}.$$

Mostriamo per induzione su $h \geq 1$ che $\operatorname{Ker} f^{k_0+h} = \operatorname{Ker} f^{k_0}$.
Per $h = 1$ segue dalla definizione di k_0.
Supponiamo allora $h > 1$ e supponiamo $\operatorname{Ker} f^{k_0+h-1} = \operatorname{Ker} f^{k_0}$.
Sia $\underline{v} \in \operatorname{Ker} f^{k_0+h}$. Scriviamo $\underline{0} = f^{k_0+h}(\underline{v}) = f^{k_0+h-1}(f(\underline{v}))$ e otteniamo $f(\underline{v}) \in \operatorname{Ker} f^{k_0+h-1} = \operatorname{Ker} f^{k_0}$. Allora $\underline{0} = f^{k_0}(f(\underline{v})) = f^{k_0+1}(\underline{v})$ e quindi $\underline{v} \in \operatorname{Ker} f^{k_0+1} = \operatorname{Ker} f^{k_0}$. Quindi, $\operatorname{Ker} f^{k_0+h} \subset \operatorname{Ker} f^{k_0}$. L'altra inclusione è sempre vera.

Torniamo alle due decomposizioni primarie.
Per $i = 1 \_ k$, con le notazioni precedenti, poniamo $W_i = \operatorname{Ker} p_i(f)^{m_i}$, (sono i termini della decomposizione primaria di V data dal polinomio caratteristico), $U_i = \operatorname{Ker} p_i(f)^{r_i}$, (sono i termini della decomposizione primaria di V data dal polinomio minimo).

Poiché $r_i \leq m_i$, $U_i \subset W_i$. Ma $n = \sum_{i=1}^{k} \dim U_i \leq \sum_{i=1}^{k} \dim W_i = n$, per cui $\dim U_i = \dim W_i$ per ogni i e quindi $W_i = U_i$ per ogni i.

Osserviamo che, per ogni i, r_i è l'indice in cui si stabilizza la successione dei nuclei delle potenze di $p_i(f)$.
Infatti, più in generale, sia $p \in \mathbb{K}[t]$ irriducibile.
Sia $k \geq 0$ l'indice per cui si stabilizza la successione dei nuclei delle potenze di

$p(f)$

$$\{0\} \subsetneqq \operatorname{Ker} p(f) \subsetneqq \operatorname{Ker} p(f)^2 \subsetneqq \cdots \subsetneqq \operatorname{Ker} p(f)^k = \operatorname{Ker} p(f)^{k+1} = \cdots$$

e sia d la molteplicità con cui p compare nella fattorizzazione in irriducibili di μ_f, ovvero $\mu_f = p^d q$ con $q \in \mathbb{K}[t]$ tale che $p \nmid q$.
Vogliamo dimostrare che $d = k$.

Se fosse $d < k$, allora $\operatorname{Ker} p(f)^d \subsetneqq \operatorname{Ker} p(f)^k$.
Per decomposizione primaria, $V = \operatorname{Ker} p(f)^d \oplus \operatorname{Ker} q(f)$, ma allora per Grassmann $\operatorname{Ker} p(f)^k \cap \operatorname{Ker} q(f) \neq \{0\}$ e quindi, sempre per decomposizione primaria, $(p^k, q) \neq 1$, ovvero $p \mid q$ ↯.

Se fosse $d > k$, allora $\operatorname{Ker} p(f)^d = \operatorname{Ker} p(f)^k$, quindi $V = \operatorname{Ker} p(f)^k \oplus \operatorname{Ker} q(f)$ da cui $p^k q \in I(f)$ (poiché $p(f)^k$ si annulla sul primo addendo e $q(f)$ si annulla sul secondo). Ma $\deg p^k q = k \deg p + \deg q < d \deg p + \deg q = \deg \mu_f$ ↯.

Poniamo $g_i = f_{\big|_{\operatorname{Ker} p_i(f)^{r_i}}}$, la restrizione di f all'i-mo sottospazio della decomposizione primaria canonica.
Notiamo che $p_i^{r_i} \in I(g_i)$, e dunque il polinomio minimo di g_i è del tipo $\mu_{g_i} = p_i^{s_i}$ con $1 \leq s_i \leq r_i$. Ma allora, i polinomi minimi dei g_i sono coprimi a due a due, per cui il polinomio minimo di f, che in generale è il minimo comune multiplo dei polinomi minimi delle restrizioni, è in effetti il prodotto di tali polinomi minimi, $\mu_f = \prod_{i=1}^{k} \mu_{g_i}$. Ne segue che $\mu_{g_i} = p_i^{r_i}$ per ogni i.
Anche il polinomio caratteristico di g_i è del tipo $p_{g_i} = p_i^{n_i}$, con $n_i \geq r_i$. Usando una base di V adattata alla decomposizione primaria canonica e la matrice di f in tale base, otteniamo che il polinomio caratteristico di f è il prodotto dei polinomi caratteristici delle restrizioni, $p_f = \prod_{i=1}^{k} p_{g_i}$. Analogamente a prima, $p_{g_i} = p_i^{m_i}$ per ogni i.
Segue quindi che $\dim \operatorname{Ker} p_i(f)^{r_i} = m_i \deg p_i$.

Specializziamo la discussione al caso degli endomorfismi trianglolabili.

Se f è triangolabile, le fattorizzazioni dei polinomi p_f e μ_f hanno una forma speciale,

$$p_f(t) = \prod_{\lambda \in sp(f)} (t-\lambda)^{m_\lambda}, \quad \mu_f(t) = \prod_{\lambda \in sp(f)}^{k} (t-\lambda)^{r_\lambda}$$

(dove di nuovo, $1 \leq r_\lambda \leq m_\lambda$), e quindi abbiamo la decomposizione primaria canonica di V:

$$V = \bigoplus_{\lambda \in sp(f)} \operatorname{Ker}(f - \lambda id_V)^{m_\lambda} = \bigoplus_{\lambda \in sp(f)} \operatorname{Ker}(f - \lambda id_V)^{r_\lambda}.$$

Il sottospazio $\operatorname{Ker}(f - \lambda id_V)^{m_\lambda} = \operatorname{Ker}(f - \lambda id_V)^{r_\lambda}$ si dice *autospazio generalizzato* di f relativo all'autovalore λ (in quanto include l'autospazio $V_\lambda(f)$) e si

indica con $V'_\lambda(f)$. Notiamo che la successione dei nuclei delle potenze di $f-\lambda id_V$ è del tipo

$$\{\underline{0}\} \subsetneqq \mathrm{Ker}(f-\lambda id_V) \subsetneqq \cdots \subsetneqq \mathrm{Ker}(f-\lambda id_V)^{r_\lambda - 1} \subsetneqq V'_\lambda(f) = V'_\lambda(f) = \cdots .$$

Se chiamiamo $g_\lambda = f_{|_{V'_\lambda(f)}}$, la restrizione di f all'autospazio generalizzato relativo all'autovalore λ, allora

- $\mu_{g_\lambda}(t) = (t-\lambda)^{r_\lambda}$.
- $p_{g_\lambda}(t) = (t-\lambda)^{m_\lambda}$;
- $\dim V'_\lambda(f) = m_\lambda$;

Ossevazioni:
► Analogamente al caso diagonalizzabile in cui V è somma diretta degli autospazi, per un endomorfismo triangolabile, V è somma diretta degli autospazi generalizzati.
► La definizione degli autospazi generalizzati ha senso anche per endomorfismi non triangolabili e gli autospazi generalizzati sono in somma diretta, ma in questo caso la somma diretta non dà tutto V.
► Come nel caso della molteplicità geometrica, anche la molteplicità algebrica si realizza come dimensione di un sottospazio.
► Lo spettro di g_λ è dato dal solo autovalore λ. La decomposizione primaria canonica è un modo di separare il contributo di ogni autovalore.

Possiamo ottenere un novo criterio per la diagonalizzabilità di un endomorfismo legato al polinomio minimo:

Proposizione
$f \in \mathrm{End}(V)$ è diagonalizzabile se e solo se μ_f è completamente fattorizzabile in $\mathbb{K}[t]$ ed ha solo radici di molteplicità 1.

Dimostrazione
Supponiamo f sia diagonalizzabile, allora si ha che $V = \bigoplus_{\lambda \in sp(f)} V_\lambda(f)$. Poiché $f_{|_{V_\lambda(f)}} = \lambda id_{V_\lambda(f)}$ il polinomio minimo di $f_{|_{V_\lambda(f)}}$ è $t-\lambda$ quindi, il polinomio minimo di f, che è il minimo comune multiplo dei polinomi minimi delle restrizioni, è $\mu_f = \prod_{\lambda \in sp(f)} (t-\lambda)$ come volulto.
Viceversa, se μ_f è completamente fattorizzabile, anche p_f lo è, perché hanno gli stessi fattori irriducibili. Quindi f è triangolabile e dato che μ_f ha solo radici di molteplicità 1, gli autospazi generalizzati coincidono con gli autospazi. Quindi V è somma diretta degli autospazi di f e dunque f è diagonalizzabile. ❑

Otteniamo anche una nuova dimostrazione che la proprietà di essere diagonalizzabile è ereditaria per i sottospazi invarianti: se il polinomio minimo di f ha solo radici di molteplicità 1, anche il polinomio minimo della restrizione di f ad

un sottospazio invariante, che divide il polinomio minimo di f, ha solo radici di molteplicità 1.

Se $f, g \in \mathrm{End}(V)$ sono endomorfismi coniugati tramite $h \in GL(V)$, $gh = hf$, allora, per ogni $\lambda \in sp(f) = sp(g)$ e per ogni $s \geq 1$, $(g-\lambda id_V)^s h = h(f-\lambda id_V)^s$ e quindi $\mathrm{Ker}(g-\lambda id_V)^s = h(\mathrm{Ker}(f-\lambda id_V)^s)$.
Le dimensioni dei nuclei di $(f-\lambda id_V)^s$ sono dunque invarianti per coniugazione.

Organizziamo queste dimensioni in una stringa strettamente crescrente di interi positivi: se $r_\lambda(f)$ è la molteplicità di λ come radice del polinomio minimo di f (cioè l'indice per cui si stabilizza la successione dei nuclei delle potenze di $f-\lambda id_V$), per $j = 1 \_ r_\lambda(f)$, poniamo $d_j(\lambda, f) = \dim \mathrm{Ker}(f-\lambda id_V)^j$ e poniamo $D_\lambda(f) = (d_1(\lambda, f), d_2(\lambda, f), \ldots, d_{r_\lambda(f)}(\lambda, f))$.
Quindi, $d_1(\lambda, f)$ è la molteplicità geometrica dell'autovalore λ (la dimensione dell'autospazio), $d_{r_\lambda(f)}(\lambda, f)$ è la molteplicità algebrica dell'autovalore λ (la dimensione dell'autospazio generalilzzato) e $0 < d_1(\lambda, f) < \cdots < d_{r_\lambda(f)}(\lambda, f)$.

Per $1 < j < r_\lambda(f)$, i $d_j(\lambda, f)$ si dicono *molteplicità intermedie* e i $\mathrm{Ker}(f-\lambda id_V)^j$ *autospazi intermedi*.
$D_\lambda(f)$ si dice la *stringa invariante* di f relativa a λ.

Allo stesso modo si definiscono le molteplicità intermedie, gli autospazi intermedi e le stringhe invarianti per una matrice quadrata. Chiaramente, le stringhe invarianti di un endomorfismo e di una matrice ad esso associata in qualche base coincidono.

Vogliamo dimostrare che lo spettro e le stringhe invarianti per ogni autovalore danno invarianti completi per la relazione di coniugazione su $\mathcal{T}(V)$.
Per farlo, mostreremo che dato $f \in \mathcal{T}(V)$, esiste una base di V tale che la matrice associata ad f in tale base ha una forma speciale, detta *forma normale di Jordan*, che dipende solo dallo spettro e dalle stringhe invarianti.

Più specificamente, una forma normale di Jordan è una matrice triangolare superiore e diagonale a blocchi i cui blocchi sono del tipo

$$J(\lambda, k) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in M(k, \mathbb{K}), \quad \lambda \in \mathbb{K}.$$

$J(\lambda, k)$ è detto *blocco di Jordan di ordine k relativo a λ* e una forma normale di Jordan è del tipo $J = diag(J(\lambda_1, k_1), \ldots, J(\lambda_s, k_s))$ per $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$,

$k_1, \dots, k_s$ interi positivi

$$J = \begin{pmatrix} J(\lambda_1, k_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2, k_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J(\lambda_s, k_s) \end{pmatrix}.$$

Una base in cui la matrice di f è in forma normale di Jordan si dice *base di Jordan* per f. In generale, non esiste un'unica base di Jordan per f, mentre vedremo che la forma normale di Jordan è unica a meno di permutare i blocchi lungo la diagonale.

Se $f \in \mathrm{End}(V)$ ed esistono sottospazi f-invarianti W_i, $i = 1\_k$, tali che $V = W_1 \oplus \cdots \oplus W_k$, allora data una base $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ adattata alla somma diretta, la matrice di f nella base $\mathcal{B}$ è diagonale a blocchi e i blocchi sono dati dalle matrici delle $f_{|_{W_i}}$ nelle basi $\mathcal{B}_i$, $M_{\mathcal{B}}^{\mathcal{B}}(f) = diag(M_{\mathcal{B}_1}^{\mathcal{B}_1}(f_{|_{W_1}}), \dots, M_{\mathcal{B}_k}^{\mathcal{B}_k}(f_{|_{W_k}}))$.
Per ottenere una forma normale di Jordan per f si può quindi cercare una forma normale di Jordan per le singole restrizioni.

Per un endomorfismo triangolabile, possiamo usare la decomposizione primaria canoica di V data dagli autospazi generalizzati. Ovvero, ci possiamo restringere al caso di un endomorfismo triangolabile con un solo autovalore $\lambda \in \mathbb{K}$.

Sia quindi W uno spazio vettoriale su $\mathbb{K}$ di dimensione m, $g \in \mathrm{End}(W)$ tale che $sp(g) = \{\lambda\}$ e $p_f(t) = (t-\lambda)^m$, per cui $\mu_f(t) = (t-\lambda)^r$, con $1 \leq r \leq m$ (nel nostro caso, W sarà un autospazio generalizzato, g la restrizione di f a tale autospazio generalizzato). Osserviamo che r è l'indice per cui si stabilizza la successione dei nuclei delle potenze di $g - \lambda id_W$.

Possiamo ulteriormente restringerci a considerare endomorfismi *nilpotenti*:
$h \in \mathrm{End}(W)$ si dice nilpotente se esiste $k > 0$ tale che $h^k = 0$. Il polinomio minimo di h divide t^k, per cui $\mu_h = t^r$ per un $1 \leq r \leq k$. Inoltre, poiché il polinomio caratteristico ha gli stessi fattori irriducibili del polinomio minimo, $p_h(t) = t^m$ e $sp(h) = \{0\}$. L'esponente r si dice l'*indice di nilpotenza* di h, è il minimo intero positivo per cui $h^r = 0$.

Poiché $(g - \lambda id_W)^r = 0$, ponendo $h = g - \lambda id_W$, h è nilpotente.
Inoltre, per ogni base $\mathcal{B}$ di W, $M_{\mathcal{B}}^{\mathcal{B}}(g) = M_{\mathcal{B}}^{\mathcal{B}}(h) + \lambda I_m$, per cui se abbiamo una forma normale di Jordan per h, una forma normale di Jordan per g si ricava da quella di h sommando λI_m. Inoltre, la stringa invariante di g relativa a λ coincide con la stringa invariante di h relativa a 0.

Non è restrittivo quindi supporre che g sia nilpotente: $sp(g) = \{0\}$, $p_g(t) = t^m$, $\mu_g(t) = t^r$ e indichiamo con $d_i = \dim \mathrm{Ker}(g^i)$, $i \geq 0$, per cui $(d_1, d_2, \dots, d_r)$ è la stringa invariante $D_0(g)$ di g relativa all'autovalore 0 ($d_r = m$).

Notiamo che una forma normale di Jordan per g è del tipo

$$J = diag(J(0,k_1),\dots,J(0,k_s))$$

e che J diventa unica se ordiniamo i blocchi per ordine decrescente, ovvero se $k_1 \geq k_2 \geq \cdots \geq k_s$. In tal caso, J è univocamente determinata dagli interi $k_1, k_2, \dots, k_s$, $\sum_{i=1}^{s} k_i = m$, ma è più conveniente usare gli interi $b_1, b_2, \dots, b_m$, dove b_h è il numero di blocchi di J di ordine h, $b_h = |\{1 \leq j \leq s |\ k_j = h\}|$. Osserviamo che può accadere che qualche $b_h = 0$ e che $\sum_{h=1}^{m} hb(h) = m$.

Vogliamo dimostrare che ogni endomorfismo nilpotente ammette basi di Jordan e che la forma normale di Jordan (resa unica come sopra) non dipende dalla scelta della base di Jordan in quanto completamente determinata da $D_0(g)$.

Mostriamo l'unicità della forma normale di Jordan assumendo l'esistenza.

Sia dunque $J = diag(J(0,k_1),\dots,J(0,k_s))$ una forma normale di Jordan per g e osserviamo che J ha polinomio minimo $\mu_J(t) = t^r$ che è completamente determinato da $D_0(g)$ (formata r interi). Inoltre, per ogni $i \geq 0$, $d_i = \dim \operatorname{Ker} J^i$. Ricordiamo che $0 < d_1 < d_2 < \ldots < d_r = m = d_{r+1} = d_{r+2} = \cdots$.
Per motivi notazionali, estendiamo la definizione dei d_i ponendo $d_i = 0$ se $i < 0$.

È immediato verificare che

$$J(0,k)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & J(0,k-1) \\ 0 & 0 \end{pmatrix}$$

e reiterando che se $1 < i < k$,

$$J(0,k)^i = \begin{pmatrix} 0 & J(0,k-i+1) \\ 0 & 0 \end{pmatrix}$$

mentre $J(0,k)^i = 0$ se $i \geq k$, cioè $J(0,k)$ è nilpotente di indice di nilpotenza k. Poiché $\operatorname{rnk} J(0,k) = k-1$, abbiamo $\operatorname{rnk} J(0,k)^i = \max\{k-i, 0\}$.

Osserviamo che per ogni $i \geq 0$, $J^i = diag(J(0,k_1)^i,\dots,J(0,k_s)^i)$ e quindi $\operatorname{rnk} J^i = \sum_{j=1}^{s} \operatorname{rnk} J(0,k_j)^i$.
Da $J^r = 0$ segue che $J(0,k_j)^r = 0$ per ogni j, ovvero $k_j \leq r$ per ogni j. Abbiamo allora $b_h = 0$ se $h > r$, ovvero J contiene solo blocchi di ordine al più r.

Dalla formula delle dimensioni di nucleo e immagine abbiamo

$$d_i = \dim \operatorname{Ker} J^i = m - \operatorname{rnk} J^i = m - \sum_{j=1}^{s} \max\{k_j - i, 0\} = m - \sum_{h=1}^{m} b_h \max\{h - i, 0\}.$$

Segue che

$$d_i - d_{i-1} = \sum_{h=1}^{m} b_h \max\{h - i + 1, 0\} - \sum_{h=1}^{m} b_h \max\{h - i, 0\} = \sum_{h \geq i} b_h$$

(ricordiamo che questa ultima somma è finita).
Otteniamo infine

$$b_i = \sum_{h \geq i} b_h - \sum_{h \geq i+1} b_h = d_i - d_{i-1} - (d_{i+1} - d_i) = 2d_i - d_{i-1} - d_{i+1}.$$

Quindi, per ogni ordine, il numero di blocchi di tale ordine presenti in J è completamente determinato da $D_0(g)$, e dunque J è completamente determinata da $D_0(g)$ come voluto.

In particolare, $b_r = 2d_r - d_{r-1} - d_{r+1} = m - d_{r-1} \neq 0$, per cui J contiene almeno un blocco di ordine r.
Inoltre, il numero di blocchi presenti in J è

$$s = \sum_{h=1}^{r} b_h = \sum_{h=1}^{r} (2d_h - d_{h-1} - d_{h+1}) = \sum_{h=1}^{r} (d_h - d_{h-1}) + \sum_{h=1}^{r} (d_h - d_{h+1}) = \\ = d_r - d_{-1} + d_1 - d_{r+1} = d_1,$$

per cui s è pari alla molteplicità geometrica dell'autovalore 0.

Veniamo adesso all'esistenza di una base di Jordan per g.

Osserviamo che per i casi estremi $r = 1$ e $r = m$ trovare una base di Jordan è semplice.
Se $r = 1$, $D_0(g) = (m)$, $\mu_g(t) = t$ per cui $g = 0$. Ogni base di W è di Jordan per g e la forna normale di Jordan ha m blocchi di ordine 1. Questo è l'unico caso in cui g è diagonalizzabile.
Se $r = m$, $D_0(g) = (1, 2, \ldots, m)$, $\mu_g(t) = t^m = p_g(t)$, per cui g è ciclico e $g^m = 0$ ma $g^{m-1} \neq 0$. Fissiamo $\underline{w} \in W$ tale che $g^{m-1}(\underline{w}) \neq \underline{0}$ (notiamo che $g^i(\underline{w}) \neq \underline{0}$ per ogni $i = 0 \_ m - 1$, mentre $g^m(\underline{w}) = \underline{0}$), allora $\underline{w}$ genera una base di W ciclica per g, $\mathcal{B} = \{\underline{w}, g(\underline{w}), \ldots, g^{m-1}(\underline{w})\}$, ovvero, $W = \mathbb{K}[g](\underline{w})$.
In questa base, la matrice di g è la matrice compagna del polinomio t^m e notiamo che $C(t^m) = J(0, m)^\top$. Rovesciando l'ordine dei vettori, la base $\overleftarrow{\mathcal{B}} = \{g^{m-1}(\underline{w}), g^{m-2}(\underline{w}), \ldots, g(\underline{w}), \underline{w}\}$ è una base di Jordan per g. La forma normale di Jordan di g contiene un solo blocco di ordine m.

Notiamo che, in generale, fissato $\underline{w} \in W$ non nullo, la restrizione di g al sottospazio g invariante $\mathbb{K}[g](\underline{w}) = \operatorname{Span}\{\underline{w}, g(\underline{w}), \ldots, g^{m-1}(\underline{w})\}$ ammette una base

ciclica (se s è il minimo intero positivo tale che $g^s(\underline{w}) = \underline{0}$, la base ciclica è data da $\underline{w}, g(\underline{w}), \dots, g^{s-1}(\underline{w})$) e quindi una forma normale di Jordan con un solo blocco. Viceversa, se $U \subset W$ è un sottospazio g-invariante per cui la restrizione di g ha una matrice associata data da un solo blocco di Jordan, allora $g_{|_U}$ ammette una base ciclica.

Dimostrare l'esistenza di una base di Jordan per g, equivale dunque a dire che W si può scrivere come somma diretta di sottospazi g-invarianti, ognuno dei quali ammette una base ciclica per la restrizione di g.

Poiché $\operatorname{Ker} g^{r-1} \subsetneq \operatorname{Ker} g^r = W$, fissiamo $\underline{w} \in W$ tale che $g^{r-1}(\underline{w}) \neq \underline{0}$.
Sia $U = \mathbb{K}[g](\underline{w}) = \operatorname{Span}(\underline{w}, g(\underline{w}), \dots, g^{r-1}(\underline{w}))$, sottospazio g-invariante con base ciclica per $g_{|_U}$. Se mostriamo che esiste Z supplementare di U g-invariante, allora possiamo concludere per induzione sulla dimensione.
Osserviamo che $U \cap \operatorname{Im} g = g(U) = \operatorname{Span}(g(\underline{w}), \dots, g^{r-1}(\underline{w}))$. L'inclusione $g(U) \subset U \cap \operatorname{Im} g$ è ovvia. Se l'inclusione fosse stretta, allora $U \cap \operatorname{Im} g = U$, ovvero $U \subset \operatorname{Im} g$, per cui si avrebbe $\underline{w} \in \operatorname{Im} g$. Ma allora $\underline{w} = g(\underline{v})$ per qualche $\underline{v} \in W$, e allora $g^{r-1}(\underline{w}) = g^r(\underline{v}) = \underline{0}$, ↯.
Ragioniamo per induzione su r: per $r = 1$ si ha che $g = 0$ e tutti i supplementari vanno bene.
Supponiamo allora $r > 1$ e che se un endomorfismo ha polinomio minimo t^{r-1} allora possiamo trovare un supplementare invariante per ogni sottospazio che ammette una base ciclica.
Notiamo che $g_{|_{\operatorname{Im} g}}$ ha polinomio minimo t^{r-1}: infatti, poiché $g^{r-1}g = g^r = 0$, $g^{r-1}(\operatorname{Im} g) = \{\underline{0}\}$ mentre $\underline{y} = g(\underline{w}) \in \operatorname{Im} g$ ma $g^{r-2}(\underline{y}) \neq \underline{0}$.
Consideriamo allora $g(U) \subset \operatorname{Im} g$ che ammette base ciclica per la restrizione di g, e per ipotesi induttiva, scriviamo $\operatorname{Im} g = g(U) \oplus Z_0$ con Z_0 g-invariante.
Poniamo $Z_1 = g^{-1}(Z_0)$ ed osserviamo che, poiché $g(Z_0) \subset Z_0$, $Z_0 \subset Z_1$.
Mostriamo che $W = U + Z_1$.
Infatti, dato $\underline{v} \in W$, scriviamo $g(\underline{v}) = g(\underline{u}) + \underline{z}$ con $\underline{u} \in U$, $\underline{z} \in Z_0$. Poiché $\underline{z} = g(\underline{v} - \underline{u})$, $\underline{z}_1 = \underline{v} - \underline{u} \in Z_1$, e $\underline{v} = \underline{u} + \underline{z}_1$ come voluto.
Mostriamo adesso che $U \cap Z_0 = \{\underline{0}\}$.
Infatti, poiché $Z_0 \subset \operatorname{Im} g$, $U \cap Z_0 \subset U \cap \operatorname{Im} g \cap Z_0 = g(U) \cap Z_0 = \{\underline{0}\}$.
Possiamo allora trovare un sottospazio $Z \subset W$ tale che $Z_0 \subset Z \subset Z_1$ e $W = U \oplus Z$, basta completare una base di $U \oplus Z_0$ a base di W usando vettori in Z_1. Allora $g(Z) \subset g(Z_1) = Z_0 \subset Z$ mostra che Z è g-invariante.

Abbiamo dunque dimostrato che ogni endomorfismo nilpotente ammette basi di Jordan. La dimostrazione però non è costruttiva nel senso che non dà una ricetta di come costruire una tale base.
Vediamo allora una seconda dimostrazione costruttiva, che inoltre scende in dettaglio sulla struttura della successione dei nuclei delle potenze di un endomorfismo.

Torniamo al caso generale e supponiamo $r > 1$.

Poiché per $i = 1\_r$, $\operatorname{Ker} g^{i-1} \subsetneq \operatorname{Ker} g^i$, scegliamo W_i un supplemetare di $\operatorname{Ker} g^{i-1}$ in $\operatorname{Ker} g^i$ e scriviamo $\operatorname{Ker} g^i = \operatorname{Ker} g^{i-1} \oplus W_i$.
Notiamo che $\dim W_i = d_i - d_{i-1}$ e che $W = \operatorname{Ker} g^r = W_r \oplus \cdots \oplus W_1$ (ma i W_i non sono in generale g-invarianti!).
Mostriamo i seguenti fatti:

Lemma

1. $g(W_i) \subset \operatorname{Ker} g^{i-1}$.
2. $\dim g(W_i) = \dim W_i$.
3. Per $i \geq 2$, $g(W_i)$ è in somma diretta con $\operatorname{Ker} g^{i-2}$

Dimostrazione
1. Dato $\underline{w} \in W_i$, si ha $g^{i-1}(g(\underline{w})) = g^i(\underline{w}) = \underline{0}$ poiché $W_i \subset \operatorname{Ker} g^i$.
2. Poiché $\operatorname{Ker} g \subset \operatorname{Ker} g^{i-1}$, $W_i \cap \operatorname{Ker} g = \{\underline{0}\}$, per cui la restrizione di g a W_i è iniettiva.
3. Dato $\underline{v} \in \operatorname{Ker} g^{i-2} \cap g(W_i)$, $\underline{v} = g(\underline{w})$, con $\underline{w} \in W_i$ e $\underline{0} = g^{i-2}(\underline{v}) = g^{i-1}(\underline{w})$, per cui $\underline{w} \in W_i \cap \operatorname{Ker} g^{i-1} = \{\underline{0}\}$. Quindi $\underline{w} = \underline{0}$ e $\underline{v} = \underline{0}$.

Grazie al lemma, possiamo costruire i W_i induttivamente partendo da $W_r = U_r$ un qualsiasi supplementare di $\operatorname{Ker} g^{r-1}$ e per $i = r-1\_1$, $W_i = g(W_{i+1}) \oplus U_i$, con U_i un qualsiasi supplementare di $\operatorname{Ker} g^{i-1} \oplus g(W_{i+1})$ in $\operatorname{Ker} g^i$ (notare che alcuni U_i potrebbero essere banali).

$$
\begin{aligned}
W = \operatorname{Ker} g^r &= \operatorname{Ker} g^{r-1} \oplus \overbrace{U_r}^{W_r} \\
\operatorname{Ker} g^{r-1} &= \operatorname{Ker} g^{r-2} \oplus \overbrace{g(U_r) \oplus U_{r-1}}^{W_{r-1}} \\
\operatorname{Ker} g^{r-2} &= \operatorname{Ker} g^{r-3} \oplus \overbrace{g^2(U_r) \oplus g(U_{r-1}) \oplus U_{r-2}}^{W_{r-2}} \\
&\vdots \\
\operatorname{Ker} g &= \overbrace{g^{r-1}(U_r) \oplus g^{r-2}(U_{r-1}) \oplus \cdots \oplus g(U_2) \oplus U_1}^{W_1}
\end{aligned}
$$

Si ha quindi $W_i = g^{r-i}(U_r) \oplus \cdots \oplus g(U_{i+1}) \oplus U_i$, e quindi W è somma diretta degli U_i e delle loro immagini reiterate tramite g.
Definiamo, per $i = 1\_r$, $Z_i = U_i \oplus g(U_i) \oplus \cdots \oplus g^{i-1}(U_i)$ (di nuovo, alcuni Z_i potrebbero essere banali) ottenendo $W = Z_r \oplus \cdots \oplus Z_1$ una decomposizione in sottospazi g-invarianti.
Osserviamo che, nel caso $U_i \neq \{\underline{0}\}$, fissata una base $\mathcal{B}_i$ per U_i, allora l'unione $\mathcal{B}_i \cup g(\mathcal{B}_i) \cup \cdots \cup g^{i-1}(\mathcal{B}_i)$ dà una base di Z_i. Poniamo $\mathcal{B}_i = \emptyset$ se $U_i = \{\underline{0}\}$. Allora l'unione $\bigcup_{\substack{i=1\_r \\ j \geq 0}} g^j(\mathcal{B}_i)$ è una base $\mathcal{B}$ di W.
Dato $\underline{w} \in \mathcal{B}_i$, la base $\mathcal{B}$ contiene $\mathcal{C}_{\underline{w}} = \{\underline{w}, g(\underline{w}), \ldots, g^{i-1}(\underline{w})\}$, e dato che $g^i(\underline{w}) = \underline{0}$, $\operatorname{Span}(\underline{w}, g(\underline{w}), \ldots, g^{i-1}(\underline{w})) = \mathbb{K}[g](\underline{w})$ è g-invariante e $\mathcal{C}_{\underline{w}}$ è una

base ciclica per la restrizione di g a tale sottospazio. Ripetendo con gli altri vettori di $\mathcal{B}_i$, si ottiene una decomposizione di Z_i in sottospazi g-invarianti di ognuno dei quali abbiamo una base ciclica per la restrizione di g. Usando le basi $\overleftarrow{\mathcal{C}_{\underline{w}}}$, ottenute rovesciando l'ordine dei vettori delle basi cicliche, si ottiene una base di Jordan per $g_{|_{Z_i}}$ composta di soli blocchi di ordine i.
Riordiniamo dunque la base $\mathcal{B}$ in modo che vi compaiano le basi $\overleftarrow{\mathcal{C}_{\underline{w}}}$ al variare di $\underline{w}$ negli insiemi $\mathcal{B}_r, \mathcal{B}_{r-1}, \ldots, \mathcal{B}_1$, in questo ordine, e otteniamo una base di Jordan per g con lungo la diagonale blocchi di Jordan di ordini decrescenti.

Osserviamo che nella forma normale di Jordan ottenuta in questa base ci sono esattamente $\dim U_i$ blocchi di ordine i per $i = 1\_\_r$ (per cui, con le notazioni precedenti $b_i = \dim U_i$), e che esistono blocchi di ordine massimo r (poiché U_r non è banale).
Usando il fatto che $\dim U_i = \dim \operatorname{Ker} g^i - \dim \operatorname{Ker} g^{i-1} - \dim W_{i+1}$, ed osservando che $\dim W_{i+1} = \dim \operatorname{Ker} g^{i+1} - \dim \operatorname{Ker} g^i$, otteniamo di nuovo la formula che lega i b_i ai d_j: $b_i = 2d_i - d_{i+1} - d_{i-1}$.
Inoltre, la base di Jordan contiene basi di tutti gli autospazi intermedi di g: una base per $\operatorname{Ker} g^k$ è data prendendo i primi k vettori in ogni base $\overleftarrow{C_{\underline{w}}}$.

Come conclusione di questa discussione possiamo enunciare il teorema di classificazione per endomorfismi triangolabili a meno di coniugazione.

Teorema
Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita.
L'unione dello spettro e, per ogni autovalore, della relativa stringa invariante, costituisce un invariante completo per la relazione di coniugazione su $\mathcal{T}(V)$.
Tale invariante completo determina in modo univoco tra le matrici che rappresentano un dato endomorfismo triangolabile la forma normale di Jordan (unica a meno di permutare i contributi dovuti ai vari autovalori). Ne segue che anche la forma normale di Jordan è un invariante completo.

Dimostrazione
Una base di Jordan per f si costruisce prendendo una base di ogni autospazio generalizzato che sia di Jordan per la restrizione di f. Queste ultime si costruiscono a partire dalle stringhe invarianti per le restrizioni.
Il teorema segue se, per ogni autovalore λ, la stringa invariante di f relativa a λ coincide con la stringa invariante di $f_{|_{V'_\lambda(f)}}$ relativa a λ.
Basta quindi osservare che, dal fatto che tutti i nuclei delle potenze di $f - \lambda id_V$ sono contenuti nell'autospazio generalizzato $V'_\lambda(f)$, segue che, per ogni $j \geq 0$, $\operatorname{Ker}(f_{|_{V'_\lambda(f)}} - \lambda id_{V'_\lambda(f)})^j = \operatorname{Ker}(f - \lambda id_V)^j$. ∎

Osservazioni:
► Per ottenere le stringhe invarianti, non c'è bisogno di conoscere l'autospazio generalizzato: per ogni autovalore λ si calcolano le dimensioni dei $\operatorname{Ker}(f - \lambda id_V)$, che danno una successione crescente, fino a che la successione si stabilizza.

▶ Se $\mathbb{K}$ è algebricamente chiuso, abbiamo un invariante completo su tutto $\mathrm{End}(V)$.
▶ Un risultato analogo si ha per le matrici quadrate triangolabili:
ogni $A \in M(n,\mathbb{K})$ triangolabile è simile ad una matrice $J(A)$ in forma normale di Jordan. Tale matrice è unica a meno di permutazione dei blocchi di Jordan che la compongono. $A, B \in M(n,\mathbb{K})$ triangolabili sono simili se e solo se $J(A) = J(B)$ (a meno di permutazione dei blocchi).
▶ Considerando i polinomi caratteristici dei singoli blocchi di una forma normale di Jordan di f, si ottiene una decomposizione del polinomio caratteristico di f del tipo

$$p_f = \prod_{\lambda \in sp(f)} \prod_{j=1}^{mg(\lambda)} \left((t-\lambda)^j\right)^{b_j(\lambda)},$$

dove $b_j(\lambda)$ è il numero di blocchi di ordine j relativi all'autovalore λ. La lista di polinomi in cui $(t-\lambda)^j$ viene ripetuto $b_j(\lambda)$ volte si dice la lista dei *divisori elementari* di f. Anche i divisori elementari sono un invariante completo per la coniugazione.
▶ La discussione fatta per la forma normale di Jordan può essere ripetuta anche nel caso in cui l'endomorfismo f non è triangolabile sempre a partire dalla decomposizione primaria canonica. In tal caso, i divisori elementari di f sono del tipo p^s con $p \in \mathbb{K}[t]$ fattore irriducibile del polinomio minimo di f, gli autospazi generalizzati sono sostituiti dai nuclei di $p(f)^r$ (r la molteplicità di p nella fattorizzazione in irriducibili di μ_f) e ognuno di tali nuclei viene decomposto in sottospazi f-invarianti ciclici (uno per ogni divisore elementare). La matrice che si ottiene scegliendo una base ciclica in ognuno di tali sottospazi è una matrice diagonale a blocchi formata dalle matrici compagne dei divisori elemetari, detta *forma normale primaria* di f. La forma normale primaria (come pure i divisori elementari) è un invariante completo per la coniugazione. Osserviamo che se nell'addendo ciclico della decomposizione dato dal divisore elementare p^s, $\deg p = d$, invece di prendere la base ciclica $\underline{u}, f(\underline{u}), \ldots, f^{sd-1}(\underline{u})$ prendiamo la base data da

$$\begin{cases} \underline{u}, f(\underline{u}), \ldots, f^{d-1}(u) \\ p(f)(\underline{u}), f(p(f)(\underline{u})), \ldots, f^{d-1}(p(f)(\underline{u})) \\ p(f)^2(u), f(p^2(f)(\underline{u})), \ldots, f^{d-1}(p^2(f)(\underline{u})) \\ \ldots \\ p^{s-1}(f)(\underline{u}), f(p(f)^{s-1}(\underline{u})), \ldots, f^{d-1}(p(f)^{s-1}(\underline{u})) \end{cases},$$

allora la matrice che si ottiene è del tipo

$$\begin{pmatrix} C(p) & 0 & \cdots & 0 \\ E & C(p) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & E & C(p) \end{pmatrix}$$

dove $E \in M(d,\mathbb{K})$ è la matrice E_{1d} con un 1 nel posto $(1,d)$, 0 altrove.

Se rovesciamo l'ordine della base otteniamo una matrice del tipo

$$\begin{pmatrix} C'(p) & E' & \cdots & 0 \\ 0 & C'(p) & \ddots & \vdots \\ \vdots & \vdots & \ddots & E' \\ 0 & 0 & \cdots & C'(p) \end{pmatrix}$$

dove $E' \in M(d, \mathbb{K})$ è la matrice E_{d1} con un 1 nel posto $(d, 1)$, 0 altrove, e

$$C'(p) = \begin{pmatrix} -a_{d-1} & 1 & \cdots & 0 \\ -a_{d-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ -a_0 & 0 & \cdots & 0 \end{pmatrix} \quad (p(t) = a_0 + \cdots + a_{d-1}t^{d-1} + t^d).$$

Tale matrice, avendo la sopradiagonale tutta composta da 1, è una estensione di un blocco di Jordan (che si ottiene quando $d = 1$).
La forma normale primaria è dunque una estensione della forma normale di Jordan (e si riduce a quest'ultima nel caso triangolabile).
▶ La forma normale primaria è in larga misura implicita perchè si basa sulla possibilità di fattorizzare effettivamente un polinomio dato in fattori irriducibili (e quindi anche sul conoscere quali siano i polinomi irriducibili di $\mathbb{K}[t]$). Ad esempio, per campi infiniti (che non siano estensioni finite di $\mathbb{Q}$) non esistono algoritmi per la fattorizzazione dei polinomi o per riconoscere se un polinomio è irriducibile. Inoltre, la forma normale primaria dipende dal campo degli scalari (perché i polinomi irriducibili cambiano se il campo cambia).
Esiste però un'altra forma normale, detta *forma normale razionale*, che è un invariante completo per coniugazione ma non si basa sulla decomposizione primaria canonica e non dipende dal campo degli scalari. Inoltre, è effettivamente calcolabile per ogni endomorfismo.

La forma normale di Jordan reale

Poiché $\mathbb{C}$ è algebricamente chiuso, per ogni spazio vettoriale V su $\mathbb{C}$ di dimensione finita abbiamo classificato completamente $\mathrm{End}(V) = \mathcal{T}(V)$ a meno di coniugazione e abbiamo classificato $M(n,\mathbb{C})$ a meno di similitudine per ogni n.
Se invece V è uno spazio vettoriale su $\mathbb{R}$ di dimensione finita almeno 2, esistono endomorfismi non triangolabili.
Sfruttando il fatto che $\mathbb{C}$ è una estensione di $\mathbb{R}$, possiamo però dare una versione reale della forma normale di Jordan, che risulta essere un invariante completo per coniugazione e per similitudine.
Vediamo un primo risultato in questa direzione.

Proposizione
Siano $A, B \in M(n,\mathbb{R})$. Allora, esiste $P \in GL(n,\mathbb{R})$ tale che $B = PAP^{-1}$ se e solo se esiste $Q \in GL(n,\mathbb{C})$ tale che $B = QAQ^{-1}$.
Ovvero, A e B sono simili come matrici reali se e solo se lo sono se considerate come matrici complesse.

Dimostrazione
Poiché $GL(n,\mathbb{R}) \subset GL(n,\mathbb{C})$ una implicazione è ovvia.
Supponiamo allora che esista $Q \in GL(n,\mathbb{C})$ tale che $B = QAQ^{-1}$.
Scriviamo $Q = R + iS$ con R, S reali. Da $B(R+iS) = (R+iS)A$, comparando parti reali e parti immaginarie abbiamo $BR = RA$ e $BS = SA$. Ne segue che $B(R+rS) = (R+rS)A$ per ogni $r \in \mathbb{R}$. Si conclude se esiste un $r \in \mathbb{R}$ per cui $R + rS$ è invertibile. Se t è un'indeterminata, consideriamo il polinomio $p(t) = \det(R + tS) \in \mathbb{R}[t] \subset \mathbb{C}[t]$. Il polinomio $p(t)$ è non nullo, in quanto non si annulla per $t = i$, e quindi esistono solo un numero finito di $r \in \mathbb{R}$ tali che $R + rS$ non è invertibile. ∎

In altre parole, la forma normale di Jordan (complessa) di una matrice reale, pensata come matrice complessa, è un invariante completo per la relazione di similitudine su $M(n,\mathbb{R})$. Questo rinforza l'idea che si possa ricavare una forma normale di Jordan reale per le matrici reali dalla loro forma normale di Jordan complessa. Inoltre pone il problema di come sollevare queste considerazioni al caso della coniugazione di endomorfismi. Questo si ottiene tramite il meccanismo della *complessificazione*.

Dato V spazio vettoriale su $\mathbb{R}$, consideriamo l'applicazione

$$\blacksquare : \mathbb{C} \times (V \times V) \to V \times V,\ (z, (\underline{v}, \underline{w})) \mapsto z\blacksquare(\underline{v}, \underline{w})$$

data da $(a+ib)\blacksquare(\underline{v}, \underline{w}) = (a\underline{v} - b\underline{w}, b\underline{v} + a\underline{w})$ per ogni $a, b \in \mathbb{R}$, $\underline{v}, \underline{w} \in V$.
Notiamo che $\blacksquare$ estende a $\mathbb{C}$ il prodotto per scalari su $V \times V$: se infatti $b = 0$, $a\blacksquare(\underline{v}, \underline{w}) = (a\underline{v}, a\underline{w})$.

Mostriamo che $V \times V$ dotato della usuale somma e del prodotto per scalari $\blacksquare$, è uno spazio vettoriale su $\mathbb{C}$.

Basta verificare la proprietà "associativa" di $\cdot$ e le due proprietà "distributive" di $\cdot$ sulle somme in $\mathbb{C}$ e in $V \times V$.
Per $a, b, c, d \in \mathbb{R}$, $\underline{v}, \underline{w}, \underline{v}', \underline{w}' \in V$ abbiamo:

$$\begin{aligned}
((a+ib)(c+id))\cdot(\underline{v},\underline{w}) &= (ac-bd+i(ad+bc))\cdot(\underline{v},\underline{w}) = \\
&= ((ac-bd)\underline{v}-(ad+bc)\underline{w},(ad+bc\underline{v}+(ac-bd)\underline{w})) = \\
&= (a(c\underline{v}-d\underline{w})-b(c\underline{w}+d\underline{v}),b(c\underline{v}-d\underline{w})+a(c\underline{w}+d\underline{v})) = \\
&= (a+ib)\cdot((c+id)\cdot(\underline{v},\underline{w})).\\
((a+ib)+(c+id))\cdot(\underline{v},\underline{w}) &= (a+c+i(b+d)\cdot(\underline{v},\underline{w}) = \\
&= ((a+c)\underline{v}-(b+d)\underline{w},(b+d)\underline{v}+(a+c)\underline{w}) = \\
&= ((a\underline{v}-b\underline{w},b\underline{v}+a\underline{w})+(c\underline{v}-d\underline{w},d\underline{v}+c\underline{w}) = \\
&= (a+ib)\cdot(\underline{v},\underline{w})+(c+id)\cdot(\underline{v}.\underline{w}).\\
(a+ib)\cdot((\underline{v},\underline{w})+(\underline{v}',\underline{w}')) &= (a+ib)\cdot(\underline{v}+\underline{v}',\underline{w}+\underline{w}') = \\
&= (a(\underline{v}+\underline{v}')-b(\underline{w}+\underline{w}'),b(\underline{v}+\underline{v}')+a(\underline{w}+\underline{w}')) = \\
&= (a\underline{v}-b\underline{w},b\underline{v}+a\underline{w})+(a\underline{v}'-b\underline{w}',b\underline{v}'+a\underline{w}') = \\
&= (a+ib)\cdot(\underline{v},\underline{w})+(a+ib)\cdot(\underline{v}',\underline{w}').
\end{aligned}$$

Lo spazio vettoriale così ottenuto si dice *complessificato di* V e si indica con $V_{\mathbb{C}}$. Come d'uso, omettiamo $\cdot$ e scriviamo $\underline{v}+i\underline{w}$ al posto di $(\underline{v},\underline{w})$. Abbiamo quindi $(a+ib)(\underline{v}+i\underline{w}) = (a\underline{v}-b\underline{w})+i(b\underline{v}+a\underline{w})$.

Identifichiamo V con la copia di V in $V\times V$ data da $V\times\{0\}$, detta *parte reale* di $V_{\mathbb{C}}$. Con un abuso di notazione, indichiamo $V\times\{0\}$ ancora con V, e indichiamo con iV la copia di V in $V\times V$ data da $\{0\}\times V$, detta *parte immaginaria* di $V_{\mathbb{C}}$. Quindi, identifichiamo il vettore $\underline{v}$ di V con il vettore $\underline{v}+i\underline{0}$ di $V_{\mathbb{C}}$ (che indicheremo semplicemente con $\underline{v}$, come pure indicheremo con $i\underline{v}$ il vettore $\underline{0}+i\underline{v}$). Notiamo che come spazio vettoriale su $\mathbb{R}$, $V_{\mathbb{C}} = V\oplus iV$.

Possiamo definire un coniugio su $V_{\mathbb{C}}$ ponendo $\overline{\underline{v}+i\underline{w}} = \underline{v}+i(-\underline{w}) = \underline{v}-i\underline{w}$. Notiamo che $\overline{\underline{z}_1+\underline{z}_2} = \overline{\underline{z}_1}+\overline{\underline{z}_2}$ e $\overline{\alpha\underline{z}_1} = \overline{\alpha}\,\overline{\underline{z}_1}$ per ogni $\underline{z}_1,\underline{z}_2\in V_{\mathbb{C}}, \alpha\in\mathbb{C}$ e quindi il coniugio è un endomorfismo di $V\times V$ (come spazio vettoriale su $\mathbb{R}$, non è un endomorfismo di $V_{\mathbb{C}}$). Inoltre $\overline{\overline{\underline{z}}} = \underline{z}$ per ogni $\underline{z}\in V_{\mathbb{C}}$, ovvero il coniugio è un isomorfismo di $V\times V$ diagonalizzabile con spettro ±1. I due autospazi sono $V = \{\underline{z}\in V_{\mathbb{C}} |\ \overline{\underline{z}} = \underline{z}\}$, $iV = \{\underline{z}\in V_{\mathbb{C}} |\ \overline{\underline{z}} = -\underline{z}\}$.
Dato $\underline{z} = \underline{v}+i\underline{w}\in V_{\mathbb{C}}$ la sua *parte reale* è $\Re(\underline{z}) = \underline{v} = \dfrac{\underline{z}+\overline{\underline{z}}}{2}\in V$, la sua *parte immaginaria* è $\Im(\underline{z}) = \underline{w} = \dfrac{\underline{z}-\overline{\underline{z}}}{2i}\in V$.

Data $\mathcal{B} = \{\underline{v}_1,\dots,\underline{v}_n\}$ una base di V (come spazio vettoriale su $\mathbb{R}$) allora $\mathcal{B}$ è anche una base di $V_{\mathbb{C}}$ (come spazio vettoriale su $\mathbb{C}$), detta *base reale* di $V_{\mathbb{C}}$.
Infatti, dati $\underline{v},\underline{w}\in V$, scriviamo $\underline{v} = \sum_{j=1}^n a_j\underline{v}_j$, $\underline{w} = \sum_{j=1}^n b_j\underline{v}_j$, con $a_j, b_j\in\mathbb{R}$, $j = 1\_n$, allora $\underline{v}+i\underline{w} = \sum_{j=1}^n (a_j+ib_j)\underline{v}_j$, per cui $\mathcal{B}$ genera $V_{\mathbb{C}}$.
Se poi $\underline{0} = \underline{0}+i\underline{0} = \sum_{j=1}^n (a_j+ib_j)\underline{v}_j = \sum_{j=1}^n a_j\underline{v}_j + i\sum_{j=1}^n b_j\underline{v}_j$ allora $\sum_{j=1}^n a_j\underline{v}_j = \sum_{j=1}^n b_j\underline{v}_j = \underline{0}$, da cui $a_j = b_j = 0$ per ogni $j = 1\_n$.
Quindi $\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V$ (mentre come spazio vettoriale su $\mathbb{R}$, $V_{\mathbb{C}} = V\times V$ ha dimensione $\dim_{\mathbb{R}} V_{\mathbb{C}} = 2\dim_{\mathbb{R}} V$).

Dato $f \in \mathrm{End}(V)$, definiamo il *complessificato* di f, $f_{\mathbb{C}} \in \mathrm{End}(V_{\mathbb{C}})$ tramite $f_{\mathbb{C}}(\underline{v} + i\underline{w}) = f(\underline{v}) + if(\underline{w})$, per ogni $\underline{v}, \underline{w} \in V$.
È immediato verificare che $f_{\mathbb{C}}$ è $\mathbb{C}$-lineare, che $f_{\mathbb{C}}|_V = f$ e che il nucleo di $f_{\mathbb{C}}$ è il complessificato del nucleo di f, $\mathrm{Ker}\, f_{\mathbb{C}} = (\mathrm{Ker}\, f)_{\mathbb{C}}$.
Inoltre, $f_{\mathbb{C}}$ rispetta il coniugio: $f_{\mathbb{C}}(\overline{\underline{z}}) = \overline{f_{\mathbb{C}}(\underline{z})}$ per ogni $\underline{z} \in V_{\mathbb{C}}$. In particolare, $\mathrm{Ker}\, \overline{f_{\mathbb{C}}} = \overline{\mathrm{Ker}\, f_{\mathbb{C}}}$, dove $\overline{f_{\mathbb{C}}}$ è la composizione del coniugio con $f_{\mathbb{C}}$, che osserviamo è $\mathbb{R}$-lineare (ma non $\mathbb{C}$-lineare).
Otteniamo $*_{\mathbb{C}} : \mathrm{End}(V) \to \mathrm{End}(V_{\mathbb{C}})$, $f \mapsto f_{\mathbb{C}}$, che è un omomorfismo di anelli: $(fg)_{\mathbb{C}} = f_{\mathbb{C}} g_{\mathbb{C}}$ per ogni $f, g \in \mathrm{End}(V)$

Se $\mathcal{B}$ è una base reale di $V_{\mathbb{C}}$ (e quindi di V), allora $M_{\mathcal{B}}^{\mathcal{B}}(f_{\mathbb{C}}) = M_{\mathcal{B}}^{\mathcal{B}}(f) \in M(n, \mathbb{R})$.
In particolare, il polinomio caratteristico di $f_{\mathbb{C}}$ coincide con il polinomo caratteristico di f ed è quindi a coefficienti reali: $p_{f_{\mathbb{C}}} = p_f \in \mathbb{R}[t] \subset \mathbb{C}[t]$. Lo stesso vale per il polinomio minimo.
Inoltre, è adesso chiaro che la proposizione ad inizio capitolo ci dice che due endomorfismi $f, g \in \mathrm{End}(V)$ sono coniugati se e solo se i loro complessificati $f_{\mathbb{C}}, g_{\mathbb{C}} \in \mathrm{End}(V_{\mathbb{C}})$ sono coniugati, proprio perché un endomorfismo reale e il suo complessificato sono rappresentati dalla stessa matrice.
Quindi, la forma normale di Jordan complessa del complessificato è un invariante completo per coniugazione per gli endomorfismi reali.

Ricordiamo che se $\alpha \in \mathbb{C}$ è una radice di un polinomio complesso a coefficienti reali, allora anche $\overline{\alpha}$ è una radice della stessa molteplicità. Suddividiamo allora le radici di $p_{f_{\mathbb{C}}}$ come $sp(f_{\mathbb{C}}) = R \cup U \cup \overline{U}$, dove R è dato dalle radici reali, U contiene metà delle radici complesse non reali, $\overline{U}$ contiene i coniugati degli elementi di U, e fattorizziamo completamente $p_{f_{\mathbb{C}}}$:

$$p_{f_{\mathbb{C}}}(t) = \prod_{\lambda \in R} (t - \lambda)^{m_\lambda} \prod_{\alpha \in U} (t - \alpha)^{n_\alpha} (t - \overline{\alpha})^{n_\alpha}.$$

Osserviamo che $R = sp(f)$ e che è possibile che R e/o U (e quindi $\overline{U}$) siano vuoti.
La fattorizzazione di p_f come prodotto di irriducibili in $\mathbb{R}[t]$ è quindi:

$$p_f(t) = \prod_{\lambda \in R} (t - \lambda)^{m_\lambda} \prod_{\alpha \in U} q_\alpha(t)^{n_\alpha},$$

dove se $\alpha \in \mathbb{C} \setminus \mathbb{R}$, $q_\alpha(t) = (t - \alpha)(t - \overline{\alpha}) = t^2 - (\alpha + \overline{\alpha})t + |\alpha|^2$ (irriducibile in $\mathbb{R}[t]$).
Confrontiamo le decomposizioni primarie canoniche di $V_{\mathbb{C}}$ e V date dalle rispettive fattorizzazioni di $p_{f_{\mathbb{C}}}$ e p_f.
Per $\lambda \in R$, $\alpha \in U$, poniamo

$$\tilde{W}_\lambda = \mathrm{Ker}(f_{\mathbb{C}} - \lambda id_{V_{\mathbb{C}}})^{m_\lambda}, \quad U_\alpha = \mathrm{Ker}(f_{\mathbb{C}} - \alpha id_{V_{\mathbb{C}}})^{n_\alpha}, \quad \hat{U}_\alpha = \mathrm{Ker}(f_{\mathbb{C}} - \overline{\alpha} id_{V_{\mathbb{C}}})^{n_\alpha},$$

$$W_\lambda = \mathrm{Ker}(f - \lambda id_V)^{m_\lambda}, \quad Z_\alpha = \mathrm{Ker}\, q_\alpha(f)^{n_\alpha}$$

per cui

$$V_{\mathbb{C}} = (\bigoplus_{\lambda\in R} \tilde{W}_\lambda) \oplus (\bigoplus_{\alpha\in U}(U_\alpha \oplus \hat{U}_\alpha)),$$

$$V = (\bigoplus_{\lambda\in R} W_\lambda) \oplus (\bigoplus_{\alpha\in U} Z_\alpha).$$

Osserviamo che $\tilde{W}_\lambda = (W_\lambda)_{\mathbb{C}}$, poiché $(f_{\mathbb{C}} - \lambda id_{V_{\mathbb{C}}})^{m_\lambda}$ è il complessificato di $(f-\lambda id_V)^{m_\lambda}$; analogamente, applicando il teorema di decomposizione primaria a $(Z_\alpha)_{\mathbb{C}} = \mathrm{Ker}(q_\alpha(f_{\mathbb{C}})^{n_\alpha})$, $U_\alpha \oplus \hat{U}_\alpha = (Z_\alpha)_{\mathbb{C}}$.
Inoltre $\hat{U}_\alpha = \overline{U_\alpha}$ (l'immagine tramite il coniugio di U_α), infatti, per ogni $\underline{z} \in V_{\mathbb{C}}$, $(\overline{f_{\mathbb{C}} - \alpha id_{V_{\mathbb{C}}}})(\underline{z}) = (f_{\mathbb{C}} - \overline{\alpha} id_{V_{\mathbb{C}}})(\overline{\underline{z}})$, e lo stesso vale per le loro potenze.
In particolare U_α e $\hat{U}_\alpha$ sono isomorfi (ma non tramite il coniugio, che non è $\mathbb{C}$-lineare!): se $\underline{u}_1, \dots, \underline{u}_h$ è una base di U_α, allora $\overline{\underline{u}_1}, \dots, \overline{\underline{u}_h}$ è una base di $\hat{U}_\alpha$.
Osserviamo anche che la restrizione di $f_{\mathbb{C}}$ a $\tilde{W}_\lambda$ o a $U_\alpha \oplus \hat{U}_\alpha$ è il complessificato della restrizione di f a W_λ o Z_α rispettivamente.

Ci siamo quindi ridotti a studiare i seguenti due casi:

1. W spazio vettoriale reale di dimensione finita, $g \in \mathrm{End}(W)$, $g_{\mathbb{C}} \in \mathrm{End}(W_{\mathbb{C}})$ con $p_g(t) = p_{g_{\mathbb{C}}}(t) = (t-\lambda)^m$, $\lambda \in \mathbb{R}$;

2. Z spazio vettoriale reale di dimensione finita, $g \in \mathrm{End}(Z)$, $g_{\mathbb{C}} \in \mathrm{End}(Z_{\mathbb{C}})$ con $p_g(t) = q_\alpha(t)^n$, $p_{g_{\mathbb{C}}}(t) = (t-\alpha)^n(t-\overline{\alpha})^n$, $\alpha \in \mathbb{C}\setminus\mathbb{R}$.

Nel caso 1, g è triangolabile e una base di W di Jordan per g è anche una base reale di $W_{\mathbb{C}}$ di Jordan per $g_{\mathbb{C}}$. Le due forme normali di Jordan coincidono.

Nel caso 2, g non è triangolabile (ha spettro vuoto). Notiamo che $Z_{\mathbb{C}} = U \oplus \overline{U}$ dove $U = \mathrm{Ker}(f_{\mathbb{C}} - \alpha id_{V_{\mathbb{C}}})^n$ è l'autospazio generalizzato di $g_{\mathbb{C}}$ relativo α e $\overline{U}$ è l'autospazio generalizzato di $g_{\mathbb{C}}$ relativo $\overline{\alpha}$.
L'osservazione chiave è che se $\mathcal{B} = \{\underline{z}_1, \dots, \underline{z}_n\}$ è una base di U di Jordan per $g_{\mathbb{C}}|_U$, allora $\overline{\mathcal{B}} = \{\overline{\underline{z}_1}, \dots, \overline{\underline{z}_n}\}$ è una base di $\overline{U}$ di Jordan per $g_{\mathbb{C}}|_{\overline{U}}$. Inoltre, esse determinano la stessa forma normale di Jordan con l'unica differenza che in un caso i blocchi di Jordan sono relativi all'autovalore α, nell'altro sono relativi all'autovalore $\overline{\alpha}$.
Infatti, fissato un indice $i = 1\_\_n$ ci sono due possibilità: $\underline{z}_i$ è un autovettore per g, $g_{\mathbb{C}}(\underline{z}_i) = \alpha\underline{z}_i$, oppure $g_{\mathbb{C}}(\underline{z}_i) = \alpha\underline{z}_i + \underline{z}_{i-1}$. Passando al coniugio otteniamo $g_{\mathbb{C}}(\overline{\underline{z}_i}) = \overline{g_{\mathbb{C}}(\underline{z}_i)} = \overline{\alpha}\,\overline{\underline{z}_i}$ oppure $g_{\mathbb{C}}(\overline{\underline{z}_i}) = \overline{\alpha}\,\overline{\underline{z}_i} + \overline{\underline{z}_{i-1}}$.
Si può anche notare che $\mathrm{Ker}(g_{\mathbb{C}} - \overline{\alpha} id_{V_{\mathbb{C}}})^k = \overline{\mathrm{Ker}(g_{\mathbb{C}} - \alpha id_{V_{\mathbb{C}}})^k}$, e quindi i due nuclei sono isomorfi per ogni k. Ne segue che le stringhe invarianti di $g_{\mathbb{C}}$ relative agli autovalori α e $\overline{\alpha}$ sono uguali.
Abbiamo ottenuto che $\mathcal{B}_{\mathbb{C}} = \mathcal{B} \cup \overline{\mathcal{B}}$ è una base di $Z_{\mathbb{C}}$ di Jordan per $g_{\mathbb{C}}$ e che, a meno di riordinare i vettori di $\mathcal{B}_{\mathbb{C}}$, la forma normale di Jordan di g è del tipo $J = diag(J(\alpha, k_1), J(\overline{\alpha}, k_1), \dots, J(\alpha, k_s)J(\overline{\alpha}, k_s))$ (s è la molteplicità geometrica dei due autovalori).
Ci siamo ridotti a studiare il caso in cui la base $\mathcal{B}_{\mathbb{C}}$ di $Z_{\mathbb{C}}$ di Jordan per $g_{\mathbb{C}}$ dia

una forma di Jordan con solo due blocchi $J = diag(J(\alpha, n), J(\overline{\alpha}, n))$.
L'idea adesso è quella di prendere le parti reali e le parti immaginarie dei vettori di $\mathcal{B}$.
Notiamo che $\mathrm{Span}(\underline{z}_i, \overline{\underline{z}_i}) = \mathrm{Span}(\Re(\underline{z}_i), \Im(\underline{z}_i))$.
Infatti, $\Re(\underline{z}_i) = \frac{\underline{z}_i + \overline{\underline{z}_i}}{2}$, $\Im(\underline{z}_i) = \frac{\underline{z}_i - \overline{\underline{z}_i}}{2i}$ e $\underline{z}_i = \Re(\underline{z}_i) + i\Im(\underline{z}_i)$, $\overline{\underline{z}_i} = \Re(\underline{z}_i) - i\Im(\underline{z}_i)$ mostrano le due inclusioni.
Possiamo allora considerare $\mathcal{B}_{\mathbb{R}} = \{\Re(\underline{z}_1), \Im(\underline{z}_1), \ldots, \Re(\underline{z}_n), \Im(\underline{z}_n)\}$ ed ottenere una base reale di $Z_{\mathbb{C}}$ (e quindi di Z). Tale base è detta *base di Jordan reale* per g e la matrice di g (o di $g_{\mathbb{C}}$) in tale base $M^{\mathcal{B}_{\mathbb{R}}}_{\mathcal{B}_{\mathbb{R}}}(g_{\mathbb{C}}) = M^{\mathcal{B}_{\mathbb{R}}}_{\mathcal{B}_{\mathbb{R}}}(g) \in M(n, \mathbb{R})$ è detta *forma normale di Jordan reale* per g.
Per vedere di che tipo sia tale matrice, analizziamo dapprima il caso in cui $\underline{z}_i$ sia un autovettore: $g_{\mathbb{C}}(\underline{z}_i) = \alpha \underline{z}_i$, $g_{\mathbb{C}}(\overline{\underline{z}_i}) = \overline{\alpha}\, \overline{\underline{z}_i}$. Allora

$$\begin{aligned} g(\Re(\underline{z}_i)) &= \frac{g_{\mathbb{C}}(\underline{z}_i) + g_{\mathbb{C}}(\overline{\underline{z}_i})}{2} = \frac{\alpha \underline{z}_i + \overline{\alpha}\, \overline{\underline{z}_i}}{2} = \\ &= \frac{\alpha(\Re(z_i) + i\Im(z_i)) + \overline{\alpha}(\Re(z_i) - i\Im(z_i))}{2} = \Re(\alpha)\Re(\underline{z}_i) - \Im(\alpha)\Im(\underline{z}_i) \\ g(\Im(\underline{z}_i)) &= \frac{g_{\mathbb{C}}(\underline{z}_i) - g_{\mathbb{C}}(\overline{\underline{z}_i})}{2i} = \frac{\alpha \underline{z}_i - \overline{\alpha}\, \overline{\underline{z}_i}}{2i} = \\ &= \frac{\alpha(\Re(z_i) + i\Im(z_i)) - \overline{\alpha}(\Re(z_i) - i\Im(z_i))}{2i} = \Im(\alpha)\Re(\underline{z}_i) + \Re(\alpha)\Im(\underline{z}_i). \end{aligned}$$

Se invece $g_{\mathbb{C}}(\underline{z}_i) = \alpha \underline{z}_i + \underline{z}_{i-1}$, $g_{\mathbb{C}}(\overline{\underline{z}_i}) = \overline{\alpha}\, \overline{\underline{z}_i} + \overline{\underline{z}_{i-1}}$, allora

$$\begin{aligned} g(\Re(\underline{z}_i)) &= \Re(\alpha)\Re(\underline{z}_i) - \Im(\alpha)\Im(\underline{z}_i) + \frac{\underline{z}_{i-1} + \overline{\underline{z}_{i-1}}}{2} = \\ &= \Re(\alpha)\Re(\underline{z}_i) - \Im(\alpha)\Im(\underline{z}_i) + \Re(\underline{z}_{i-1}) \\ g(\Im(\underline{z}_i)) &= \Im(\alpha)\Re(\underline{z}_i) + \Re(\alpha)\Im(\underline{z}_i) + \frac{\underline{z}_{i-1} - \overline{\underline{z}_{i-1}}}{2i} = \\ &= \Re(\alpha)\Re(\underline{z}_i) - \Im(\alpha)\Im(\underline{z}_i) + \Im(\underline{z}_{i-1}) \end{aligned}$$

Otteniamo che la forma normale reale è triangolare superiore a blocchi, con lungo diagonale n copie della matrice $\begin{pmatrix} \Re(\alpha) & \Im(\alpha) \\ -\Im(\alpha) & \Re(\alpha) \end{pmatrix}$ e sopra la diagonale $n-1$ copie della matrice I_2.
Operativamente, per ottenere la forma normale di Jordan reale, si eliminano dalla forma normale di Jordan complessa i blocchi relativi a $\overline{\alpha}$ e si "esplodono" i blocchi relativi a α (che raddoppiano di taglia). Per esplodere una matrice complessa, rimpiazzare ogni elemento $z \in \mathbb{C}$ con la matrice $\begin{pmatrix} \Re(z) & \Im(z) \\ -\Im(z) & \Re(z) \end{pmatrix}$.

Osserviamo che le matrici di taglia 2×2 reali nella forma $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ formano un campo isomorfo a $\mathbb{C}$.

Tornando al caso generale, la forma normale di Jordan reale di un endomorfismo reale contiene gli usuali blocchi di Jordan relativi agli autovalori reali e l'esplosione di metà dei blocchi della forma normale di Jordan complessa relativi agli autovalori non reali.

Inoltre, la forma normale di Jordan reale, essendo univocamente determinata dalla forma normale di Jordan del complessificato, è un invariante completo per coniugazione per endomorfismi reali.

Osserviamo che dalla discussione sopra si ha che trovare la forma normale di Jordan complessa di un endomorfismo reale o una base di Jordan complessa è reso più semplice dal fatto che i contributi di un autovalore complesso non reale e del suo coniugato sono uguali: per trovare i blocchi di Jordan relativi all'autovalore $\overline{\alpha}$ basta coniugare i blocchi di Jordan relativi all'autovalore α; per trovare una base di Jordan per i blocchi di Jordan relativi all'autovalore $\overline{\alpha}$ basta coniugare una base di Jordan per i blocchi di Jordan relativi all'autovalore α.

Prodotti Scalari

Conveniamo che nel seguito tutti gli spazi vettoriali saranno di dimensione finita, anche se gran parte delle definizioni che daremo e parte della teoria che svilupperemo hanno senso anche in dimensione arbitraria.

Dati V, W spazi vettoriali su $\mathbb{K}$, consideriamo $Bil(V \times W, \mathbb{K})$, l'insieme delle applicazioni bilineari da $V \times W$ a $\mathbb{K}$, i cui elementi si chiamano *forme bilineari*. Ricordiamo che $Bil(V \times W, \mathbb{K})$ è uno spazio vettoriale su $\mathbb{K}$.

Una $\phi \in Bil(V \times W, \mathbb{K})$ è una $\phi : V \times W \to \mathbb{K}$ lineare in entrambi gli argomenti: per ogni $\underline{v}_1, \underline{v}_2 \in V$, $\underline{w}_1, \underline{w}_2 \in W$, $\mu \in \mathbb{K}$,

$$\phi(\underline{v}_1 + \underline{v}_2, \underline{w}_1) = \phi(\underline{v}_1, \underline{w}_1) + \phi(\underline{v}_2, \underline{w}_1),$$

$$\phi(\underline{v}_1, \underline{w}_1 + \underline{w}_2) = \phi(\underline{v}_1, \underline{w}_1) + \phi(\underline{v}_1, \underline{w}_2),$$

$$\phi(\mu\underline{v}_1, \underline{w}_1) = \phi(\underline{v}_1, \mu\underline{w}_1) = \mu\phi(\underline{v}_1, \underline{w}_1).$$

Ricordiamo che queste richieste sono equivalenti a dire che comunque fissiamo il primo argomento, facendo variare il secondo otteniamo un funzionale $\phi(\underline{v}_0, \cdot)$ su W, mentre comunque fissiamo il secondo argomento, facendo variare il primo otteniamo un funzionale $\phi(\cdot, \underline{w}_0)$ su V.
Abbiamo quindi due applicazioni

$$F_{\phi s} : V \to W^*, \quad \underline{v}_0 \mapsto \phi(\underline{v}_0, \cdot),$$

$$F_{\phi d} : W \to V^*, \quad \underline{w}_0 \mapsto \phi(\cdot, \underline{w}_0),$$

entrambe lineari, dette *omomorfismi di rappresentazione*.
Infatti, mostriamo che $F_{\phi s}$ è lineare, l'argomento è identico per $F_{\phi d}$: per ogni $\underline{v}_1, \underline{v}_2 \in V$,
$F_{\phi s}(\underline{v}_1 + \underline{v}_2) = \phi(\underline{v}_1 + \underline{v}_2, \cdot) = \phi(\underline{v}_1, \cdot) + \phi(\underline{v}_2, \cdot) = F_{\phi s}(\underline{v}_1) + F_{\phi s}(\underline{v}_2)$;
per ogni $\underline{v} \in V$, $\mu \in \mathbb{K}$,
$F_{\phi s}(\mu\underline{v}) = \phi(\mu\underline{v}, \cdot) = \mu\phi(\underline{v}, \cdot) = \mu F_{\phi s}(\underline{v})$.

Analogamente al nucleo di una applicazione lineare, possiamo definire il *radicale destro* e il *radicale sinistro* di una forma bilineare (l'immagine di una forma bilineare non nulla è $\mathbb{K}$): data $\phi \in Bil(V \times W, \mathbb{K})$ poniamo

$$Rad_d(\phi) = \operatorname{Ker} F_{\phi d} = \{\underline{w} \in W |\ \phi(\underline{v}, \underline{w}) = 0\ \forall \underline{v} \in V\}$$

$$Rad_s(\phi) = \operatorname{Ker} F_{\phi s} = \{\underline{v} \in V |\ \phi(\underline{v}, \underline{w}) = 0\ \forall \underline{w} \in W\}$$

ed è immediato che $Rad_d(\phi)$ è un sottospazio di W, $Rad_s(\phi)$ è un sottospazio di V. Inoltre notiamo che

$$Rad_d(\phi) = \bigcap_{\underline{v} \in V} \operatorname{Ker} F_{\phi s}(\underline{v}) = Z(\operatorname{Im} F_{\phi s}),$$

$$Rad_s(\phi) = \bigcap_{\underline{w}\in W} \operatorname{Ker} F_{\phi d}(\underline{w})) = Z(\operatorname{Im} F_{\phi d}),$$

per cui, passando all'annullatore, siamo in grado di determinare le immagini degli omomorfismi di reppresentazione:

$$\operatorname{Im} F_{\phi d} = Ann(Rad_s(\phi)), \qquad \operatorname{Im} F_{\phi s} = Ann(Rad_d(\phi)).$$

Osserviamo che in dimensione infinita, i nuclei degli degli omomorfismi di rappresentazione continuano a coincidere con i radicali, mentre le loro immagini sono, in generale, solo contenute negli annullatori dei radicali.

ϕ si dice *non degenere a destra* se $Rad_d(\phi) = \{\underline{0}\}$, *non degenere a sinistra* se $Rad_s(\phi) = \{\underline{0}\}$. Inoltre, ϕ si dice *non degenere* se lo è sia a destra che a sinistra.

L'applicazione $^\top : Bil(V \times W, \mathbb{K}) \to Bil(W \times V, \mathbb{K})$, $\phi \mapsto \phi^\top$ che scambia gli argomenti, $\phi^\top(\underline{w}, \underline{v}) = \phi(\underline{v}, \underline{w})$ per ogni $\underline{v} \in V$, $\underline{w} \in W$, è un isomorfismo e ovviamente $F_{\phi^\top d} = F_{\phi s}$, $F_{\phi^\top s} = F_{\phi d}$, per cui $Rad_d(\phi^\top) = Rad_s(\phi)$, $Rad_s(\phi^\top) = Rad_d(\phi)$.

Date $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$, $\mathcal{D} = \{\underline{w}_1, \dots, \underline{w}_m\}$ basi di V e W rispettivamente, per ogni $\phi \in Bil(V \times W, \mathbb{K})$ costruiamo la matrice di taglia $n \times m$

$$M_{\mathcal{BD}}(\phi) = \big(\phi(\underline{v}_i, \underline{w}_j)\big)_{\substack{i=1\_\_n \\ j=1\_\_m}}$$

detta *matrice associata a* ϕ (o *che rappresenta* ϕ) *nelle basi* $\mathcal{B}$ *e* $\mathcal{D}$.

Se scriviamo $\underline{v} \in V$ come combinazione lineare dei $\underline{v}_i$ e scriviamo $\underline{w} \in W$ come combinazione lineare dei $\underline{w}_j$, abbiamo

$$\phi(\underline{v}, \underline{w}) = \phi(\sum_{i=1}^n a_i \underline{v}_i, \sum_{j=1}^m b_j \underline{w}_j) = \sum_{i=1}^n a_i \phi(\underline{v}_i, \sum_{j=1}^m b_j \underline{w}_j) =$$

$$= \sum_{i=1}^n \sum_{j=1}^m a_i b_j \phi(\underline{v}_i, \underline{w}_j) = [\underline{v}]_{\mathcal{B}}^\top M_{\mathcal{BD}}(\phi) [\underline{w}]_{\mathcal{D}}$$

Inoltre, per ogni $\phi_1, \phi_2 \in Bil(V \times W, \mathbb{K})$, $\mu \in \mathbb{K}$,

$$(\phi_1 + \phi_2)(\underline{v}_i, \underline{w}_j) = \phi_1(\underline{v}_i, \underline{w}_j) + \phi_2(\underline{v}_i, \underline{w}_j), \text{ e } (\mu\phi_1)(\underline{v}_i, \underline{w}_j) = \mu\phi_1(\underline{v}_i, \underline{w}_j),$$

quindi

$$M_{\mathcal{BD}}(\phi_1 + \phi_2) = M_{\mathcal{BD}}(\phi_1) + M_{\mathcal{BD}}(\phi_2), \text{ e } M_{\mathcal{BD}}(\mu\phi) = \mu M_{\mathcal{BD}}(\phi).$$

Otteniamo $M_{\mathcal{BD}} : Bil(V \times W, \mathbb{K}) \to M(n, m, \mathbb{K})$ lineare che è un isomorfismo.
Infatti, se $\phi \in Bil(V \times W, \mathbb{K})$ è tale che $M_{\mathcal{BD}}(\phi) = 0$, allora per ogni $\underline{v} \in V$, $\underline{w} \in W$, $\phi(\underline{v}, \underline{w}) = [\underline{v}]_{\mathcal{B}}^\top 0 [\underline{w}]_{\mathcal{D}} = 0$, per cui $\phi = 0$, quindi $M_{\mathcal{BD}}$ è iniettiva.
Data $M \in M(n, m, \mathbb{K})$, definiamo $\phi : V \times W \to \mathbb{K}$ ponendo, per ogni $\underline{v} \in V$,

$\underline{w} \in W$, $\phi(\underline{v}, \underline{w}) = [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}}$. ϕ è bilineare in quanto:

- $\phi(\underline{v}_1 + \underline{v}_2, w) = [\underline{v}_1 + \underline{v}_2]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} = ([\underline{v}_1]_{\mathcal{B}} + [\underline{v}_2]_{\mathcal{B}})^{\top} M [\underline{w}]_{\mathcal{D}} =$
  $= ([\underline{v}_1]_{\mathcal{B}}^{\top} + [\underline{v}_2]_{\mathcal{B}}^{\top}) M [\underline{w}]_{\mathcal{D}} = [\underline{v}_1]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} + [\underline{v}_2]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} =$
  $= \phi(\underline{v}_1, \underline{w}) + \phi(\underline{v}_2, \underline{w})$;
- $\phi(\mu \underline{v}, w) = [\mu \underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} = \mu [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} = \mu \phi(\underline{v}, \underline{w})$;
- $\phi(\underline{v}, \underline{w}_1 + \underline{w}_2) = [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}_1 + \underline{w}_2]_{\mathcal{D}} = [\underline{v}]_{\mathcal{B}}^{\top} M ([\underline{w}_1]_{\mathcal{D}} + [\underline{w}_2]_{\mathcal{D}}) =$
  $= [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}_1]_{\mathcal{D}} + [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}_2]_{\mathcal{D}} = \phi(\underline{v}, \underline{w}_1) + \phi(\underline{v}, \underline{w}_2)$;
- $\phi(\underline{v}, \mu \underline{w}) = [\underline{v}]_{\mathcal{B}}^{\top} M [\mu \underline{w}]_{\mathcal{D}} = \mu [\underline{v}]_{\mathcal{B}}^{\top} M [\underline{w}]_{\mathcal{D}} = \mu \phi(\underline{v}, \underline{w})$;

per ogni $\underline{v}, \underline{v}_1, \underline{v}_2 \in V, \underline{w}, \underline{w}_1, \underline{w}_2 \in W, \mu \in \mathbb{K}$.
Inoltre, la matrice associata a ϕ nelle basi $\mathcal{B}$ e $\mathcal{D}$ è M, quindi $M_{\mathcal{B}\mathcal{D}}$ è surgettiva.

Abbiamo quindi $\dim Bil(V \times W, \mathbb{K}) = \dim V \dim W$.

Osserviamo che $M_{\mathcal{D}\mathcal{B}}(\phi^{\top}) = M_{\mathcal{B}\mathcal{D}}(\phi)^{\top}$, ovvero che abbiamo il diagramma commutativo

$$\begin{array}{ccc} Bil(V,W) & \xrightarrow{\top} & Bil(W,V) \\ {\scriptstyle M_{\mathcal{B}\mathcal{D}}}\downarrow & \circlearrowleft & \downarrow{\scriptstyle M_{\mathcal{D}\mathcal{B}}} \\ M(n,m,\mathbb{K}) & \xrightarrow{\top} & M(m,n,\mathbb{K}) \end{array}$$

Nel caso $V = \mathbb{K}^n$, $W = \mathbb{K}^m$, usando le basi canoniche otteniamo l'isomorfismo canonico $M_{\mathcal{C}an\mathcal{C}an} : Bil(\mathbb{K}^n \times \mathbb{K}^m, \mathbb{K}) \to M(n, m, \mathbb{K})$, per cui ogni forma bilineare ϕ su $\mathbb{K}^n \times \mathbb{K}^m$ si scrive in modo unico come $\phi = \Phi_A$, dove $\Phi_A(\underline{x}, \underline{y}) = \underline{x}^{\top} A \underline{y}$, per ogni $\underline{x} \in \mathbb{K}^n, \underline{y} \in \mathbb{K}^m$, con $A \in M(n, m, \mathbb{K})$.

Osservazioni
▶ Notiamo che, coerentemente con la definizione di matrice associata ad una forma bilineare, per ogni i, j si ha $\underset{i-}{\underset{|}{A}}\overset{j}{\underset{|}{}} = \underline{e}_i^{\top} A \underline{e}_j = \Phi_A(\underline{e}_i, \underline{e}_j)$.
▶ Questo dà direttamente l'iniettività di $M_{\mathcal{C}an\mathcal{C}an}$:
$\underline{x}^{\top} A \underline{y} = 0$ per ogni $\underline{x} \in \mathbb{K}^n, \underline{y} \in \mathbb{K}^m \iff A = 0$
(facendo variare $\underline{x}$ e $\underline{y}$ sulle basi canoniche si ottengono tutti i coefficienti di A).
▶ Inoltre, $\underline{x}^{\top} A \underline{y} = \underline{0}$ per ogni $\underline{y} \in \mathbb{K}^m \iff A^{\top} \underline{x} = 0$
(facendo variare $\underline{y}$ sulla base canonica si ottiene che per ogni colonna di A, $x^{\top} \underset{|}{\overset{j}{A}} = 0$ e quindi $\underline{x}^{\top} A = 0$, poi si passa alla trasposta).
▶ Analogamente, $\underline{x}^{\top} A \underline{y} = \underline{0}$ per ogni $\underline{x} \in \mathbb{K}^n \iff A \underline{y} = 0$
(facendo variare $\underline{x}$ sulla base canonica si ottiene che per ogni riga di A, $\underset{i-}{A}\underline{y} = 0$ e quindi $A\underline{y} = 0$).

▶ In generale, se $A = M_{\mathcal{BD}}(\phi)$ otteniamo il diagramma commutativo

$$\begin{array}{ccc} V \times W & \xrightarrow{\phi} & \mathbb{K} \\ {\scriptstyle [\ ]_{\mathcal{B}}}\downarrow \quad \downarrow{\scriptstyle [\ ]_{\mathcal{D}}} & \circlearrowleft & \downarrow{\scriptstyle id_{\mathbb{K}}} \\ \mathbb{K}^n \times \mathbb{K}^m & \xrightarrow{\Phi_A} & \mathbb{K} \end{array}$$

Se $\mathcal{B}'$, $\mathcal{D}'$ sono altre basi per V e W, allora per ogni $\underline{v} \in V, \underline{w} \in W$

$$\begin{aligned} [\underline{v}]_{\mathcal{B}'}^{\top} M_{\mathcal{B}'\mathcal{D}'}(\phi)[\underline{w}]_{\mathcal{D}'} = \phi(\underline{v},\underline{w}) = [\underline{v}]_{\mathcal{B}}^{\top} M_{\mathcal{BD}}(\phi)[\underline{w}]_{\mathcal{D}} = \\ = (M_{\mathcal{B}}^{\mathcal{B}'}(id_V)[\underline{v}]_{\mathcal{B}'})^{\top} M_{\mathcal{BD}}(\phi)(M_{\mathcal{D}}^{\mathcal{D}'}(id_W)[\underline{w}]_{\mathcal{D}'}) = \\ = [\underline{v}]_{\mathcal{B}'}^{\top}(M_{\mathcal{B}}^{\mathcal{B}'}(id_V)^{\top} M_{\mathcal{BD}}(\phi) M_{\mathcal{D}}^{\mathcal{D}'}(id_W))[\underline{w}]_{\mathcal{D}'} \end{aligned}$$

per cui

$$M_{\mathcal{B}'\mathcal{D}'}(\phi) = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)^{\top} M_{\mathcal{BD}}(\phi) M_{\mathcal{D}}^{\mathcal{D}'}(id_W)$$

esprime come cambia la matrice associata ad una forma bilineare cambiando le basi.
Osserviamo che il rango della matrice associata ad una forma bilineare ϕ non dipende dalle basi scelte.
È quindi ben definito il rango di ϕ come $\operatorname{rnk}\phi = \operatorname{rnk} M_{\mathcal{BD}}(\phi)$ ($\mathcal{B}$ una qualsiasi base di V, $\mathcal{D}$ una qualsiasi base di W).

Posto $A = M_{\mathcal{BD}}(\phi)$, osserviamo che $Rad_d(\phi)$ corrisponde, tramite l'isomorfismo delle coordinate nella base $\mathcal{D}$, al sottospazio $\{\underline{y} \in \mathbb{K}^m |\ \underline{x}^{\top} A \underline{y} = 0\ \forall \underline{x} \in \mathbb{K}^n\}$, che per quanto osservato sopra è $\operatorname{Ker} A$ ed ha quindi dimensione $\dim W - \operatorname{rnk}\phi$.
Quindi, ϕ è non degenere a destra se e solo se $\operatorname{rnk}\phi = \dim W$.
Analogamente, $Rad_s(\phi)$ corrisponde, tramite l'isomorfismo delle coordinate nella base $\mathcal{B}$, al sottospazio $\{\underline{x} \in \mathbb{K}^n |\ \underline{x}^{\top} A \underline{y} = 0\ \forall \underline{y} \in \mathbb{K}^m\}$, che per quanto osservato sopra è $\operatorname{Ker} A^{\top}$ ed ha quindi dimensione $\dim V - \operatorname{rnk}\phi$.
Quindi, ϕ è non degenere a sinistra se e solo se $\operatorname{rnk}\phi = \dim V$.
Osserviamo che ϕ non degenere $\Rightarrow \dim V = \dim W$ e che se $\dim V = \dim W$, ϕ è non degenere a destra se e solo se lo è a sinistra.

Da qui in avanti ci occuperemo del caso $W = V$, $\dim V = n$, e indicheremo $Bil(V \times V, \mathbb{K})$ semplicemente con $Bil(V)$, che chiameremo lo spazio delle forme bilineari su V.

Per $\phi \in Bil(V)$, oltre ai radicali destri e sinistri, che in questo caso hanno la stessa dimensione e diventano

$$Rad_d(\phi) = \{\underline{w} \in V |\ \phi(\underline{v},\underline{w}) = 0\ \forall \underline{v} \in V\}$$

$$Rad_s(\phi) = \{\underline{v} \in V |\ \phi(\underline{v},\underline{w}) = 0\ \forall \underline{w} \in V\},$$

è ben definito il *cono isotropo* di ϕ,

$$CI(\phi) = \{\underline{v} \in V |\ \phi(\underline{v},\underline{v}) = 0\}$$

i cui elementi si dicono *vettori isotropi* (per ϕ).

Notiamo che, in generale, il cono isotropo non è un sottospazio (è però chiuso per prodotto per scalari e contiene $\underline{0}$) e che contiene i due radicali.
La forma bilineare ϕ si dice *anisotropa* se $\underline{0} \in V$ è l'unico vettore isotropo, ovvero se $CI(\phi) = \{\underline{0}\}$. Notiamo che se ϕ è anisotropa, è anche non degenere.

Fissata una base $\mathcal{B}$ di V, indichiamo con $M_{\mathcal{B}}(\phi) = M_{\mathcal{B}\mathcal{B}}(\phi)$, detta la matrice associata alla forma bilineare ϕ nella base $\mathcal{B}$. Chiaramente, $M_{\mathcal{B}}(\phi) \in M(n, \mathbb{K})$ è quadrata. $M_{\mathcal{B}} : Bil(V) \to M(n, \mathbb{K})$ è un isomorfismo e quindi $Bil(V)$ ha dimensione n^2.
Ad esempio, data $A \in M(n, \mathbb{K})$, la matrice associata a $\Phi_A \in Bil(\mathbb{K}^n)$ nella base canonica è proprio A.
Se $\mathcal{B}'$ è un'altra base di V, abbiamo $M_{\mathcal{B}'}(\phi) = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)^\top M_{\mathcal{B}}(\phi) M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$.
Come nel caso generale, è ben definito $\operatorname{rnk}\phi$ e coincide con il rango di una qualsiasi matrice associata a ϕ. In particolare, ϕ è non degenere se e solo se una qualsiasi matrice associata a ϕ è invertibile.

Definizione:
Una $\phi \in Bil(V)$ si dice:
• *simmetrica* o *prodotto scalare* se $\phi^\top = \phi$, ovvero se per ogni $\underline{v}, \underline{w} \in V$, $\phi(\underline{v}, \underline{w}) = \phi(\underline{w}, \underline{v})$.
• *antisimmetrica* se $\phi^\top = -\phi$, ovvero se per ogni $\underline{v}, \underline{w} \in V$, $\phi(\underline{v}, \underline{w}) = -\phi(\underline{w}, \underline{v})$.

Ad esempio, la formula $\langle \underline{x}, \underline{y} \rangle = \underline{x}^\top \underline{y}$ definisce un prodotto scalare $\langle \cdot, \cdot \rangle$ su $\mathbb{K}^n$, detto *prodotto scalare standard*. Notiamo che $\langle \cdot, \cdot \rangle = \Phi_{I_n}$, quindi $\langle \cdot, \cdot \rangle$ è non degenere.

L'insieme dei prodotti scalari su V si indica con $PS(V)$ ed è un sottospazio di $Bil(V)$. Il valore di ϕ sulla coppia $(\underline{v}, \underline{w})$, $\phi(\underline{v}, \underline{w})$, si dice *prodotto scalare* (tramite ϕ) di $\underline{v}$ e $\underline{w}$.
L'insieme delle forme bilineari antisimmetriche su V si indica con $A(V)$ ed è un sottospazio di $Bil(V)$.

Osservazioni:
▶ Se $\mathbb{K}$ ha caratteristica diversa da 2, una ϕ antisimmetrica è alternante, e quindi una 2-forma su V: $A(V) = \Lambda^2(V)$.
▶ Per $V = \mathbb{K}^n$, i prodotti scalari sono dati dalle Φ_A con $A \in M(n, \mathbb{K})$ simmetrica; le forme bilineari antisimmetriche sono date dalle Φ_A con $A \in M(n, \mathbb{K})$ antisimmetrica.
▶ Più in generale, se $\phi \in Bil(V)$ è un prodotto scalare (risp. una forma bilineare antisimmetrica) allora $M_{\mathcal{B}}(\phi)$ è simmetrica (risp. antisimmetrica) per ogni base $\mathcal{B}$ di V.
▶ Viceversa, se esiste una base $\mathcal{B}$ di V per cui $M_{\mathcal{B}}(\phi)$ è simmetrica (risp. antisimmetrica), allora ϕ è un prodotto scalare (risp. una forma bilineare antisimmetrica).
▶ Se $\mathbb{K}$ ha caratteristica 2, $PS(V) = A(V)$. In effetti, la teoria in questo caso differisce sostanzialmente dal caso generale.

▶ Se $\mathbb{K}$ ha caratteristica diversa da 2, $Bil(V) = PS(V) \oplus A(V)$.
Infatti, in maniera analoga alla trasposta di matrici quadrate, l'endomorfismo $^\top : Bil(V) \to Bil(V)$ ha ordine 2, $(^\top)^2 = id_{Bil(V)}$, quindi è diagonalizzabile con autospazi $V_1(^\top) = PS(V)$, $V_{-1}(^\top) = A(V)$.
Notare che questo corrisponde (tramite $M_{\mathcal{B}}$) al fatto che lo spazio delle matrici quadrate è la somma diretta dello spazio delle matrici simmetriche e dello spazio delle matrici antisimmetriche: $M(n, \mathbb{K}) = S(n, \mathbb{K}) \oplus A(n, \mathbb{K})$.
Più esplicitamente, data $\phi \in Bil(V)$, $\phi + \phi^\top$ è un prodotto scalare, $\phi - \phi^\top$ è antisimmetrica e $\phi = \frac{1}{2}(\phi + \phi^\top) + \frac{1}{2}(\phi - \phi^\top)$.
Si ha quindi $\dim PS(V) = \frac{n(n+1)}{2}$, $\dim A(V) = \frac{n(n-1)}{2}$.

Osserviamo che se ϕ è una forma antisimmetrica $F_{\phi s} = -F_{\phi d}$, mentre se ϕ è un prodotto scalare $F_{\phi s} = F_{\phi d}$ che denoteremo semplicemente con F_ϕ.
In entrambi i casi, $Rad_d(\phi) = Rad_s(\phi)$, che quindi indicheremo semplicamemnte con $Rad(\phi)$, il *radicale* di ϕ

$$Rad(\phi) = \{\underline{v} \in V |\ \phi(\underline{v}, \underline{w}) = 0 \text{ per ogni } \underline{w} \in V\}$$

e per un prodotto scalare abbiamo $\operatorname{Ker} F_\phi = Rad(\phi)$ e $\operatorname{Im} F_\phi = Ann(Rad(\phi))$.

In entrambi i casi, $Rad(\phi)$ corrisponde, tramite l'isomorfismo delle coordinate in una base $\mathcal{B}$ di V, al nucleo della matrice associata a ϕ nella base $\mathcal{B}$:

$$[Rad(\phi)]_{\mathcal{B}} = \operatorname{Ker} M_{\mathcal{B}}(\phi).$$

Ricordiamo che ϕ è non degenere se e solo se $M_{\mathcal{B}}(\phi)$ è invertibile.

Da qui in avanti ci occuperemo della teoria dei prodotti scalari in dimesione finita.
Come notazione, indicheremo con (V, ϕ) uno spazio vettoriale V di dimensione finita munito del fissato prodotto scalare $\phi \in PS(V)$.

Definizione:
Siano (V, ϕ) e (W, ψ) due spazi vettoriali su $\mathbb{K}$ muniti di prodotto scalare.
Una $F : V \to W$ lineare si dice *isometria* se è un isomorfismo ed inoltre $\psi(F(\underline{v}_1), F(\underline{v}_2)) = \phi(\underline{v}_1, \underline{v}_2)$ per ogni $\underline{v}_1, \underline{v}_2 \in V$. Se esiste una tale isometria, (V, ϕ) e (W, ψ) si dicono *isometrici* (a volte abbreviato dicendo che ϕ e ψ sono isometrici).

Ad esempio, se $\mathcal{B}$ è una base di V, $[\ ]_{\mathcal{B}}$ è una isometria tra (V, ϕ) e $(\mathbb{K}^n, \Phi_{M_{\mathcal{B}}(\phi)})$.

Osserviamo che se (V, ϕ) e (W, ψ) sono isometrici, sono anche isomorfi. Inoltre, la relazione di essere isometrici è una specie di relazione di equivalenza sulla classe degli spazi vettoriali su $\mathbb{K}$ muniti di un prodotto scalare. Infatti, id_V è una isometria; l'inversa di una isometria è una isometria; la composizione di isometrie è una isometria.
Il nostro scopo è quello di studiare gli spazi vettoriali di dimensione finita muniti di prodotto scalare a meno di isometrie, ovvero trovare invarianti completi che

ci permettano di dire se due tali spazi siano isometrici o meno.
Vedremo una classificazione completa nei casi $\mathbb{K} = \mathbb{C}$ e $\mathbb{K} = \mathbb{R}$, ma molte delle cose che diremo saranno valide per ogni campo (anche se dovremo richiedere quasi subito $char(\mathbb{K}) \neq 2$).

Se F è una isometria tra (V, ϕ) e (W, ψ), allora $F(Rad(\phi)) = Rad(\psi)$. Infatti,
$$\begin{aligned} F(\underline{v}) \in Rad(\psi) &\iff \psi(F(\underline{v}), \underline{w}) = 0 \text{ per ogni } \underline{w} \in W \\ &\iff \psi(F(\underline{v}), F(\underline{u})) = 0 \text{ per ogni } \underline{u} \in V \text{ (essendo } F \text{ surgettiva)} \\ &\iff \phi(\underline{v}, \underline{u}) = 0 \text{ per ogni } \underline{u} \in V \iff \underline{v} \in Rad(\phi). \end{aligned}$$

Abbiamo quindi alcuni invarianti per isometria: la dimensione di V, la dimensione del radicale di ϕ, e quindi il rango di ϕ, $\operatorname{rnk}\phi = \dim V - \dim Rad(\phi)$. Notiamo che anche la proprietà di essere non degenere è invariante per isometria.

La condizione di essere una isometria la possiamo esplicitare anche in forma matriciale.
Sia F una isometria tra (V, ϕ) e (W, ψ) e siano $\mathcal{B}$ una base di V, $\mathcal{D}$ una base di W.
Per ogni $\underline{v}_1, \underline{v}_2 \in V$ abbiamo

$$\phi(\underline{v}_1, \underline{v}_2) = [\underline{v}_1]_{\mathcal{B}}^{\top} M_{\mathcal{B}}(\phi)[\underline{v}_2]_{\mathcal{B}}$$
$$\begin{aligned} \psi(F(\underline{v}_1), F(\underline{v}_2)) &= [F(\underline{v}_1)]_{\mathcal{D}}^{\top} M_{\mathcal{D}}(\psi)[F(\underline{v}_2)]_{\mathcal{D}} = \\ &= (M_{\mathcal{D}}^{\mathcal{B}}(F)[\underline{v}_1]_{\mathcal{B}})^{\top} M_{\mathcal{D}}(\psi)(M_{\mathcal{D}}^{\mathcal{B}}(F)[\underline{v}_2]_{\mathcal{B}}) = \\ &= [\underline{v}_1]_{\mathcal{B}}^{\top} \left(M_{\mathcal{D}}^{\mathcal{B}}(F)^{\top} M_{\mathcal{D}}(\psi) M_{\mathcal{D}}^{\mathcal{B}}(F)\right) [\underline{v}_2]_{\mathcal{B}} \end{aligned}$$

per cui
$$M_{\mathcal{B}}(\phi) = M_{\mathcal{D}}^{\mathcal{B}}(F)^{\top} M_{\mathcal{D}}(\psi) M_{\mathcal{D}}^{\mathcal{B}}(F).$$

Notiamo che questa formula è dello stesso tipo della formula che esprime come cambia la matrice di un prodotto scalare al variare della base: se $\mathcal{B}, \mathcal{B}'$ sono basi di V,
$$M_{\mathcal{B}'}(\phi) = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)^{\top} M_{\mathcal{B}}(\phi) M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$$
e che le due formule sono simili alle formule che esprimono la coniugazione o il cambio di base per endomorfismi (in queste la matrice trasposta sostituisce la matrice inversa).

Possiamo allora definire una ulteriore relazione di equivalenza su $M(n, \mathbb{K})$, detta *congruenza*:
$A, B \in M(n, \mathbb{K})$ si dicono *congruenti*, si scrive $A \equiv B$, se esite $P \in GL(n, \mathbb{K})$ tale che $B = P^{\top} A P$.
Notiamo che se A è simmetrica e $B \equiv A$, allora anche B è simmetrica. Possiamo quindi restringere la relazione di congruenza allo spazio delle matrici simmetriche $S(n, \mathbb{K})$ (lo stesso vale per lo spazio delle matrici antisimmetriche $A(n, \mathbb{K})$).
Studiare i prodotti scalari su $\mathbb{K}^n$ a meno di isometrie è la stessa cosa che studiare il quoziente $S(n, \mathbb{K})/_{\equiv}$.

Osserviamo inoltre che, essendo la congruenza un caso particolare di DS-equivalenza, il rango è un invariante per congruenza.

In modo del tutto analogo a quanto visto per la coniugazione/similitudine, abbiamo la seguente:

Proposizione
Siano $\phi \in PS(V)$, $\psi \in PS(W)$. I seguenti fatti sono equivalenti:

1. (V, ϕ) e (W, ψ) sono isometrici;
2. Per ogni base $\mathcal{B}$ di V e per ogni base $\mathcal{D}$ di W, $M_{\mathcal{B}}(\phi)$ e $M_{\mathcal{D}}(\psi)$ sono congruenti;
3. Esiste un base $\mathcal{B}$ di V ed esiste una base $\mathcal{D}$ di W tali che $M_{\mathcal{B}}(\phi)$ e $M_{\mathcal{D}}(\psi)$ sono congruenti;
4. Esiste un base $\mathcal{B}$ di V ed esiste una base $\mathcal{D}$ di W tali che $M_{\mathcal{B}}(\phi) = M_{\mathcal{D}}(\psi)$.

Osserviamo che fissato un isomorfismo $f : V \to W$, e dato un prodotto scalare ψ su W, possiamo definire un prodotto scalare $f^*\psi$ su V tramite la formula

$$f^*\psi(\underline{v}_1, \underline{v}_2) = \psi(f(\underline{v}_1), f(\underline{v}_2)), \quad \text{per ogni } \underline{v}_1, \underline{v}_2 \in V.$$

$f^*\psi$ si dice *rimontato* o *pull-back* di ψ tramite f.
È immediato osservare che f diventa una isometria da $(V, f^*\psi)$ a (W, ψ) che sono quindi isometrici.
Con queste notazioni, $F \in \mathrm{Hom}(V, W)$ isomorfismo è una isometria tra (V, ϕ) e (W, ψ) se e solo se $\phi = F^*\psi$.

Supponiamo di avere due spazi vettoriali su $\mathbb{K}$ muniti di prodotti scalari, (W, ϕ) e (Z, ψ) con W e Z isomorfi. Fissiamo un terzo spazio vettoriale V isomorfo a W e a Z tramite gli isomorfismi $f : V \to W$, $g : V \to Z$.
Possiamo munire V di due prodotti scalari, $f^*\phi$ e $g^*\psi$. Osserviamo che se $F \in GL(V)$ è una isometria tra $(V, f^*\phi)$ e $(V, g^*\psi)$, allora $G = g \circ F \circ f^{-1}$ è una isometria tra (W, ϕ) e (Z, ψ):

$$\begin{aligned}\psi(g(F(f^{-1}(\underline{w}_1))), g(F(f^{-1}(\underline{w}_2)))) &= g^*\psi(F(f^{-1}(\underline{w}_1)), F(f^{-1}(\underline{w}_2))) = \\ &= f^*\phi(f^{-1}(\underline{w}_1), f^{-1}(\underline{w}_2)) = \phi(\underline{w}_1, \underline{w}_2)\end{aligned}$$

per ogni $\underline{w}_1, \underline{w}_2 \in W$.
Allo stesso modo, se G è una isometria tra (W, ϕ) e (Z, ψ), allora $F = g^{-1} \circ G \circ f$ è una isometria tra $(V, f^*\phi)$ e $(V, g^*\psi)$.

Questo ci dice che per studiare la relazione di essere isometrici, possiamo ricondurci al caso di un solo spazio vettoriale V munito di due prodotti scalari $\phi, \psi \in PS(V)$. In questo caso, se esiste una isometria tra (V, ϕ) e (V, ψ) diremo semplicemente che ϕ e ψ sono isometrici.

La proposizione precedente in questo caso diventa:

Proposizione
Siano $\phi, \psi \in PS(V)$. I seguenti fatti sono equivalenti:

1. ϕ e ψ sono isometrici;
2. Per ogni base $\mathcal{B}$ di V, $M_{\mathcal{B}}(\phi)$ e $M_{\mathcal{B}}(\psi)$ sono congruenti;
3. Esiste un base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(\phi)$ e $M_{\mathcal{B}}(\psi)$ sono congruenti;
4. Esistono basi $\mathcal{B}, \mathcal{B}'$ di V tali che $M_{\mathcal{B}}(\phi) = M_{\mathcal{B}'}(\psi)$.

Se $W \subset V$ è un sottospazio, denotiamo con $\phi|_W$ la restrizione di ϕ a $W \times W$, $\phi|_W : W \times W \to \mathbb{K}$, $\phi|_W(\underline{w}_1, \underline{w}_2) = \phi(\underline{w}_1, \underline{w}_2)$ per ogni $\underline{w}_1, \underline{w}_2 \in W$. Diremo impropriamente che $\phi|_W$ è la restrizione di ϕ a W. È evidente che $\phi|_W$ è un prodotto scalare su W.
Ad esempio, la restrizione di ϕ a qualsiasi sottospazio contenuto in $Rad(\phi)$ dà il prodotto scalare nullo.
Notiamo che la mappa di restrizione $|_W : PS(V) \to PS(W)$ è lineare.

Dato (V, ϕ), consideriamo lo spazio quoziente $V/Rad(\phi)$, con proiezione al quoziente $\pi : V \to V/Rad(\phi)$.
Il prodotto scalare ϕ passa al quoziente, definendo $\bar{\phi}([\underline{v}_1], [\underline{v}_2]) = \phi(\underline{v}_1, \underline{v}_2)$ per ogni $[\underline{v}_1], [\underline{v}_2] \in V/Rad(\phi)$.
La bilinearità e la simmetria di $\bar{\phi}$ sono ovvie, controlliamo la buona definizione: se $\underline{w}_1, \underline{w}_2 \in Rad(\phi)$,

$$\begin{aligned}\phi(\underline{v}_1 + \underline{w}_1, \underline{v}_2 + \underline{w}_2) &= \phi(\underline{v}_1 + \underline{w}_1, \underline{v}_2) + \phi(\underline{v}_1 + \underline{w}_1, \underline{w}_2) = \\ &= \phi(\underline{v}_1 + \underline{w}_1, \underline{v}_2) = \\ &= \phi(\underline{v}_1, \underline{v}_2) + \phi(\underline{w}_1, \underline{v}_2) = \\ &= \phi(\underline{v}_1, \underline{v}_2).\end{aligned}$$

Il prodotto scalare $\bar{\phi}$ è non degenere, infatti, se $[\underline{v}] \in Rad(\bar{\phi})$, allora si ha $\phi(\underline{v}, \underline{w}) = \bar{\phi}([\underline{v}], [\underline{w}]) = 0$ per ogni $\underline{w} \in V$, per cui $\underline{v} \in Rad(\phi)$ e quindi $[\underline{v}] = [\underline{0}]$.

$(V/Rad(\phi), \bar{\phi})$ si dice *lo spazio non degenere (canonicamente) associato a* (V, ϕ).

Osserviamo che se $U \subset V$ è un sottospazio supplementare di $Rad(\phi)$, allora $\pi|_U : (U, \phi|_U) \to (V/Rad(\phi), \bar{\phi})$ è una isometria (infatti $\pi|_U$ è un isomorfismo). Ne segue che $\phi|_U$ è non degenere.
Inoltre, dati U_1, U_2 due supplementari di $Rad(\phi)$, $(U_1, \phi|_{U_1})$ e $(U_2, \phi|_{U_2})$ sono

canonicamente isometrici (tramite ${\pi|_{U_2}}^{-1} \circ \pi|_{U_1}$).
Questo lo possiamo vedere anche direttamente: $V = Rad(\phi) \oplus U_1 = Rad(\phi) \oplus U_2$, e dati $\underline{u}_1, \underline{u}_1' \in U_1$, scriviamo $\underline{u}_1 = \underline{w} + \underline{u}_2$, $\underline{u}_1' = \underline{w}' + \underline{u}_2'$ con $\underline{w}, \underline{w}' \in Rad(\phi)$, $\underline{u}_2, \underline{u}_2' \in U_2$. Allora $({\pi|_{U_2}}^{-1} \circ \pi|_{U_1})(\underline{u}_1) = \underline{u}_2$, $({\pi|_{U_2}}^{-1} \circ \pi|_{U_1})(\underline{u}_1') = \underline{u}_2'$ e si ha $\phi(\underline{u}_1, \underline{u}_1') = \phi(\underline{w} + \underline{u}_2, \underline{w}' + \underline{u}_2') = \phi(\underline{u}_2, \underline{u}_2')$.

Notiamo che se (V, ϕ) e (W, ψ) sono isometrici allora gli spazi non degeneri associati a (V, ϕ) e (W, ψ) sono isometrici.
Infatti, data una isometria $F : (V, \phi) \to (W, \psi)$, $F(Rad(\phi)) = Rad(\psi)$ e quindi F passa al quoziente definendo

$$\bar{F} : {V}/{Rad(\phi)} \to {W}/{Rad(\psi)}, \quad \bar{F}([\underline{v}]) = [F(\underline{v})]$$

che dà una isometria tra $({V}/{Rad(\phi)}, \bar{\phi})$ e $({W}/{Rad(\psi)}, \bar{\psi})$.
Viceversa, se $\dim V = \dim W$ e gli spazi non degeneri associati a (V, ϕ) e (W, ψ) sono isometrici, allora (V, ϕ) e (W, ψ) sono isometrici.
Infatti, fissiamo V_1 un supplementare di $Rad(\phi)$, e fissiamo W_1 un supplementare di $Rad(\psi)$. Data una isometria $\bar{F} : {V}/{Rad(\phi)} \to {W}/{Rad(\psi)}$, la composizione di isometrie $g = {\pi|_{W_1}}^{-1} \circ \bar{F} \circ \pi|_{V_1}$ dà una isometria tra $(V_1, \phi|_{V_1})$ e $(W_1, \psi|_{W_1})$. Poiché ${V}/{Rad(\phi)}$ è isomorfo a ${W}/{Rad(\psi)}$ e V è isomorfo a W, allora anche $Rad(\phi)$ e $Rad(\psi)$ sono isomorfi. Fissiamo un isomorfimo $h : Rad(\phi) \to Rad(\psi)$ e costruiamo l'isometria $F : V \to W$ richiedendo $F|_{V_1} = g$, $F|_{Rad(\phi)} = h$ (osserviamo che poiché $\phi|_{Rad(\phi)}$ e $\psi|_{Rad(\psi)}$ sono nulli, un qualsiasi isomorfismo tra $Rad(\phi)$ e $Rad(\psi)$ è un'isometria).
Infatti, dati $\underline{v}, \underline{v}' \in V$, scriviamo $\underline{v} = \underline{v}_1 + \underline{w}$, $\underline{v}' = \underline{v}_1' + \underline{w}'$, con $\underline{v}_1, \underline{v}_1' \in V_1$, $\underline{w}, \underline{w}' \in Rad(\phi)$, e allora

$$\begin{aligned}
\psi(F(\underline{v}), F(\underline{v}')) &= \psi(g(\underline{v}_1) + h(\underline{w}), g(\underline{v}_1') + h(\underline{w}')) = \\
&= \psi(g(\underline{v}_1), g(\underline{v}_1')) = \phi(\underline{v}_1, \underline{v}_1') = \\
&= \phi(\underline{v}_1 + \underline{w}, \underline{v}_1' + \underline{w}') = \\
&= \phi(\underline{v}, \underline{v}').
\end{aligned}$$

Quindi, per lo studio dei prodotti scalari a meno di isometrie possiamo ridurci al caso non degenere.
In particolare, la dimensione del radicale e la classe di isometria dello spazio non degenere associato (o la classe di isometria della restrizione ad un supplementare del radicale) sono invarianti completi per isometria.

Ortogonalità

Sia (V, ϕ) uno spazio vettoriale di dimensione finita, $\dim V = n > 0$, munito di prodotto scalare.

Due vettori $\underline{v}, \underline{w} \in V$ si dicono *ortogonali* (per il prodotto scalare ϕ), in simboli $\underline{v} \perp \underline{w}$, se $\phi(\underline{v}, \underline{w}) = 0$.
Ad esempio, un vettore $\underline{v} \in V$ è isotropo $\iff \underline{v} \perp \underline{v}$.
Due sottoinsiemi $S, T \subset V$ si dicono ortogonali, e si scrive $S \perp T$, se per ogni $\underline{s} \in S, \underline{t} \in T$, $\underline{s} \perp \underline{t}$.
L'insieme dei vettori di V ortogonali a $\underline{v} \in V$ si indica con $\underline{v}^\perp$. Osserviamo che $\underline{v}^\perp = \mathrm{Ker}(\phi(\underline{v}, \cdot)) = \mathrm{Ker}\, F_\phi(\underline{v})$ è un sottospazio.
Ad esempio, $\underline{v} \in Rad(\phi) \iff \underline{v}^\perp = V$.
Più in generale:

Definizione:
Sia $W \subset V$ un sottospazio.
L'*ortogonale* di W è

$$W^\perp = \{\underline{v} \in V |\ \phi(\underline{v}, \underline{w}) = 0\ \forall \underline{w} \in W\}$$

Se vogliamo evidenziare il ruolo di ϕ, scriviamo $W^{\perp_\phi}$.

Notiamo che $W^\perp = \bigcap_{\underline{w} \in W} \underline{w}^\perp = \bigcap_{\underline{w} \in W} \mathrm{Ker}(F_\phi(\underline{w})) = Z(F_\phi(W))$ è un sottospazio di V che contiene $Rad(\phi) = V^\perp$. Passando all'annullatore, $F_\phi(W) = Ann(W^\perp)$.

Valgono quindi proprietà simili a quelle che valevano per Z e Ann: se $U, W \subset V$ sono sottospazi,

• $W \subset U \Rightarrow W^\perp \supset U^\perp$,
infatti, $F_\phi(W) \subset F_\phi(U)$ e Z rovescia le inclusioni;

• $(W + U)^\perp = W^\perp \cap U^\perp$,
infatti, $F_\phi(W + U) = F_\phi(W) + F_\phi(U)$ e Z manda somme in intersezioni;

• $(W \cap U)^\perp \supset W^\perp + U^\perp$,
infatti, $F_\phi(W \cap U) \subset F_\phi(W) \cap F_\phi(U)$, quindi $Z(F_\phi(W \cap U)) \supset Z(F_\phi(W) \cap F_\phi(U))$ e Z mada intersezioni in somme;

• se ϕ è non degenere, $(W \cap U)^\perp = W^\perp + U^\perp$,
infatti, F_ϕ è un isomorfismo e allora $F_\phi(W \cap U) = F_\phi(W) \cap F_\phi(U)$;

• se $\underline{w}_1, \ldots, \underline{w}_m$ generano W, $W^\perp = \bigcap_{i=1}^m \underline{w}_i^\perp$,
infatti, $Z(F_\phi(W)) = Z(\{F_\phi(\underline{w}_1), \ldots, F_\phi(\underline{w}_m)\})$.

Tutte queste proprietà possono essere dimostrate direttamente dalla definizione di sottospazio ortogonale.

Ad esempio, per l'ultima: l'inclusione $W^\perp \subset \bigcap_{i=1}^m \underline{w}_i^\perp$ è ovvia; viceversa, se $\underline{v} \in V$

è ortogonale ad ogni $\underline{w}_i$, allora dato $\underline{w} \in W$, scrivendo $\underline{w} = \sum_{i=1}^{m} \alpha_i \underline{w}_i$, $\alpha_i \in \mathbb{K}$, abbiamo $\phi(\underline{v}, \underline{w}) = \sum_{i=1}^{m} \alpha_i \phi(\underline{v}, \underline{w}_i) = 0$.

La descrizione di $W^\perp$ come luogo di zeri ci fornisce un facile modo di calcolarne la dimensione:

$$\begin{aligned}\dim W^\perp &= \dim V - \dim F_\phi(W) = \dim V - \dim W + \dim(\mathrm{Ker}(F_\phi) \cap W) = \\ &= \dim V - \dim W + \dim(Rad(\phi) \cap W)\end{aligned}$$

È istruttivo ricavare questo risultato direttamente, senza usare Z. La dimostrazione, anche se alquanto più laboriosa, coinvolge idee e metodi comunque interessanti.

Se $W \subset V$ è un sottospazio, consideriamo la restrizione di ϕ a W. In generale, $\phi_{|W}$ può avere caratteristiche completamente diverse da quelle di ϕ. In particolare, può capitare che ϕ sia non degenere mentre $\phi_{|W}$ è degenere, o viceversa, che $\phi_{|W}$ sia non degenere anche se ϕ non lo è.
In effetti, direttamente della definizione

$$Rad(\phi_{|W}) = \{\underline{w} \in W |\ \phi(\underline{w}, \underline{u}) = 0\ \forall \underline{u} \in W\} = W \cap W^\perp$$

mostra che non c'è relazione tra $Rad(\phi_{|W})$ e $Rad(\phi)$. Osserviamo però che $Rad(\phi) \cap W \subset Rad(\phi_{|W})$.

Affrontiamo prima il caso particolare in cui ϕ è non degenere.
Vogliamo dimostrare che allora $\dim W^\perp = \dim V - \dim W$.
Poniamo $n = \dim V$, $m = \dim W$.
Fissiamo $\mathcal{D} = \{\underline{w}_1, \ldots, \underline{w}_m\}$ una base di W e osserviamo che $W^\perp = \bigcap_{i=1}^{m} \underline{w}_i^\perp$.
Completiamo tale base ad una base $\mathcal{B}$ di V e consideriamo $A = M_{\mathcal{B}}(\phi)$. Osserviamo che il minore di A dato dalle prime m righe e m colonne (dato dai prodotti scalari $\phi(\underline{w}_i, \underline{w}_j)$) è la matrice di $\phi_{|W}$ nella base $\mathcal{D}$.
Tramite l'isomorfismo delle coordinate nella base $\mathcal{B}$ (che è una isometria tra V munito di ϕ e $\mathbb{K}^n$ munito di Φ_A), W è isomorfo a $\mathrm{Span}(\underline{e}_1, \ldots, \underline{e}_m)$ e $W^\perp$ è isomorfo a $\bigcap_{i=1}^{m} \underline{e}_i^\perp$ (questi ortogonali rispetto a Φ_A).
Quindi $\underline{x} \in \mathbb{K}^n$ appartiene all'immagine di $W^\perp$ tramite l'isomorfismo delle coordinate nella base $\mathcal{B}$ se e solo se $\underline{e}_i^\top A \underline{x} = 0$ per $i = 1 \_ m$.
Ovvero, se A' è la sottomatrice di A data dalle prime m righe, $\dim W^\perp$ è uguale alla dimensione dello spazio delle soluzioni del sistema lineare $A'\underline{x} = \underline{0}$.
Ora, la matrice A è invertibile, quindi le righe di A sono linearmente indipendenti e lo sono anche quelle di A', che ha quindi rango m. Ne segue che $\dim W^\perp = n - m$ come voluto.

Notiamo che arriviamo alla stessa conclusione se supponiamo $\phi_{|W}$ non degenere.

Infatti, in questo caso, la matrice A' contiene il minore invertibile di ordine m $M_{\mathcal{D}}(\phi_{|_W})$, per cui anche in queste ipotesi A' ha rango m.
Notiamo che anche in questo caso, il risultato è coerente con la formula già trovata, essendo $Rad(\phi) \cap W = \{\underline{0}\}$.

Affrontiamo adesso il caso generale, cercando di ricondurlo al caso non degenere.

Ricordiamo che se U è un supplementare di $Rad(\phi)$, $V = U \oplus Rad(\phi)$, allora $\phi_{|_U}$ è non degenere.
Vediamo una dimostrazione alternativa. Dato $\underline{u} \in U$, $\underline{u} \neq \underline{0}$, allora $\underline{u}$ non appartiene al radicale di ϕ, per cui esiste $\underline{v}_0 \in V$ tale che $\phi(\underline{u}, \underline{v}_0) \neq 0$. Scrivendo $\underline{v}_0 = \underline{u}_0 + \underline{w}_0$ con $\underline{u}_0 \in U$, $\underline{w}_0 \in Rad(\phi)$, $0 \neq \phi(\underline{u}, \underline{v}_0) = \phi(\underline{u}, \underline{w}_0 + \underline{u}_0) = \phi(\underline{u}, \underline{w}_0) + \phi(\underline{u}, \underline{u}_0) = \phi(\underline{u}, \underline{u}_0)$. Quindi $\underline{u}$ non appartiene al radicale di $\phi_{|_U}$.

Scriviamo $W = Z \oplus (W \cap Rad(\phi))$, dove Z è un supplementare di $W \cap Rad(\phi)$ in W, e osserviamo che $W^\perp = Z^\perp$.
Infatti, $Z \subset W \Rightarrow W^\perp \subset Z^\perp$; inoltre, se $\underline{v} \in Z^\perp$, scriviamo $\underline{w} \in W$ come $\underline{w} = \underline{z} + \underline{u}$, con $\underline{z} \in Z$, $\underline{u} \in Rad(\phi)$, e allora $\phi(\underline{v}, \underline{w}) = \phi(\underline{v}, \underline{z} + \underline{u}) = \phi(\underline{v}, \underline{z}) = 0$.

Osserviamo che $Z \cap Rad(\phi) = \{\underline{0}\}$, per cui Z è contenuto in un supplementare U di $Rad(\phi)$. Consideriamo quindi $Z^{\perp_{\phi_{|U}}}$, l'ortogonale di Z tramite $\phi_{|_U}$, che è un sottospazio di U di dimensione $\dim U - \dim Z$, essendo $\phi_{|_U}$ non degenere.

Se mostriamo che $Z^\perp = Z^{\perp_\phi} = Z^{\perp_{\phi_{|U}}} \oplus Rad(\phi)$, allora
$$\begin{aligned}\dim W^\perp = \dim Z^\perp &= \dim U - \dim Z + \dim Rad(\phi) = \dim V - \dim Z = \\ &= \dim V - (\dim W - \dim W \cap Rad(\phi))\end{aligned}$$
come voluto.
Sia allora $\underline{v} \in Z^{\perp_\phi}$ e scriviamo $\underline{v} = \underline{u} + \underline{w}$ con $\underline{u} \in U$, $\underline{w} \in Rad(\phi)$, allora per ogni $\underline{z} \in Z$, $0 = \phi(\underline{v}, \underline{z}) = \phi(\underline{v}, \underline{u})$, mostra che $\underline{u} \in Z^{\perp_{\phi_{|U}}}$.
Viceversa, dato $\underline{v} + \underline{w}$ con $\underline{v} \in Z^{\perp_{\phi_{|U}}}$, $\underline{w} \in Rad(\phi)$, per ogni $\underline{z} \in Z$, $\phi(\underline{v} + \underline{w}, \underline{z}) = \phi(\underline{v}, \underline{z}) + \phi(\underline{w}, \underline{z}) = 0 + 0 = 0$, ovvero $\underline{v} + \underline{w} \in Z^{\perp_\phi}$.

Come corollario della formula della dimensione dell'ortogonale, determiniamo $(W^\perp)^\perp$ per $W \subset V$ sottospazio.

Notiamo che $W \subset (W^\perp)^\perp$ (gli elementi di W sono ortogonali a quelli di $W^\perp$ per definizione).
Per motivi dimensionali, se ϕ è non degenere vale l'uguaglianza, $W = (W^\perp)^\perp$, visto che, essendo $Rad(\phi) = \{\underline{0}\}$,

$$\dim(W^\perp)^\perp = \dim V - \dim W^\perp = \dim V - (\dim V - \dim W) = \dim W.$$

Osserviamo che, in quetso caso, $Ann(W) = Ann((W^\perp)^\perp) = F_\phi(W^\perp)$.

In generale si ha
$$(W^\perp)^\perp = W + Rad(\phi).$$

Infatti, $Rad(\phi) \subset (W^\perp)^\perp$ e quindi $W + Rad(\phi) \subset (W^\perp)^\perp$, visto che $(W^\perp)^\perp$ è un sottospazio.
Calcolando la dimensione, e ricordando che $Rad(\phi) \subset W^\perp$,

$$\begin{aligned}\dim(W^\perp)^\perp &= \dim V - \dim W^\perp + \dim(W^\perp \cap Rad(\phi)) = \\ &= \dim V - (\dim V - \dim W + \dim(W \cap Rad(\phi)) + \dim Rad(\phi) = \\ &= \dim W + \dim Rad(\phi) - \dim(W \cap Rad(\phi)) = \\ &= \dim(W + Rad(\phi)).\end{aligned}$$

In generale, quindi,

$$\begin{aligned}F_\phi(W^\perp) = Ann((W^\perp)^\perp) &= Ann(W + Rad(\phi)) = \\ &= Ann(W) \cap Ann(Rad(\phi)) = \\ &= Ann(W) \cap \mathrm{Im}(F_\phi).\end{aligned}$$

Notiamo che se $f : (V, \phi) \to (W, \psi)$ è una isometria, per ogni $U \subset V$ sottospazio, $f(U^{\perp_\phi}) = (f(U))^{\perp_\psi}$.
Infatti, $\phi(\underline{v}, \underline{w}) = 0$ per ogni $\underline{w} \in U \iff \psi(f(\underline{v}), f(\underline{w})) = 0$ per ogni $\underline{w} \in U$, e posto $\underline{z} = f(\underline{w}), \iff \psi(f(\underline{v}), \underline{z}) = 0$ per ogni $\underline{z} \in f(U)$.

Inoltre, se $V = U_1 \oplus U_2$ con U_1, U_2 sottospazi ortogonali, $U_1 \perp U_2$, e $W = Z_1 \oplus Z_2$ con Z_1, Z_2 sottospazi ortogonali, $Z_1 \perp Z_2$, se $f_i : U_i \to Z_i$ è una isometria tra $(U_i, \phi_{|_{U_i}})$ e $(Z_i, \psi_{|_{Z_i}}), i = 1, 2$, allora possiamo costruire una isometria f tra (V, ϕ) e (W, ψ) richiedendo che $f_{|_{U_i}} = f_i$, $i = 1, 2$. Infatti, dati $\underline{v}, \underline{v}' \in V$, scriviamo $\underline{v} = \underline{u}_1 + \underline{u}_2$, $\underline{v}' = \underline{u}'_1 + \underline{u}'_2$ con $\underline{u}_i, \underline{u}'_i \in U_i$, e quindi

$$\begin{aligned}\psi(f(\underline{v}), f(\underline{v}')) &= \psi(f_1(\underline{u}_1) + f_2(\underline{u}_2), f_1(\underline{u}'_1) + f_2(\underline{u}'_2)) = \\ &= \psi(f_1(\underline{u}_1), f_1(\underline{u}'_1)) + \psi(f_1(\underline{u}_1), f_2(\underline{u}'_2)) + \\ &\quad + \psi(f_2(\underline{u}_2), f_1(\underline{u}'_1)) + \psi(f_2(\underline{u}_2), f_2(\underline{u}'_2)) = \\ &= \phi(\underline{u}_1, \underline{u}'_1) + \phi(\underline{u}_2, \underline{u}'_2),\end{aligned}$$

poiché $f_1(\underline{u}_1), f_1(\underline{u}'_1) \in Z_1 \perp Z_2 \ni f_2(\underline{u}_2), f_2(\underline{u}'_2)$.
Allo stesso modo,

$$\phi(\underline{v}, \underline{v}') = \phi(\underline{u}_1, \underline{u}'_1) + \phi(\underline{u}_2, \underline{u}'_2),$$

poiché $\underline{u}_1, \underline{u}'_1 \in U_1 \perp U_2 \ni \underline{u}_2, \underline{u}'_2$.

Lo stesso vale per decomposizioni di V e W in somma diretta di più di due sottospazi a due a due ortogonali.

Se $W_1, \dots, W_k \subset V$ sono sottospazi in somma diretta ed inoltre sono a due a due ortogonali, $W_i \perp W_j$ (cioè $W_i \subset W_j^\perp$) se $i \neq j$, usiamo la notazione $W_1 \oplus^\perp \cdots \oplus^\perp W_k$ per indicare la somma diretta, che si dice *somma diretta ortogonale*.
Notiamo che in una base adattata alla somma diretta, la matrice della restrizione di ϕ ad una somma diretta ortogonale è diagonale a blocchi, e i blocchi sulla diagonale sono matrici associate alle restrizioni di ϕ ai singoli addendi. Quindi

$\operatorname{rnk}(\phi|_{W_1\oplus^\perp\cdots\oplus^\perp W_k}) = \sum_{i=1}^{k} \operatorname{rnk}(\phi|_{W_i})$. In particolare $\phi|_{W_1\oplus^\perp\cdots\oplus^\perp W_k}$ è non degenere se e solo se $\phi|_{W_i}$ è non degenere per ogni $i = 1\_k$.

Ad esempio, se $W \subset V$ è un sottospazio, $W \perp W^\perp$ per definizione, ma non sempre i due sottospazi sono in somma diretta.

Proposizione
Dato $W \subset V$ sottospazio, allora W e $W^\perp$ sono in somma diretta se e solo se $\phi|_W$ è non degenere. In tal caso, $V = W \oplus^\perp W^\perp$

Dimostrazione
La prima affermazione è chiara in quanto $Rad(\phi|_W) = W \cap W^\perp$.
Resta da vedere che in tal caso, $V = W \oplus W^\perp$.
Ma $\dim W^\perp = \dim V - \dim W + \dim(Rad(\phi) \cap W) = \dim V - \dim W$, per cui $\dim(W \oplus W^\perp) = \dim W + \dim W^\perp = \dim V$, come voluto. ■

Nel caso particolare in cui $\dim W = 1$, se $\underline{v}_0$ genera W, allora $\phi|_W$ è non degenere se e solo se $\underline{v}_0$ non è isotropo.
Infatti, dato $a \in \mathbb{K}$, $\phi(a\underline{v}_0, a\underline{v}_0) = a^2\phi(\underline{v}_0, \underline{v}_0)$, per cui ci sono due possibilità: $\underline{v}_0 \in CI(\phi)$ e quindi $\phi|_W = 0$; oppure $\underline{v}_0 \notin CI(\phi)$ e $\phi|_W$ è non degenere (ovvero, l'unica matrice di taglia 1×1 non invertibile è la matrice nulla).
Nel caso $\underline{v}_0 \notin CI(\phi)$, allora $V = \operatorname{Span}(\underline{v}_0) \oplus \underline{v}_0^\perp$. Possiamo allora esplicitamente trovare la decomposizione di ogni vettore in termini di tale somma diretta.
Dato $\underline{v} \in V$, scriviamo $\underline{v} = \alpha\underline{v}_0 + \underline{w}$ con $\alpha \in \mathbb{K}$ e $\underline{w} \in \underline{v}_0^\perp$. Facendo il prodotto scalare per $\underline{v}_0$ otteniamo $\phi(\underline{v}, \underline{v}_0) = \alpha\phi(\underline{v}_0, \underline{v}_0)$, da cui $\alpha = \dfrac{\phi(\underline{v}, \underline{v}_0)}{\phi(\underline{v}_0, \underline{v}_0)}$, e dunque $\underline{w} = \underline{v} - \alpha\underline{v}_0$.
Verifichiamo che $\underline{w} \in \underline{v}_0^\perp$:

$$\begin{aligned}\phi(\underline{w}, \underline{v}_0) &= \phi(\underline{v}, \underline{v}_0) - \alpha\phi(\underline{v}_0, \underline{v}_0) = \\ &= \phi(\underline{v}, \underline{v}_0) - \frac{\phi(\underline{v}, \underline{v}_0)}{\phi(\underline{v}_0, \underline{v}_0)}\phi(\underline{v}_0, \underline{v}_0) = \\ &= 0.\end{aligned}$$

Il coefficiente $c(\underline{v}, \underline{v}_0) = \dfrac{\phi(\underline{v}, \underline{v}_0)}{\phi(\underline{v}_0, \underline{v}_0)}$ si dice *coefficiente di Fourier* di $\underline{v}$ rispetto a $\underline{v}_0$.
Abbiamo quindi una espressione esplicita per le due proiezioni date dalla somma diretta ortogonale $V = \operatorname{Span}(\underline{v}_0) \oplus \underline{v}_0^\perp$, dette *proiezioni ortogonali*. Esse sono date da:

$$p_{\operatorname{Span}(\underline{v}_0)}(\underline{v}) = c(\underline{v}, \underline{v}_0)\underline{v}_0, \ \forall \underline{v} \in V,$$
$$p_{\underline{v}_0^\perp}(\underline{v}) = \underline{v} - c(\underline{v}, \underline{v}_0)\underline{v}_0, \ \forall \underline{v} \in V.$$

Notiamo che $\underline{v}_0 \in CI(\phi) \iff \underline{v}_0 \in \underline{v}_0^\perp$, per cui in questo caso non abbiamo la decomposizione di sopra.

Dato (V, ϕ), la *forma quadratica* associata a ϕ è la mappa

$$q_\phi : V \to \mathbb{K}, \quad \underline{v} \mapsto \phi(\underline{v}, \underline{v}).$$

Osserviamo che $q_\phi^{-1}\{0\} = CI(\phi)$.
Per ogni $\underline{v}, \underline{w} \in V$, $\phi(\underline{v}+\underline{w}, \underline{v}+\underline{w}) = \phi(\underline{v}, \underline{v}) + \phi(\underline{v}, \underline{w}) + \phi(\underline{w}, \underline{v}) + \phi(\underline{w}, \underline{w})$.
Otteniamo la *formula di polarizzazione*:

$$2\phi(\underline{v}, \underline{w}) = q_\phi(\underline{v}+\underline{w}) - q_\phi(\underline{v}) - q_\phi(\underline{w}).$$

Notiamo che se $\mathbb{K}$ ha caratteristica diversa da 2, la forma quadratica determina completamente il prodotto scalare. In particolare, $\phi = 0$ se e solo se $q_\phi = 0$.

L'ipotesi sulla caratteristica è essenziale, come mostra il seguente esempio:
$\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$, $V = \mathbb{K}^2$, $\phi = \Phi_M$ con $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
Allora $q_\phi(\binom{x}{y}) = (x\ y)M\binom{x}{y} = xy + xy = 0$. La forma quadratica è quindi nulla, mentre il prodotto scalare non lo è (anzi, è non degenere!).

Da ora in avanti, considereremo solo campi di caratterisitica diversa da 2 per poter applicare quanto sopra.

Definizione
Una base di V, $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$, si dice *ortogonale* per ϕ se $\underline{v}_i \perp \underline{v}_j$ per ogni $i, j = 1\_n$, $i \neq j$. Abbrevieremo dicendo che $\mathcal{B}$ è una base ortogonale di (V, ϕ) (o solo base ortogonale di V, sottintendendo ϕ).
Si ha quindi $\phi(\underline{v}_i, \underline{v}_j) = 0$ per ogni $i, j = 1\_n$, $i \neq j$, in altre parole, la matrice $M_\mathcal{B}(\phi)$ di ϕ nella base $\mathcal{B}$ è diagonale.

Ad esempio, la base canonica è una base ortogonale per il prodotto scalare standard su $\mathbb{K}^n$.

Osserviamo che ϕ è non degenere se e solo se una base ortogonale $\mathcal{B}$ di (V, ϕ) non contiene vettori isotropi: in entrambi i casi, la matrice diagonale $M_\mathcal{B}(\phi)$ è invertibile.
Di contro, se ϕ è degenere, ogni base ortogonale $\mathcal{B}$ di (V, ϕ) deve necessariamente contenere una base di $Rad(\phi)$.
Infatti, poiché il rango della matrice diagonale $M_\mathcal{B}(\phi)$ vale $n - \dim Rad(\phi)$, $\mathcal{B}$ contiene esattamente $\dim Rad(\phi)$ vettori isotropi (linearmente indipendenti). Tali vettori, essendo ortogonali a tutti i vettori di una base di V (ognuno di essi è ortogonale a sé stesso poiché isotropo e agli altri poiché la base è ortogonale), appartengono al radicale di ϕ, e quindi ne sono una base.

Se $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ è una base ortogonale di (V, ϕ), dato $\underline{v} \in V$ scriviamo $\underline{v} = \sum_{i=1}^{n} \alpha_i \underline{v}_i$ con $\alpha_i \in K$. Allora, visto che $\underline{v}_i \perp \underline{v}_j$ se $i \neq j$, $\phi(\underline{v}, \underline{v}_i) = \alpha_i \phi(\underline{v}_i, \underline{v}_i)$, per cui, essendo $\underline{v}_i$ nono isotropo, $\alpha_i = c(\underline{v}, \underline{v}_i)$. Dunque, $[\underline{v}]_\mathcal{B} = \begin{pmatrix} c(\underline{v}, \underline{v}_1) \\ \vdots \\ c(\underline{v}, \underline{v}_n) \end{pmatrix}$.

Teorema (Esistenza di basi ortogonali)
Ogni (V, ϕ) ammette una base ortogonale.

Dimostrazione
Per induzione su $n = \dim V$
Se $n = 1$, tutte le basi di V sono ortogonali.
Supponiamo quindi $n > 1$.
Se la forma quadratica $q_\phi = 0$, allora $\phi = 0$ e tutte le basi di V sono ortogonali. Altrimenti, esiste $\underline{v} \in V$ tale che $q_\phi(\underline{v}) \neq 0$. Tale $\underline{v}$ è quindi non isotropo e abbiamo $V = \mathrm{Span}(\underline{v}) \oplus \underline{v}^\perp$. Consideriamo $\underline{v}^\perp$, che ha dimensione $n-1$, munito del prodotto scalare $\phi_{|_{\underline{v}^\perp}}$. Per ipotesi induttiva, esiste una base ortogonale di $(\underline{v}^\perp, \phi_{|_{\underline{v}^\perp}})$, $\underline{v}_2, \ldots, \underline{v}_n$. Allora la base di V $\underline{v}_1 = \underline{v}, \underline{v}_2, \ldots, \underline{v}_n$ è una base ortogonale di (V, ϕ). ■

Notiamo di nuovo che l'ipotesi restrittiva che $\mathbb{K}$ abbia caratteristica diversa da 2 è essenziale, senza questa ipotesi possono non esistere basi ortogonali. Riprendendo l'esempio precedente, Φ_M è non degenere ma $CI(\Phi_M) = \mathbb{K}^2$, per cui non possono esistere basi ortogonali.

Notiamo che il teorema di esistenza delle basi ortogonali, applicato a $(\mathbb{K}^n, \Phi_A)$ con $A \in S(n, \mathbb{K})$, dà immediatamente il seguente risultato:

Corollario
Ogni matrice simmetrica (a coefficienti in un campo di caratteristica diversa da 2) è congruente ad una matrice diagonale:
$\forall\ A \in S(n, \mathbb{K})$ esiste $P \in GL(n, \mathbb{K})$ tale che $P^\top AP = D$ diagonale.

■

Osserviamo che se $W_1, \ldots, W_k \subset V$ sono sottospazi in somma diretta ortogonale, e $\mathcal{B}_i$ è una base ortogonale di W_i, $i = 1\_k$, allora $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ è una base ortogonale di $W_1 \oplus^\perp \cdots \oplus^\perp W_k$.
Otteniamo una dimostrazione alternativa del fatto che la restrizione di ϕ alla somma diretta è non degenere se e solo se le restrizioni di ϕ ai singoli addendi lo sono: infatti, la base $\mathcal{B}$ non continene vettori isotropi se e solo se lo stesso vale per le basi $\mathcal{B}_i$.

Dato $W \subset V$ un sottospazio tale che $\phi_{|_W}$ è non degenere, e quindi abbiamo la decomposizione in somma diretta ortogonale $V = W \oplus^\perp W^\perp$, siano p_W e $p_{W^\perp}$ le due proiezioni date dalla somma diretta, che si dicono *proiezioni ortogonali*.
Se $\underline{w}_1, \ldots, \underline{w}_k$ è una base ortogonale di W, allora possiamo dare una formulazione esplicita per le due proiezioni ortogonali.
Dato $\underline{v} \in V$, $\underline{z} = \underline{v} - \sum_{i=1}^{k} c(\underline{v}, \underline{w}_i)\underline{w}_i \in W^\perp$. Infatti, per ogni $i = 1\_k$,

$\phi(\underline{z}, \underline{w}_i) = \phi(\underline{v}, \underline{w}_i) - c(\underline{v}, \underline{w}_i)\phi(\underline{w}_i, \underline{w}_i) = 0$. Allora, per ogni $\underline{v} \in V$,

$$p_W(\underline{v}) = \sum_{i=1}^{k} c(\underline{v}, \underline{w}_i)\underline{w}_i,$$

$$p_{W^\perp}(\underline{v}) = \underline{v} - \sum_{i=1}^{k} c(\underline{v}, \underline{w}_i)\underline{w}_i.$$

Raffiniamo la dimostrazione del teorema di esistenza di basi ortogonali e formuliamo un algoritmo che permetta di costruire una base ortogonale $\mathcal{B}'$ a partire da una base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$.
Se $\phi = 0$, poniamo $\mathcal{B}' = \mathcal{B}$.
Altrimenti, individuiamo un vettore non isotropo $\underline{v}$.
Se esiste $i = 1\_n$ tale che $\underline{v}_i$ è non isotropo, poniamo $\underline{v} = \underline{v}_i$.
Altrimenti, se tutti i vettori della base $\mathcal{B}$ sono isotropi, scegliamo $\underline{v}_i, \underline{v}_j$ tali che $\phi(\underline{v}_i, \underline{v}_j) \neq 0$. Tali i, j esistono in quanto, poiché $\phi \neq 0$, $M_{\mathcal{B}}(\phi) \neq 0$, e basta scgliere i, j tali che $\overline{M_{\mathcal{B}}(\phi)\big|_i^j} \neq 0$.
Allora $\underline{v} = \underline{v}_i + \underline{v}_j$ non è isotropo:

$$q_\phi(\underline{v}_i + \underline{v}_j) = q_\phi(\underline{v}_i) + q_\phi(\underline{v}_j) + 2\phi(\underline{v}_i, \underline{v}_j) = 2\phi(\underline{v}_i, \underline{v}_j) \neq 0.$$

In entrambi i casi, a meno di sostituire un $\underline{v}_i$ con $\underline{v}_i + \underline{v}_j$ e/o di riordinare i vettori della base, possiamo supporre che $\underline{v}_1$ sia non isotropo.
Se $\dim V = 1$, poniamo $\mathcal{B}' = \{\underline{v}_1\}$.
Altrimenti, abbiamo $V = \mathrm{Span}(\underline{v}_1) \oplus \underline{v}_1^\perp$ e possiamo iterare il procedimento su $\underline{v}_1^\perp$ se troviamo esplicitamente una base di $\underline{v}_1^\perp$.
Per questo, proiettiamo $\underline{v}_2, \dots, \underline{v}_n$ su $\underline{v}_1^\perp$ usando la proiezione ortogonale data dalla somma diretta. Otteniamo $\underline{w}_2^{(1)}, \dots, \underline{w}_n^{(1)} \in \underline{v}_1^\perp$, $\underline{w}_i^{(1)} = \underline{v}_i - c(\underline{v}_i, \underline{v}_1)\underline{v}_1$, che sono linearmente indipendenti (e quindi una base di $\underline{v}_1^\perp$); infatti, per $\alpha_i \in \mathbb{K}$,

$$\underline{0} = \sum_{i=2}^{n} \alpha_i \underline{w}_i^{(1)} = \sum_{i=2}^{n} \alpha_i \underline{v}_i + (\sum_{i=2}^{n} \alpha_i c(\underline{v}_i, \underline{v}_1))\underline{v}_1 \quad \Rightarrow \quad \alpha_i = 0 \ \forall i = 2\_n.$$

Osserviamo che $\mathcal{B}^{(1)} = \{\underline{v}_1' = \underline{v}_1, \underline{w}_2^{(1)}, \dots, \underline{w}_n^{(1)}\}$ è una base di V in cui il primo vettore è ortogonale agli altri.
Si itera quindi l'algoritmo sostituendo (V, ϕ) con $(v_1^\perp, \phi_{|_{\underline{v}_1^\perp}})$, e la base $\mathcal{B}$ con la base $\underline{w}_2^{(1)}, \dots, \underline{w}_n^{(1)}$: se $\phi_{|_{\underline{v}_1^\perp}} = 0$, $\mathcal{B}' = \mathcal{B}^{(1)}$, altrimenti si produce una base $\mathcal{B}^{(2)} = \{\underline{v}_1', \underline{v}_2', \underline{w}_3^{(2)}, \dots, \underline{w}_n^{(2)}\}$ di V in cui il primi due vettori sono ortogonali tra di loro e agli altri, e $\underline{w}_3^{(2)}, \dots, \underline{w}_n^{(2)}$ è una base di $\mathrm{Span}(\underline{v}_1', \underline{v}_2')^\perp$.
Dopo al più $n - 1$ iterazioni otteniamo una base ortogonale $\mathcal{B}'$ di (V, ϕ).

Osserviamo che se ϕ è anisotropo, l'algoritmo si semplifica, perché il primo vettore della base non è isotropo (e si evita la fase iniziale di determinazione di un

vettore isotropo). In tal caso, l'algoritmo prende il nome di *ortogonalizzazione di Gram-Schmidt.*
In questo caso, al primo passo otteniamo la base $\mathcal{B}^{(1)}$

$$\begin{aligned}
&\underline{v}_1,\\
&\underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1,\\
&\vdots\\
&\underline{v}_n - c(\underline{v}_n, \underline{v}_1)\underline{v}_1
\end{aligned}$$

(in cui il primo vettore è ortogonale a tutti gli altri).
Si nota che le basi $\mathcal{B}$ e $\mathcal{B}^{(1)}$ danno la stessa bandiera. Ovvero, la matrice di cambio di base è triangolare superiore con 1 sulla diagonale.
Questo rimane vero anche per i passi successivi (che sono formalmente uguali al primo).
Ad esempio, al secondo passo otteniamo la base $\mathcal{B}^{(2)}$

$$\begin{aligned}
&\underline{v}_1,\\
&\underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1,\\
&\underline{v}_3 - c(\underline{v}_3, \underline{v}_1)\underline{v}_1 - c(\underline{v}_3 - c(\underline{v}_3, \underline{v}_1)\underline{v}_1, \underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1)(\underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1),\\
&\vdots\\
&\underline{v}_n - c(\underline{v}_n, \underline{v}_1)\underline{v}_1 - c(\underline{v}_n - c(\underline{v}_n, \underline{v}_1)\underline{v}_1, \underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1)(\underline{v}_2 - c(\underline{v}_2, \underline{v}_1)\underline{v}_1)
\end{aligned}$$

(in cui il primo e il secondo vettore sono ortogonali tra di loro e a tutti gli altri), da cui si vede che la bandiera data da $\mathcal{B}^{(2)}$ è la stessa di quella data sia da $\mathcal{B}^{(1)}$ che da $\mathcal{B}$, e che le matrici di cambio di base sono triangolari superiori con 1 sulla diagonale.
Quindi, tutte le basi $\mathcal{B}^{(i)}$ danno la stessa bandiera della base iniziale $\mathcal{B}$, e le varie matrici di cambio di base sono triangolari superiori con 1 sulla diagonale.
Otteniamo che $\mathcal{B}$ e $\mathcal{B}'$ danno la stessa bandiera e che la matrice di cambio di base da $\mathcal{B}$ a $\mathcal{B}'$ è triangolare superiore con 1 sulla diagonale.

Abbiamo immediatamente il seguente

Corollario
Sia (V, ϕ) con ϕ anisotropo e sia $f \in \mathrm{End}(V)$ triangolabile.
Allora esiste una base ortogonale di (V, ϕ) che triangola f.

Dimostrazione
Fissata una base $\mathcal{B}$ che triangola f, ovvero la cui bandiera è f-invariante, produciamo tramite l'algoritmo la base ortogonale $\mathcal{B}'$, che avendo la stessa bandiera di $\mathcal{B}$, ha ancora bandiera f-invariante e quindi triangola f. ∎

Il gruppo ortogonale

Dato (V, ϕ), definiamo il *gruppo ortogonale* di ϕ

$$O(\phi) = \{f \in GL(V) |\ \phi(\underline{v}, \underline{w}) = \phi(f(\underline{v}), f(\underline{w}))\ \forall \underline{v}, \underline{w} \in V\}.$$

Notiamo che si tratta delle isometrie da (V, ϕ) a (V, ϕ), per cui è immediato che si tratti di un sottogruppo di $GL(V)$ (che viene detto anche gruppo delle isometrie di ϕ).
Nel caso $V = \mathbb{K}^n$, $\phi = \Phi_M$ con $M \in S(n, \mathbb{K})$,

$$O(\Phi_M) = \{P \in GL(n, \mathbb{K}) |\ P^\top M P = M\}.$$

In particolare, per $M = I_n$ si ottiene il *gruppo ortogonale matriciale classico*

$$O(n, \mathbb{K}) = \{P \in GL(n, \mathbb{K}) |\ P^{-1} = P^\top\}.$$

Se (V, ϕ), (W, ψ) sono isometrici, allora $O(\phi)$ e $O(\psi)$ sono isomorfi. Infatti, se $F : V \to W$ è una isometria, allora la coniugazione con F dà un isomorfismo tra i gruppi ortogonali, $F \circ * \circ F^{-1} : O(\phi) \to O(\psi)$. In particolare, $O(\phi)$ è isomorfo a $O(\Phi_M)$, dove $M = M_{\mathcal{B}}(\phi)$, per ogni base $\mathcal{B}$ di V.

Sia (V, ϕ) uno spazio vettoriale sul campo $\mathbb{K}$ di caratteristica diversa da 2, con $\dim V = n$, munito del prodotto scalare ϕ.

Se $\underline{v} \in V$ è non isotropo, usando la decomposizione $V = \mathrm{Span}(\underline{v}) \oplus \underline{v}^\perp$, definiamo $\rho_{\underline{v}} : V \to V$ lineare ponendo $\rho_{\underline{v}}(\underline{v}) = -\underline{v}$, $\rho_{\underline{v}}|_{\underline{v}^\perp} = id_{\underline{v}^\perp}$.
$\rho_{\underline{v}}$ si dice *riflessione ortogonale parallela a* $\underline{v}$ o *rispetto a* $\underline{v}^\perp$.

Notiamo che $\rho_{\underline{v}} = \rho_{\lambda \underline{v}}$ per ogni $\lambda \in \mathbb{K}$, $\lambda \neq 0$, e che $\rho_{\underline{v}}^2 = id_V$. Quindi $\rho_{\underline{v}}$ è diagonalizzabile con spettro $\{\pm 1\}$ e autospazi $V_{-1}(\rho_{\underline{v}}) = \mathrm{Span}(\underline{v})$, $V_1(\rho_{\underline{v}}) = \underline{v}^\perp$. Il luogo dei punti fissi di $\rho_{\underline{v}}$ è dunque un iperpiano. Notiamo anche che le riflessioni ortogonali hanno determinante pari a -1.

Esplicitamente, dato $\underline{w} \in V$, scriviamo $\underline{w} = c(\underline{w}, \underline{v})\underline{v} + \underline{z}$, dove sappiamo che $\underline{z} \perp \underline{v}$, e allora si ha $\rho_{\underline{v}}(\underline{w}) = -c(\underline{w}, \underline{v})\underline{v} + \underline{z} = \underline{w} - 2c(\underline{w}, \underline{v})\underline{v}$.

Mostriamo che $\rho_{\underline{v}} \in O(\phi)$:
dati $\underline{v}_1, \underline{v}_2 \in V$, scriviamo $\underline{v}_i = c(\underline{v}_i, \underline{v})\underline{v} + \underline{z}_i$, $i = 1, 2$, $(\underline{z}_i \perp \underline{v})$. Allora si ha

$$\begin{aligned}\phi(\rho_{\underline{v}}(\underline{v}_1), \rho_{\underline{v}}(\underline{v}_2)) &= \phi(-c(\underline{v}_1, \underline{v})\underline{v} + \underline{z}_1, -c(\underline{v}_2, \underline{v})\underline{v} + \underline{z}_2) = \\ &= c(\underline{v}_1, \underline{v})c(\underline{v}_2, \underline{v})\phi(\underline{v}, \underline{v}) + \phi(\underline{z}_1, \underline{z}_2) = \\ &= \phi(c(\underline{v}_1, \underline{v})\underline{v} + \underline{z}_1, c(\underline{v}_2, \underline{v})\underline{v} + \underline{z}_2) = \\ &= \phi(\underline{v}_1, \underline{v}_2);\end{aligned}$$

inoltre, $\phi(\rho_{\underline{v}}(\underline{v}_1), \underline{v}_2) = \phi(\rho_{\underline{v}}^2(\underline{v}_1), \rho_{\underline{v}}(\underline{v}_2)) = \phi(\underline{v}_1, \rho_{\underline{v}}(\underline{v}_2))$ (si dice che $\rho_{\underline{v}}$ è *autoaggiunta*).

Se $\underline{v}_1, \underline{v}_2 \in V$ sono non isotropi e ortogonali, allora $\rho_{\underline{v}_1} \circ \rho_{\underline{v}_2} = \rho_{\underline{v}_2} \circ \rho_{\underline{v}_1}$.

Infatti, posto $W = \mathrm{Span}(\underline{v}_1, \underline{v}_2)$, $\phi|_W$ è non degenere e abbiamo la decomposizione $V = \mathrm{Span}(\underline{v}_1) \oplus^\perp \mathrm{Span}(\underline{v}_2) \oplus^\perp W^\perp$.
Scrivendo $\underline{w} \in V$ come $\underline{w} = \alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \underline{z}$, con $\alpha_1, \alpha_2 \in \mathbb{K}$, $\underline{z} \in W^\perp$,

$$\begin{aligned}\rho_{\underline{v}_1}(\rho_{\underline{v}_2}(\underline{w})) &= \rho_{\underline{v}_1}(\alpha_1\underline{v}_1 - \alpha_2\underline{v}_2 + \underline{z}) = \\ &= -\alpha_1\underline{v}_1 - \alpha_2\underline{v}_2 + \underline{z} = \\ &= \rho_{\underline{v}_2}(-\alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \underline{z}) = \\ &= \rho_{\underline{v}_2}(\rho_{\underline{v}_1}(\underline{w})).\end{aligned}$$

Notiamo che $\rho_{\underline{v}_1} \circ \rho_{\underline{v}_2}|_W = -id_W$, $\rho_{\underline{v}_1} \circ \rho_{\underline{v}_2}|_{W^\perp} = id_{W^\perp}$.
Questo si generalizza immediatamente ad un numero finito di vettori non isotropi $\underline{v}_1, \ldots, \underline{v}_k \in V$ a due a due ortogonali : le riflessioni ortogonali $\rho_{\underline{v}_1}, \ldots, \rho_{\underline{v}_k}$ commutano tra di loro e la loro composizione se ristretta a $W = \mathrm{Span}(\underline{v}_1, \ldots, \underline{v}_k)$ vale $-id_W$, mentre se ristretta a $W^\perp$ vale $id_{W^\perp}$

Poiché per ogni riflessione ortogonale ρ, $id_V = \rho^2$, si ha che id_V è composizione di un numero finito di riflessioni ortogonali.
Se ϕ è non degenere, data $\underline{v}_1, \ldots, \underline{v}_n$ una base ortogonale di V (interamente composta da vettori non isotropi), $\rho_{\underline{v}_1} \circ \cdots \circ \rho_{\underline{v}_n} = -id_V$, per cui anche $-id_V$ è composizione di un numero finito di riflessioni ortogonali.
Osserviamo che $-id_V$ non si può scrivere come composizione di m riflessioni ortogonali con $m < n$, in quanto l'intersezione degli iperpiani dei punti fissi di tali riflessioni ortogonali, che è contenuto nel luogo dei punti fissi della loro composizione, avrebbe dimensione almeno $n - m > 0$ mentre il luogo dei punti fissi di $-id_V$ è dato dal solo $\underline{0}$.
Vogliamo vedere che, nel caso ϕ sia non degenere, id_V e $-id_V$ non sono casi particolari.

L'osservazione chiave è la seguente.
Se $\underline{v}, \underline{w} \in V$ sono tali che $q_\phi(\underline{v}) = q_\phi(\underline{w})$ e $\underline{z} = \underline{v} - \underline{w}$ non è isotropo, allora $\rho_{\underline{z}}(\underline{v}) = \underline{w}$ (e applicando $\rho_{\underline{z}}$, $\rho_{\underline{z}}(\underline{w}) = \underline{v}$).
Infatti, scrivendo $\underline{v} = \frac{1}{2}(\underline{v} + \underline{w}) + \frac{1}{2}(\underline{v} - \underline{w})$ e osservando che $(\underline{v} + \underline{w}) \perp (\underline{v} - \underline{w})$ in quanto $\phi(\underline{v} + \underline{w}, \underline{v} - \underline{w}) = q_\phi(\underline{v}) - q_\phi(\underline{w}) + \phi(\underline{v}, \underline{w}) - \phi(\underline{v}, \underline{w}) = 0$, si ha che $\rho_{\underline{z}}(\underline{v}) = \frac{1}{2}(\underline{v} + \underline{w}) - \frac{1}{2}(\underline{v} - \underline{w}) = \underline{w}$.
Analogamente, se $\underline{u} = \underline{v} + \underline{w}$ non è isotropo, allora $\rho_{\underline{u}}(\underline{v}) = -\underline{w}$ (e applicando $\rho_{\underline{u}}$, $\rho_{\underline{u}}(\underline{w}) = -\underline{v}$).

Teorema
Se ϕ è non degenere, $O(\phi)$ è generato dalle riflessioni ortogonali, ovvero ogni $f \in O(\phi)$ è composizione di un numero finito di riflessioni ortogonali.

Dimostrazione
Ragioniamo per induzione su $n = \dim V$.
Per $n = 1$, $O(\phi) = \{id_V, -id_V\}$ e abbiamo finito. Infatti, data $f \in O(\phi)$, esiste $\lambda \in \mathbb{K}$ tale che $f = \lambda id_V$. Ma allora, scelto $\underline{v} \in V$ non nullo $0 \neq \phi(\underline{v}, \underline{v}) = \phi(f(\underline{v}), f(\underline{v})) = \lambda^2\phi(\underline{v}, \underline{v})$ mostra che $\lambda = \pm 1$.

Possiamo quindi supporre $n > 1$.
Supponiamo esista $\underline{v}_0 \in V$ non isotropo tale che $f(\underline{v}_0) = \underline{v}_0$ e scriviamo $V = \mathrm{Span}(\underline{v}_0) \oplus \underline{v}_0^\perp$.
Essendo $\underline{v}_0^\perp$ f-invariante ($\underline{w} \perp \underline{v}_0 \Rightarrow f(\underline{w}) \perp f(\underline{v}_0) = \underline{v}_0$) ed essendo $\phi_{|_{\underline{v}_0^\perp}}$ non degenere (ad esempio perché $Rad(\phi)_{|_{\underline{v}_0^\perp}} = \underline{v}_0^\perp \cap (\underline{v}_0^\perp)^\perp = \underline{v}_0^\perp \cap \mathrm{Span}(\underline{v}_0) = \{\underline{0}\}$), allora per ipotesi induttiva, $f_{|_{\underline{v}_0^\perp}}$ è composizione di un numero finito di riflessioni ortogonali $\tilde{\rho}_1, \dots, \tilde{\rho}_m \in O(\phi_{|_{\underline{v}_0^\perp}})$. Se $\tilde{\rho}_j$ è la riflessione ortogonale di $O(\phi_{|_{\underline{v}_0^\perp}})$ parallela al vettore non isotropo $\underline{w}_j \in \underline{v}_0^\perp$, sia ρ_j la riflessione ortogonale di $O(\phi)$ parallela allo stesso vettore $\underline{w}_j$. Osserviamo che $\rho_j(\underline{v}_0) = \underline{v}_0$ per ogni j, poiché $\underline{v}_0 \perp \underline{w}_j$.
Allora f è la composizione di $\rho_1, \dots, \rho_m$ (infatti questa composizione lascia fisso $\underline{v}_0$, come f, e su $\underline{v}_0^\perp$ coincide con $f_{|_{\underline{v}_0^\perp}}$).
Se tale $\underline{v}_0$ non esiste, allora fissiamo $\underline{v}$ non isotropo. Si avrà che $f(\underline{v}) \neq \underline{v}$.
Supponiamo che $\underline{w} = f(\underline{v}) - \underline{v}$ sia non isotropo (in modo da poter usare la riflessione ortogonale parallela a $\underline{w}$).
Allora, poiché $q_\phi(f(\underline{v})) = q_\phi(\underline{v})$, $\rho_{\underline{w}}(f(\underline{v})) = \underline{v}$.
Per il caso precedente, $g = \rho_{\underline{w}} \circ f \in O(\phi)$ è composizione di un numero finito di riflessioni ortogonali e quindi anche $f = \rho_{\underline{w}} \circ g$ lo è.
Se invece $f(\underline{v}) - \underline{v}$ è isotropo, allora $\underline{w}' = f(\underline{v}) + \underline{v}$ non lo è.
Infatti, se fossero entrambi isotropi, allora da
$0 = q_\phi(f(\underline{v}) - \underline{v}) = q_\phi(f(\underline{v})) + q_\phi(\underline{v}) - 2\phi(\underline{v}, f(\underline{v})) = 2q_\phi(\underline{v}) - 2\phi(\underline{v}, f(\underline{v}))$ e
$0 = q_\phi(f(\underline{v}) + \underline{v}) = q_\phi(f(\underline{v})) + q_\phi(\underline{v}) + 2\phi(\underline{v}, f(\underline{v})) = 2q_\phi(\underline{v}) + 2\phi(\underline{v}, f(\underline{v}))$ otterremmo $4q_\phi(\underline{v}) = 0$, ma $\underline{v}$ è non isotrpo ↯.
Allora, di nuovo poiché $q_\phi(f(\underline{v})) = q_\phi(\underline{v})$, $\rho_{\underline{w}'}(f(\underline{v})) = -\underline{v}$.
Posto $g = (-id_V) \circ \rho_{\underline{w}'} \circ f \in O(\phi)$, g è tale che $g(\underline{v}) = \underline{v}$.
Per il caso precedente, g è composizione di un numero finito di riflessioni ortogonali e quindi anche $f = \rho_{\underline{w}'} \circ (-id_V) \circ g$ lo è. ■

Osserviamo che, ripercorrendo la dimostrazione, esiste una costante intera $C(n)$ (indipendente da ϕ) tale che ogni $f \in O(\phi)$ può essere espresso come composizione di $k \leq C(n)$ riflessioni ortogonali (qui pensiamo che id_V sia composizione di 0 riflessioni ortogonali). Ragionando di nuovo per induzione si ha $C(1) = 1$, $C(n) \leq C(n-1) + n + 1$ (nel caso peggiore in cui si debba usare $-id_V$ che è composizione di n riflessioni ortogonali). Si può quindi prendere $C(n)$ definito per induzione da $C(1) = 1$, $C(n) = C(n-1) + n + 1$, ovvero $C(n) = \frac{n(n+3)}{2} - 1$.
Volendo trovare la costante minima $\tilde{C}(n)$, abbiamo già visto che $\tilde{C}(n) \geq n$. Per i prodotti scalari anisotropi, sempre dalla dimostrazione (in questo caso, non si deve usare $-id_V$), $\tilde{C}(n) = n$.
È un risultato tutt'altro che banale che vale sempre $\tilde{C}(n) = n$. Daremo una dimostrazione più avanti.

Due sottospazi $W, U \subset V$ si dicono *congruenti* se esiste $f \in O(\phi)$ tale che $f(W) = U$.

È immediato verificare che questo definisce una relazione di equivalenza sull'insieme dei sottospazi di V e osserviamo che se due sottospazi W, U sono congruenti allora $(W, \phi_{|_W})$ e $(U, \phi_{|_U})$ sono astrattamente isometrici, infatti la restrizione $f_{|_W} : W \to U$ dà una isometria.

Vogliamo dimostrare che è vero anche il viceversa: se esiste una isometria astratta $g : W \to U$, allora W e U sono congruenti. In effetti vogliamo dimostare di più: esiste $f \in O(\phi)$ tale che $f_{|_W} = g$, ovvero ogni isometria astratta tra sottospazi si estende ad un elemento del gruppo ortogonale.

Nel seguito considereremo un sottospazio di V sempre come munito della restrizione di ϕ; ad esempio, diremo semplicemente che W e U sono isometrici, la restrizione di ϕ essendo sottintesa.

Supponiamo di avere una isometria $g : W \to U$.
Facciamo l'ulteriore ipostesi che $\phi_{|_W}$ sia non degenere (e quindi che anche $\phi_{|_U}$ è non degenere).
In questo caso, possiamo usare quanto visto per le riflessioni ortogonali per concludere.
Ragioniamo per induzione su $m = \dim W = \dim U$. Per $m = 0$, $W = U = \{\underline{0}\}$, $g = 0$ e qualsiasi $f \in O(\phi)$ estende g.
Sia allora $m > 0$, e fissiamo una base orogonale $\underline{w}_1, \dots, \underline{w}_m$ di W. Allora, $\underline{u}_1 = g(\underline{w}_1), \dots, \underline{u}_m = g(\underline{w}_m)$ è una base ortogonale di U. Notiamo che i $\underline{w}_i$ e gli $\underline{u}_j$ sono tutti vettori non isotropi.
Consideriamo $W' = \mathrm{Span}(\underline{w}_1, \dots, \underline{w}_{m-1})$ e $U' = \mathrm{Span}(\underline{u}_1, \dots, \underline{u}_{m-1})$. Poiché $g(W') = U'$, abbiamo che $g_{|_{W'}}$ è una isometria tra W' e U' e quindi per ipotesi induttiva si estende a $f' \in O(\phi)$.
Se $f'(\underline{w}_m) = \underline{u}_m$, allora $f'_{|_W} = g$, e possiamo prendere $f = f'$.
Se invece $f'(\underline{w}_m) \neq \underline{u}_m$, allora consideriamo $\underline{v} = f'(\underline{w}_m) - \underline{u}_m$ e $\underline{z} = f'(\underline{w}_m) + \underline{u}_m$.
Poiché $\underline{w}_m$ non è isotropo, $\underline{v}$ e $\underline{z}$ non possono essere entrambi isotropi: se lo fossero avremmo,
$0 = q_\phi(f'(\underline{w}_m)) + q_\phi(\underline{u}_m) - 2\phi(\underline{u}_m, f'(\underline{w}_m)) = 2q_\phi(\underline{w}_m) - 2\phi(\underline{u}_m, f'(\underline{w}_m))$ e
$0 = q_\phi(f'(\underline{w}_m)) + q_\phi(\underline{u}_m) + 2\phi(\underline{u}_m, f'(\underline{w}_m)) = 2q_\phi(\underline{w}_m) + 2\phi(\underline{u}_m, f'(\underline{w}_m))$,
in quanto $\underline{u}_m = g(\underline{w}_m)$ e g, f' sono isometrie, da cui $4q_\phi(\underline{w}_m) = 0$ ↯.
Se $\underline{v}$ non è isotropo, allora $\rho_{\underline{v}}(f'(\underline{w}_m)) = \underline{u}_m$, mentre per $i = 1 \_ m - 1$, $\rho_{\underline{v}}(f'(\underline{w}_i)) = f'(\underline{w}_i)$, infatti $f'(\underline{w}_i) \perp \underline{v}$ in quanto

$$\begin{aligned}\phi(f'(\underline{w}_i), \underline{v}) &= \phi(\underline{w}_i, \underline{w}_m) - \phi(f'(\underline{w}_i), \underline{u}_m) \\ &= \phi(\underline{w}_i, \underline{w}_m) - \phi(g(\underline{w}_i), g(\underline{w}_m)) \\ &= 0.\end{aligned}$$

Quindi $f = \rho_{\underline{v}} \circ f' \in O(\phi)$ estende g.
Se invece $\underline{v}$ è isotropo, allora $\underline{z}$ non è isotropo e $\rho_{\underline{z}}(f'(\underline{w}_m)) = -\underline{u}_m$. Componendo con $\rho_{\underline{u}_m}$ abbiamo $\rho_{\underline{u}_m}(\rho_{\underline{z}}(f'(\underline{w}_m))) = \underline{u}_m$, mentre per $i = 1 \_ m - 1$ abbiamo $\rho_{\underline{u}_m}(\rho_{\underline{z}}(f'(\underline{w}_i))) = f'(\underline{w}_i)$. Infatti $\rho_{\underline{z}}(f'(\underline{w}_i)) \perp \underline{u}_m$ in quanto

$\phi(\rho_{\underline{z}}(f'(\underline{w}_i)), \underline{u}_m) = \phi(f'(\underline{w}_i), \rho_{\underline{z}}(\underline{u}_m)) = \phi(f'(\underline{w}_i), -f'(\underline{w}_m)) =$
$= -\phi(\underline{w}_i, \underline{w}_m) = 0.$
Inoltre, $f'(\underline{w}_i) \perp \underline{z}$ in quanto
$\phi(f'(\underline{w}_i), \underline{z}) = \phi(\underline{w}_i, \underline{w}_m) + \phi(f'(\underline{w}_i), \underline{u}_m) = \phi(\underline{w}_i, \underline{w}_m) + \phi(g(\underline{w}_i), g(\underline{w}_m)) =$
$2\phi(\underline{w}_i, \underline{w}_m) = 0.$
Quindi $f = \rho_{\underline{u}_m} \circ \rho_{\underline{z}} \circ f' \in O(\phi)$ estende g.

Per ricondurre il caso generale al caso non degenere appena studiato, dobbiamo sviluppare la cosiddetta teoria di Witt, che porterà anche a individuare nuovi invarianti per isometria.

Teoria di Witt

Sia (V, ϕ) uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$ di caratteristica diversa da 2 e ϕ non degenere.

Vogliamo vedere come dato un sottospazio W di V si possa, in modo costruttivo, trovare un sottospazio $W' \supset W$ tale che $\phi_{|_{W'}}$ sia non degenere. Vogliamo inoltre che W' abbia dimensione minima possibile.

Definizione:
Dato $W \subset V$ sottospazio, una *estensione non degenere* di W è un sottospazio $\widetilde{W} \subset V$ tale che $W \subset \widetilde{W}$ e $\phi_{|_{\widetilde{W}}}$ è non degenere.
Un *completamento non degenere* di W è una estensione non degenere di dimensione minima.

Notiamo che V è una estensione non degenere di qualsiasi sottospazio, quindi i completamenti non degeneri esistono. Inoltre, se $\phi_{|_W}$ è non degenere, allora W è l'unico completamento non degenere di W.

Definizione:
Un *piano iperbolico* è uno spazio vettoriale su $\mathbb{K}$ di dimensione 2 munito di un prodotto scalare non degenere e non anisotropo.
Un sottospazio $H \subset V$ si dice *piano iperbolico* se $(H, \phi_{|_H})$ lo è, cioè se $\dim H = 2$, $\phi_{|_H}$ è non degenere e $CI(\phi_{|_H}) \neq \{\underline{0}\}$.

Notiamo che tutti i piani iperbolici sono isometrici, in quanto ogni piano iperbolico ammette una base $\mathcal{B}$, detta *base iperbolica*, per cui la matrice che rappresenta ϕ in tale base è $M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
Infatti, se (H, ϕ) è un piano iperbolico, poiché ϕ non è anisotropo, esiste $\underline{v} \in H$ isotropo non nullo. Completiamo $\underline{v}$ a base di H con $\underline{w}$. La matrice di ϕ in tale base sarà del tipo $\begin{pmatrix} 0 & \phi(\underline{v}, \underline{w}) \\ \phi(\underline{v}, \underline{w}) & \phi(\underline{w}, \underline{w}) \end{pmatrix}$, per cui, essendo ϕ non degenere, $\phi(\underline{v}, \underline{w}) \neq 0$.
Cerchiamo una base iperbolica del tipo $\{\underline{v}, a\underline{v} + b\underline{w}\}$, $a, b \in \mathbb{K}$, $b \neq 0$.
Si deve quindi avere che

$$1 = \phi(\underline{v}, a\underline{v} + b\underline{w}) = a\phi(\underline{v}, \underline{v}) + b\phi(\underline{v}, \underline{w}) = b\phi(\underline{v}, \underline{w})$$

$$0 = q_\phi(a\underline{v} + b\underline{w}) = a^2 q_\phi(\underline{v}) + b^2 q_\phi(\underline{w}) + 2ab\phi(\underline{v}, \underline{w}) = b^2 q_\phi(\underline{w}) + 2a,$$

da cui $b = \frac{1}{\phi(\underline{v}, \underline{w})}$, $a = -\frac{1}{2} b^2 q_\phi(\underline{w})$.

Dato $W \subset V$ sottospazio tale che $\phi_{|_W}$ è degenere, fissiamo U un supplementare di $Rad(\phi_{|_W})$ in W, $W = U \oplus^\perp Rad(\phi_{|_W})$, e fissiamo una base $\underline{z}_1, \ldots, \underline{z}_k$ di $Rad(\phi_{|_W})$ ($\dim Rad(\phi_{|_W}) = k$). Ricordiamo che $\phi_{|_U}$ è non degenere e gli $\underline{z}_i$

sono isotropi.
Poniamo $Z = U \oplus \mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-1})$.
Mostriamo che esiste $\underline{d}_k \in Z^\perp$ ma $\underline{d}_k \notin \underline{z}_k^\perp$. Infatti, se fosse $Z^\perp \subset \underline{z}_k^\perp$, allora passando agli ortogonali e ricordando che ϕ è non degenere, si avrebbe $\mathrm{Span}(\underline{z}_k) \subset Z$, contro il fatto che $\mathrm{Span}(\underline{z}_k)$ e Z sono in somma diretta ↯.
Notiamo che $\underline{z}_k$ e $\underline{d}_k$ sono indipendenti: $\underline{z}_k \neq \underline{0}$ e se fosse $\underline{d}_k \in \mathrm{Span}(\underline{z}_k)$, allora $\phi(\underline{z}_k, \underline{d}_k) = 0$ ↯. Quindi, $H_k = \mathrm{Span}(\underline{z}_k, \underline{d}_k)$ ha dimensione 2.
La matrice della restrizione $\phi_{|_{H_k}}$ nella base $\mathcal{D} = \{\underline{z}_k, \underline{d}_k\}$ è dunque del tipo $M_{\mathcal{D}}(\phi_{|_{H_k}}) = \begin{pmatrix} 0 & a \\ a & b \end{pmatrix}$, dove $a = \phi(\underline{z}_k, \underline{d}_k) \neq 0$, quindi $\phi_{|_{H_k}}$ è non degenere e H_k è un piano iperbolico.
Osserviamo che procedendo come in precedenza, si ottiene una base iperbolica di H_k, che contiene $\underline{z}_k$: esplicitamente, la base iperbolica è data da $\underline{z}_k$ e da $\underline{t}_k = -\frac{1}{2}\frac{b}{a^2}\underline{z}_k + \frac{1}{a}\underline{d}_k$.
Mostriamo che Z e H_k sono in somma diretta.
Infatti, notiamo che $\underline{d}_k$ non appartiene a W, in quanto gli elementi di W sono ortogonali a $\underline{z}_k$ mentre $\underline{d}_k$ non lo è, e quindi, se per $\alpha, \beta \in \mathbb{K}$, $\alpha\underline{z}_k + \beta\underline{d}_k = \underline{u} \in Z$, allora $\beta\underline{d}_k = \underline{u} - \alpha\underline{z}_k \in W$, da cui $\beta = 0$, e allora da $\alpha\underline{z}_k \in Z$ segue $\alpha = 0$.
Inoltre, $\underline{z}_k, \underline{d}_k \in Z^\perp$, per cui $Z \perp H_k$.
Poniamo allora $W_1 = Z \oplus^\perp H_k \supset W$ e $U_1 = U \oplus^\perp H_k$. Osserviamo che $\phi_{|_{U_1}}$ è non degenere, essendo le restrizioni di ϕ a U e a H_k non degeneri.
Mostriamo che $Rad(\phi_{|_{W_1}}) = \mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-1})$.
Infatti, $\underline{z}_1, \ldots, \underline{z}_{k-1}$ sono ortogonali agli elementi di W_1 essendo ortogonali agli elementi di W e a $\underline{d}_k$, quindi $\mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-1}) \subset Rad(\phi_{|_{W_1}})$; inoltre, essendo $\phi_{|_{U_1}}$ non degenere, $Rad(\phi_{|_{W_1}}) \cap U_1 = \{0\}$, per cui usando Grassmann, $\dim Rad(\phi_{|_{W_1}}) \leq \dim W_1 - \dim U_1 = \dim \mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-1})$, da cui l'uguaglianza.
Possiamo allora reiterare il procedimento partendo da W_1, usando U_1 come supplememtare di $Rad(\phi_{|_{W_1}})$ e $\underline{z}_1, \ldots, \underline{z}_{k-1}$ come base di $Rad(\phi_{|_{W_1}})$. Costruiamo quindi $W_2 = U \oplus^\perp H_k \oplus^\perp H_{k-1} \oplus^\perp \mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-2})$, dove H_{k-1} è un piano iperbolico e il radicale di $\phi_{|_{W_2}}$ è $\mathrm{Span}(\underline{z}_1, \ldots, \underline{z}_{k-2})$.
Iterando k volte, otteniamo $W_k = U \oplus^\perp H_k \oplus^\perp \cdots \oplus^\perp H_1$, dove gli H_i sono piani iperbolici, che è una estensione non degenere di W di dimensione $\dim U + 2k = \dim W + \dim Rad(\phi_{|_W})$.
W_k è in realtà un completamento non degenere di W. Infatti, se $\widetilde{W}$ è una qualsiasi estensione non degenere di W, possiamo ripetere la costruzione rimpiazzando V con $\widetilde{W}$: troviamo una estensione non degenere di W isomorfa a W_k contenuta in $\widetilde{W}$, per cui $\dim \widetilde{W} \geq \dim W + \dim Rad(\phi_{|_W})$. W_k realizza quindi il minimo delle dimensioni delle estensioni non degeneri di W, e quindi è un completamento non degenere di W.
Di più, applicando quanto sopra con $\widetilde{W}$ un completamento non degenere di W, otteniamo che i completamenti non degeneri di W sono tutti della forma $U \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_k$, dove gli H_i sono piani iperbolici, U è un supplementare

di $Rad(\phi_{|_W})$ in W e $k = \dim Rad(\phi_{|_W})$.
I completamenti non degeneri di W sono tutti isomorfi, avendo tutti la stessa dimensione $\dim W + \dim Rad(\phi_{|_W})$ e sono anche tutti isometrici, infatti tutti i piani iperbolici sono isometrici e due supplementari di $Rad(\phi_{|_W})$ in W, U, U', sono isometrici: una isometria tra $U \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_k$ e $U' \oplus^\perp H'_1 \oplus^\perp \cdots \oplus^\perp H'_k$ si costruisce sui singoli addendi della somma diretta ortogonale, usando una isometria tra U e U' e, per ogni i, una isometria tra i piani iperbolici H_i e H'_i. Usando quanto sappiamo sulle estensioni delle isometrie nel caso non degenere, otteniamo inoltre che i completamenti non degeneri di W sono tutti congruenti.

Torniamo al problema delle estensioni delle isometrie.
Con le notazioni usate nel caso non degenere, sia $g : W \to U$ una isometria astratta. Sia $Z \subset W$ un supplememtare di $Rad(\phi_{|_W})$, $W = Z \oplus^\perp Rad(\phi_{|_W})$, e applichiamo g a tale decomposizione. Essendo g una isometria abbiamo $U = g(Z) \oplus^\perp g(Rad(\phi_{|_W})) = g(Z) \oplus^\perp Rad(\phi_{|_U})$. Abbiamo dunque che $g(Z) \subset U$ è un supplementare di $Rad(\phi_{|_U})$ in U.
Sia $\mathcal{D}$ una base di W adattata alla somma diretta di cui sopra, allora $g(\mathcal{D})$ è una base di U adattata alla somma diretta di cui sopra. Siano $\widehat{W}$ e $\widehat{U}$ i completamenti non degeneri di W e U ottenuti usando le basi $\mathcal{D}$ e $g(\mathcal{D})$. $\widehat{W} = Z \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_k$, $\widehat{U} = g(Z) \oplus^\perp H'_1 \oplus^\perp \cdots \oplus^\perp H'_k$ (dove gli H_i e gli H'_i sono piani iperbolici e $k = \dim Rad(\phi_{|_W}) = \dim Rad(\phi_{|_U})$). Notiamo che g si può esendere ad una isometria $\hat{g}$ tra $\widehat{W}$ e $\widehat{U}$: lavorando su ogni singolo addendo, $g_{|_Z}$ dà una isometria tra Z e $g(Z)$ e, per ogni $i = 1\_k$, in H_i abbiamo una base iperbolica $\underline{z}_i, \underline{t}_i$ (con $\underline{z}_i \in Rad(\phi_{|_W})$) e in H'_i abbiamo una base iperbolica $g(\underline{z}_i), \underline{t}'_i$ e per definire una isometria tra H_i e H'_i basta mandare $\underline{z}_i$ in $g(\underline{z}_i)$ e t_i in t'_i.
Adesso, possiamo applicare il caso non degenere all'isometria astratta $\hat{g} : \widehat{W} \to \widehat{U}$ per ottenere $f \in O(\phi)$ che estende $\hat{g}$ e quindi estende anche g.

Come corollario otteniamo che se $W, U \subset V$ sono sottospazi isometrici, allora anche gli ortogonali $W^\perp, U^\perp$ sono isometrici. Infatti, W e U sono congruenti tramite una $f \in O(\phi)$, $f(W) = U$. Allora $f(W^\perp) = f(U^\perp)$, e quindi anche gli ortogonali sono congruenti tramite f e dunque isometrici.

Definizione:
Dato (V, ϕ) l'*indice di Witt* di ϕ è

$$W(\phi) = \max\{\dim W |\ W \subset V \text{ sottospazio tale che } \phi_{|_W} = 0\}.$$

Per una matrice simmetrica A, poniamo $W(A) = W(\Phi_A)$.

Notiamo che l'indice di Witt è invariante per isometria/congruenza e che $W(\phi) = 0$ se e solo se ϕ è anisotropo. Inoltre, $W(\phi) = W(\lambda\phi)$ per ogni $\lambda \in \mathbb{K}$, $\lambda \neq 0$.

Poiché $\phi_{|W} = 0$ se e solo se W è contenuto nel cono isotropo di ϕ, $W(\phi)$ è anche la massima dimensione di un sottospazio contenuto in $CI(\phi)$.

La definizione ha senso anche per prodotti scalari degeneri.
In tal caso, osserviamo che se W è un sottospazio tale che $\phi_{|W} = 0$, allora anche $\phi_{|W+Rad(\phi)} = 0$, per cui se W realizza l'indice di Witt ($\dim W = W(\phi)$), $W \supset Rad(\phi)$.
Mostriamo che $W(\phi) = W(\bar{\phi}) + \dim Rad(\phi)$, dove $\bar{\phi}$ è il prodotto scalare non degerene su $V\!/_{Rad(\phi)}$ canonicamente associato a ϕ.
Ricordiamo che se U è un supplementare di $Rad(\phi)$, $(U, \phi_{|U})$ è isometrico a $(V\!/_{Rad(\phi)}, \bar{\phi})$, quindi basta dimostrare che $W(\phi) = W(\phi_{|U}) + \dim Rad(\phi)$.
Sia $Z \subset U$ un sottospazio che realizza l'indice di Witt di $\phi_{|U}$, ovvero $\dim Z = W(\phi_{|U})$ e $\phi_{|Z} = 0$. Allora $Z' = Z \oplus^{\perp} Rad(\phi)$ è un sottospazio tale che $\phi_{|Z'} = 0$, e quindi $W(\phi) \geq \dim Z' = W(\phi_{|U}) + \dim Rad(\phi)$.
Se $\hat{W} \subset V$ realizza l'indice di Witt di ϕ, scriviamo $\hat{W} = U' \oplus Rad(\phi)$, ed osserviamo che U' è contenuto in un supplementare del radicale isometrico a U, per cui $W(\phi_{|U}) \geq \dim U' = W(\phi) - \dim Rad(\phi)$.

In alternativa, notiamo che $\hat{W} + U = V$.
Chiaramante, $\phi_{|W\cap U} = 0$, per cui $W(\phi_{|U}) \geq \dim W \cap U$. Ma $\dim W \cap U = \dim W + \dim U - \dim(W+U) = W(\phi) + n - \dim Rad(\phi) - n = W(\phi) - \dim Rad(\phi)$, da cui $W(\phi) \leq W(\phi_{|U}) + \dim Rad(\phi)$.

Osserviamo infine che se W realizza l'indice di Witt, $W \subset W^{\perp}$, da cui

$$W(\phi) \leq \frac{\dim V + \dim Rad(\phi)}{2}.$$

Tornando al caso ϕ non degenere, se $W \subset V$ è un sottospazio tale che $\phi_{|W} = 0$, il suo completamento non degenere è nella forma $\widehat{W} = H_1 \oplus^{\perp} \cdots \oplus^{\perp} H_m$, dove $m = \dim W$. Otteniamo quindi che $2m = \dim \widehat{W} \leq \dim V$, ovvero $\dim W \leq \frac{\dim V}{2}$ (abbiamo di nunovo $W(\phi) \leq \frac{\dim V}{2}$).
In particolare, se W realizza l'indice di Witt, il suo completamento non degenere è nella forma $\widehat{W} = H_1 \oplus^{\perp} \cdots \oplus^{\perp} H_w$, dove $w = \dim W = W(\phi)$. Se poniamo $A = \widehat{W}^{\perp}$ otteniamo una decomposizione di V in somma diretta ortogonale $V = A \oplus^{\perp} H_1 \oplus^{\perp} \cdots \oplus^{\perp} H_w$.
Osserviamo che A è anisotropo, infatti, se $\underline{v} \in A$, $\underline{v} \neq \underline{0}$, fosse isotropo, allora la restrizione di ϕ a $\mathrm{Span}(\underline{v}) \oplus^{\perp} W$ sarebbe nulla, contro la massimalità di W ↯.
Una decomposizione di V in somma diretta ortogonale con queste caratteristiche $V = A' \oplus^{\perp} H'_1 \oplus^{\perp} \cdots \oplus^{\perp} H'_s$, con una una parte anisotropa A' e s piani iperbolici H'_j, si dice una *decomposizione di Witt* di (V, ϕ) (notiamo che è possibile che sia $A' = \{\underline{0}\}$ o che $s = 0$).
Abbiamo appena mostrato che le decomposizioni di Witt esistono, costruendone

una che contiene $w = W(\phi)$ piani iperbolici.
Vogliamo mostrare che due decomposizioni di Witt di (V, ϕ) sono congruenti, ovvero che se $V = A \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_w$ (decomposizione di Witt costruita a partire da un sottospazio che realizza l'indice di Witt) e $V = A' \oplus^\perp H'_1 \oplus^\perp \cdots \oplus^\perp H'_s$ è un'altra decomposizione di Witt allora $s = w$ ed esiste $f \in O(\phi)$ tale che $f(A) = A'$ e $f(H_i) = H'_i$ per $i = 1\_w$.

Scegliendo un vettore isotropo $\underline{v}_j$ in ognuno dei piani iperbolici H'_j otteniamo $U = \mathrm{Span}(\underline{v}_1, \ldots, \underline{v}_s)$ di dimensione s e tale che $\phi_{|_U} = 0$, per cui $s \leq w$.
Supponiamo per assurdo che $s < w$.
Allora $H_1 \oplus^\perp \cdots \oplus^\perp H_s$ e $H'_1 \oplus^\perp \cdots \oplus^\perp H'_s$ sono isometrici (tramite una isometria che manda H_i in H'_i, $i = 1\_s$) e quindi congruenti tramite $f \in O(\phi)$ ($f(H_i) = H'_i$, $i = 1\_s$).
Ma allora, passando agli ortogonali, $f(A \oplus^\perp H_{s+1} \oplus^\perp \cdots \oplus^\perp H_w) = A'$ ma A' è anisotropo, mentre nei piani iperbolici $H_{s+1}, \ldots, H_w$ esistono vettori isotropi non nulli ↯.
Quindi $s = w$ e $f(A) = A'$ come voluto.

L'indice e la decomposizione di Witt (che valgono su un campo $\mathbb{K}$ arbitrario, $char\mathbb{K} \neq 2$) riducono lo studio dei prodotti scalari non degeneri a meno di isometria al caso dei prodotti scalari anisotropi. La struttura di questi ultimi dipende fortemente dalle proprietà algebriche del campo $\mathbb{K}$.
Ad esempio, se $\dim V = 2W(\phi)$ (si dice che (V, ϕ) è *neutro*) allora V è somma diretta ortogonale di $W(\phi)$ piani iperbolici e quindi tutti gli spazi neutri della stessa dimensione sono isometrici.
Otteniamo che l'indice di Witt e la classe di isometria della parte anisotropa di una decomposizione di Witt sono invarianti completi per isometria nel caso non degenere.

Nel caso generale (in cui ϕ non è necessariamente non degenere), una decomposizione di Witt è una decomposizione di V del tipo

$$V = Rad(\phi) \oplus^\perp A \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_w,$$

con A anisotropo e gli H_j piani iperbolici, ottenuta da una decomposizione di Witt di un supplementare del radicale. Anche in questo caso, due decomposizioni di Witt sono congruenti, ma notiamo che il numero di piani iperbolici è $w = W(\phi) - \dim Rad(\phi)$.
Otteniamo che la dimensione del radicale, l'indice di Witt e la classe di isometria della parte anisotropa di una decomposizione di Witt sono invarianti completi per isometria.

Osservazione: la parte finale di questo capitolo non è materia di esame.

Abbiamo adesso tutti gli strumenti per stimare il numero di riflessioni ortogonali che servono per esprimere un elemento del gruppo ortogonale.
Ricordiamo che esiste una minima costante $\tilde{C}(n)$ tale che, se ϕ è un prodotto scalare non degenere su uno spazio di dimensione n, ogni $f \in O(\phi)$ è composizione di $\tilde{C}(n)$ riflessioni ortogonali. Abbiamo già osservato che $\tilde{C}(n) \geq n$ e adesso vogliamo dimostrare che $\tilde{C}(n) = n$.
Ripercorrendo la dimostrazione del fatto che il gruppo ortogonale è generato dalle riflessioni ortogonali, si presentavano tre casi:
- esiste $\underline{v} \in V$ non isotropo tale che $f(\underline{v}) = \underline{v}$: in tal caso bastavano al più $\tilde{C}(n-1)$ riflessioni ortogonali per esprimere f;
- esiste $\underline{v} \in V$ non isotropo tale che $f(\underline{v}) \neq \underline{v}$ ma $f(\underline{v}) - \underline{v}$ non è isotropo: in tal caso bastavano al più $\tilde{C}(n-1)+1$ riflessioni ortogonali per esprimere f;
questi due casi li chiamiamo "favorevoli", per distinguerli dal seguente caso "sfavorevole"
- per ogni $\underline{v} \in V$ non isotropo $f(\underline{v}) \neq \underline{v}$ e $f(\underline{v}) - \underline{v}$ è isotropo: in tal caso servivano al più $\tilde{C}(n-1)+1+n$ riflessioni ortogonali per esprimere f;
Analizziamo meglio il caso "sfavorevole".

Proposizione
Sia (V, ϕ), con $\dim V = n$ e ϕ non degenere, e sia $f \in O(\phi)$. Se per ogni $\underline{v} \in V$, $\underline{v}$ non isotropo, $f(\underline{v}) - \underline{v} \neq 0$ e $f(\underline{v}) - \underline{v}$ è isotropo, allora (V, ϕ) è neutro e $\det f = 1$.

Dimostrazione
Per $n = 1$, le ipotesi della proposizione non si verificano mai, quindi è evidentemente vera.
Analizziamo il caso $n = 2$.
Per ipotesi esiste un vettore isotropo non nullo. Quindi (V, ϕ) è un piano iperbolico (e quindi neutro).
Fissiamo una base iperbolica $\mathcal{B} = \{\underline{z}, \underline{w}\}$ per V, $M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = H$. Poiché $f \in O(\phi)$, il cono isotropo di ϕ è f-invariante. Ne segue che la matrice P di f nella base $\mathcal{B}$ può essere di una delle due forme

$$P = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \ \lambda\mu = 1$$

oppure

$$P = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix}, \ \lambda\mu = 1 \ .$$

Nel primo caso $\det P = \lambda\mu = 1$, nel secondo $\det P = -\lambda\mu = -1$. Vogliamo escludere che $\det P = -1$. Un vettore $\underline{v} = \alpha\underline{z} + \beta\underline{w}$, $\alpha, \beta \in \mathbb{K}$, non è isotropo se e solo se $\alpha \neq 0$ e $\beta \neq 0$. Se fosse $\det P = -1$,

$$f(\underline{v}) - \underline{v} = (\lambda\beta - \alpha)\underline{z} + (\mu\alpha - \beta)\underline{w},$$

è isotropo se e solo se $\alpha = \lambda\beta$. Applichiamo queste considerazioni ai vettori non isotropi $z + w$, $z - w$. Avremmo $1 = \lambda$ e $1 = -\lambda$ che è impossibile (perché la caratteristica di $\mathbb{K}$ è diversa da 2). Osserviamo, d'altra parte, che se $\det P = 1$, e $\underline{v}$ non è isotropo,

$$g(\underline{v}) - \underline{v} = (\lambda - 1)\alpha\underline{z} + (\mu - 1)\beta\underline{w}$$

è isotropo se e solo se $\lambda = \mu = 1$, per cui $f = id_V$.
Dunque anche per $n = 2$ le ipotesi della proposizione non si verificano mai.

Sia ora $n > 2$. Poniamo

$$W = \mathrm{Ker}(f - id_V)$$

ed oesserviamo che $W \subset CI(\phi)$, è f-invariante e $f_{|_W} = id_W$. Infatti, per ipotesi, se $\underline{v}$ non è isotropo, $f(\underline{v}) \neq \underline{v}$, quindi tutti i vettori di $\mathrm{Ker}(f - id_V)$ sono isotropi.
Vogliamo dimostrare che $\dim(V) = 2\dim(W)$, per cui W realizza l'indice di Witt di ϕ e dunque (V, ϕ) è neutro.
La dimostrazione di questa affermazione segue dai seguenti due fatti:
(1) Per ogni $\underline{u} \in W$, per ogni $\underline{w} = f(\underline{v}) - \underline{v} \in \mathrm{Im}(f - id_V)$, $\phi(\underline{u}, \underline{w}) = 0$.
(2) La restrizione di ϕ su $\mathrm{Im}(f - id_V)$ è nulla.
Dimostriamo (1):

$$\begin{aligned}\phi(\underline{u}, f(\underline{v}) - \underline{v}) = \phi(\underline{u}, f(\underline{v})) - \phi(\underline{u}, \underline{v}) = \phi(\underline{u}, f(\underline{v}) - \phi(f(\underline{u}), f(\underline{v})) = \\ = \phi(\underline{u} - f(\underline{u}), f(\underline{v})) = \phi(\underline{0}, f(\underline{v})) = 0.\end{aligned}$$

Dimostriamo (2):
se $\underline{v}$ non è isotropo, per ipotesi, $f(\underline{v}) - \underline{v}$ è isotropo. Se invece $\underline{v}$ è isotropo, essendo $\dim(\underline{v}^\perp) = n - 1 > n/2$ (perché $n > 2$), si ha $\dim(\underline{v}^\perp) > W(\phi)$, per cui esiste $\underline{w} \in v^\perp$ non isotropo. Ne segue che

$$q_\phi(\underline{v} + \underline{w}) = q_\phi(\underline{v} - \underline{w}) = q_\phi(\underline{w}) \neq 0.$$

Applicando l'ipotesi a questi tre vettori non isotropi, abbiamo che:
$f(\underline{w}) - \underline{w}$ è isotropo;
$f(\underline{v} + \underline{w}) - (\underline{v} + \underline{w}) = (f(\underline{v}) - \underline{v}) + (f(\underline{w}) - \underline{w})$ è isotropo;
$f(\underline{v} - \underline{w}) - (\underline{v} - \underline{w}) = (f(\underline{v}) - \underline{v}) - (f(\underline{w}) - \underline{w})$ è isotropo.
Calcolando:

$$\begin{aligned}0 = \phi((f(\underline{v}) - \underline{v}) + (f(\underline{w}) - \underline{w}), (f(\underline{v}) - \underline{v}) + (f(\underline{w}) - \underline{w})) + \\ + \phi((f(\underline{v}) - \underline{v}) - (f(\underline{w}) - \underline{w}), (f(\underline{v}) - \underline{v}) - (f(\underline{w}) - \underline{w})) = \\ 2\phi(f(\underline{v}) - \underline{v}, f(\underline{v}) - \underline{v}) + 2\phi(f(\underline{w}) - \underline{w}, f(\underline{w}) - \underline{w}) = 2\phi(f(\underline{v}) - \underline{v}, f(\underline{v}) - \underline{v})\ .\end{aligned}$$

Quindi $f(\underline{v}) - \underline{v}$ è isotropo. In conclusione, tutti i vettori $\underline{v} \in \mathrm{Im}(f - id_V)$ sono isotropi.
Dai punti (1) e (2) e dal fatto che ϕ è nulla su W, deduciamo:

$$\mathrm{Im}(f - id_V) \subset W^\perp \text{ e } W \subset W^\perp$$

$$W \subset (\mathrm{Im}(f - id_V))^\perp \text{ e } \mathrm{Im}(f - id_V) \subset (\mathrm{Im}(f - id_V))^\perp.$$
$$\dim(W^\perp) = n - \dim(W) = \dim(\mathrm{Im}(f - id_V)),$$

ne segue che

$$\mathrm{Im}(f - id_V) = \mathrm{Ker}(f - id_V)^\perp.$$

Poiché $W \subset W^\perp$, $n - \dim W \geq \dim W$,

$$\dim V \geq 2\dim(W).$$

Analogamente

$$W = (\mathrm{Im}(f - id_V))^\perp;$$

poiché $\mathrm{Im}(f - id_V) \subset (\mathrm{Im}(f - id_V))^\perp$, $\dim(W) \geq n - \dim(W)$,

$$2\dim(W) \geq \dim(V).$$

In conclusione

$$W = \mathrm{Im}(f - id_V) \text{ e } \dim V = 2\dim W$$

come volevamo.
Abbiamo dimostrato che (V, ϕ) è neutro; inoltre, per costruzione, disponiamo di un sottospazio W che realizza l'indice di Witt, $W(\phi) = n/2$, e tale che $f_{|_W} = id_W$. Resta da dimostrare che $\det(f) = 1$. V è un completamento non degenere di W,

$$V = H_1 \oplus^\perp \cdots \oplus^\perp H_{n/2}$$

e disponiamo di una base di V $\{\underline{z}_1, \underline{w}_1, \ldots, \underline{z}_{n/2}, \underline{w}_{n/2}\}$ adattata alla somma diretta tale che ogni coppia $\{\underline{z}_j, \underline{w}_j\}$ è una base iperbolica del piano iperbolico H_j, mentre $\{\underline{z}_1, \ldots, \underline{z}_{n/2}\}$ è una base di W. Permutando i vettori di questa base di V, consideriamo la base

$$\mathbb{B} = \{\underline{z}_1, \ldots, \underline{z}_{n/2}, \underline{w}_1, \ldots, \underline{w}_{n/2}\}.$$

Allora

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & I_{n/2} \\ I_{n/2} & 0 \end{pmatrix}.$$

La matrice che rappresenta f è della forma

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} I_{n/2} & B \\ 0 & C \end{pmatrix};$$

$$f \in O(\Phi) \Leftrightarrow \begin{pmatrix} 0 & C \\ C^\top & C^\top B + B^\top C \end{pmatrix} = \begin{pmatrix} 0 & I_{n/2} \\ I_{n/2} & 0 \end{pmatrix};$$

in particolare,

$$C^\top = I_{n/2} = C$$

da cui

$$\det M_{\mathcal{B}}^{\mathcal{B}}(f) = 1.$$

Teorema (Cartan-Dieudonné)
Sia (V, ϕ) con $\dim V = n$ e ϕ non degenere. Ogni $f \in O(\Phi)$ è composizione al più di n riflessioni ortogonali: $\tilde{C}(n) = n$.

Dimostrazione
Supponiamo $f \neq id_V$ e procediamo per induzione su $n \geq 1$. Per $n = 1$ abbiamo visto che $C(1) = 1$. Sia ora $n > 1$. Se esiste $\underline{v}$ non isotropo tale che $f(\underline{v}) = \underline{v}$ oppure $f(\underline{v}) \neq \underline{v}$ e $f(\underline{v}) - \underline{v}$ è non isotropo, allora siamo in uno dei casi "favorevoli" e, per ipotesi induttiva, bastano al più $\tilde{C}(n-1) = n-1$ oppure $\tilde{C}(n-1)+1 = n$ riflessioni ortogonali per esprimere f.
Altrimenti, siamo nel caso "sfavorevole" e applicando la proposizione precedente, (V, ϕ) è neutro e $\det f = 1$. Sia ρ una qualsiasi riflessione ortogonale (parallela a qualche $\underline{w}$ non isotropo). Allora $\det(\rho \circ f) = -1$ e, sempre applicando la proposizione, $\rho \circ f$ ricade nei casi "favorevoli" precedenti e quindi è composizione di k riflessioni ortogonali $\rho_1, \ldots, \rho_k$, $k \leq \dim V = 2W(\phi)$.
Ne segue che $f = \rho \circ \rho_1 \circ \cdots \circ \rho_k$ e che $1 = \det f = (-1)^{k+1}$, per cui k è dispari e dunque $k < 2W(\phi)$, ovvero $k + 1 \leq \dim V = n$.
Anche nel caso "sfavorevole" quindi bastano al più n riflessioni ortogonali.
Sapendo che $\tilde{C}(n) \geq n$, per induzione si conclude che $\tilde{C}(n) = n$.

Classificazione dei prodotti scalari reali e complessi.

Classificare i prodotti scalari a meno di isometria nel caso $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$ è reso facile dalla possibilità di caratterizzare completamente i prodotti scalari anisotropi.
Nel seguito, V sarà uno spazio vettoriale su $\mathbb{K}$, dove $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$ e ϕ un prodotto scalare su V. Per differenziare i due casi, diremo che ϕ è un prodotto scalare reale o complesso rispettivamente.

Definizione
Un prodotto scalare reale ϕ si dice:

- *definito positivo* se per ogni $\underline{v} \in V$, $\underline{v} \neq \underline{0}$, $q_\phi(\underline{v}) = \phi(\underline{v}, \underline{v}) > 0$;
- *definito negativo* se per ogni $\underline{v} \in V$, $\underline{v} \neq \underline{0}$, $q_\phi(\underline{v}) < 0$;
- *semidefinito positivo* se per ogni $\underline{v} \in V$, $q_\phi(\underline{v}) \geq 0$;
- *semidefinito negativo* se per ogni $\underline{v} \in V$, $q_\phi(\underline{v}) \leq 0$.

Se ϕ è definito positivo, data una base ortogonale $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ di (V, ϕ), abbiamo $q_\phi(\underline{v}_i) > 0$, per ogni $i = 1\_n$.
Sostituiamo $\underline{v}_i$ con $\underline{v}'_i = t_i \underline{v}_i$ scegliendo $t_i \in \mathbb{R}$ tale che $t_i^2 = \frac{1}{q_\phi(\underline{v}_i)}$. Allora, $\phi(\underline{v}'_i, \underline{v}'_i) = t_i^2 \phi(\underline{v}_i, \underline{v}_i) = 1$, mentre $\phi(\underline{v}'_i, \underline{v}'_j) = 0$ se $i \neq j$. La base $\mathcal{B}' = \{\underline{v}'_1, \ldots, \underline{v}'_n\}$ si dice *base ortonormale* di V (per ϕ) ed è caratterizzata dal fatto che $M_{\mathcal{B}'}(\phi) = I_n$.
Poiché ogni prodotto scalare definito positivo su V ammette una base ortonormale, tutti i prodotti scalari definiti positivi su V sono isometrici fra loro.
Se ϕ è definito negativo, allora $-\phi$ è definito positivo e operando come sopra si trova una base $\mathcal{B}'$ di V tale che $M_{\mathcal{B}'}(\phi) = -I_n$.
Poiché ogni prodotto scalare definito negativo su V ammette una base di questo tipo, tutti i prodotti scalari definiti negativi su V sono isometrici fra loro.

Osserviamo che i prodotti scalari definiti (sia positivi che negativi) sono necessariamente anisotropi. È vero anche il viceversa, per cui i prodotti scalari anisotropi su $\mathbb{R}$ sono esattamente quelli definiti. Invece, su $\mathbb{C}$, i prodotti scalari anisotropi sono possibili solo in dimensione 1.

Proposizione
1) Un prodotto scalare reale ϕ è anisotropo se e solo se è definito.
2) Un prodotto scalare complesso ϕ è anisotropo se e solo se è non degenere e $\dim V = 1$.

Dimostrazione
Le implicazioni verso sinistra sono ovvie, dimostriamo quelle verso destra.
1) Mostriamo che se ϕ non è definito, allora non è anisotropo.
Negare che ϕ sia definito significa che esistono $\underline{v}, \underline{w} \in V$ non nulli tali che $\phi(\underline{v}, \underline{v}) \geq 0$ e $\phi(\underline{w}, \underline{w}) \leq 0$. Se uno dei due è isotropo abbiamo finito, altrimenti

abbiamo $\phi(\underline{v},\underline{v}) > 0$ e $\phi(\underline{w},\underline{w}) < 0$. Per $t \in \mathbb{R}$, consideriamo $\underline{z} = \underline{w} + t\underline{v}$. Se fosse $\underline{z} = \underline{0}$, allora $0 > \phi(\underline{w},\underline{w}) = \phi(-t\underline{v},-t\underline{v}) = t^2\phi(\underline{v},\underline{v}) \geq 0$, ↯. Quindi $\underline{z} \neq \underline{0}$ e $\phi(\underline{z},\underline{z}) = \phi(\underline{w},\underline{w}) + 2t\phi(\underline{w},\underline{v}) + t^2\phi(\underline{v},\underline{v})$. Poiché $\phi(\underline{w},\underline{v})^2 - \phi(\underline{v},\underline{v})\phi(\underline{w},\underline{w}) > 0$, esistono due valori di t per cui $\underline{z}$ è isotropo.
2) Se fosse $\dim V \geq 2$, ragionando come sopra, in ogni sottospazio di dimensione 2 esisterebbero vettori isotropi non nulli ↯. Quindi $\dim V = 1$ e se ϕ fosse degenere allora sarebbe nullo e quindi ogni vettore di V sarebbe isotropo ↯. ❑

Quindi, se ϕ è un prodotto scalare reale una decomposizione di Witt è del tipo $V = Rad(\phi) \oplus^\perp A \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_w$ con $\phi_{|_A}$ definito, dove ricordiamo che sono possibili i casi $Rad(\phi) = \{\underline{0}\}$, $A = \{\underline{0}\}$, $w = 0$. Ad esempio, ϕ è definito se e solo se $Rad(\phi) = \{\underline{0}\}$ e $w = 0$; mentre ϕ è semidefinito se e solo se $w = 0$.
Definiamo il *segno* di ϕ, indicato con $sgn(\phi)$, come $+1$ se $A \neq \{\underline{0}\}$ e $\phi_{|_A}$ è definito positivo, -1 se $A \neq \{\underline{0}\}$ e $\phi_{|_A}$ è definito negativo, 0 se $A = \{\underline{0}\}$.
Allora la dimensione del radicale (o il rango), l'indice di Witt e il segno di un prodotto scalare reale sono invarianti completi per isometria sulla classe degli spazi vettoriali di dimensione fissata n.

Se invece ϕ è un prodotto scalare complesso una decomposizione di Witt è del tipo $V = Rad(\phi) \oplus^\perp A \oplus^\perp H_1 \oplus^\perp \cdots \oplus^\perp H_w$ con $A = \{\underline{0}\}$ o $\dim A = 1$ e $\phi_{|_A}$ non degenere. Osserviamo che l'addendo anisotropo ha dimensione 0 se e solo se $\dim Rad(\phi)$ e $\dim V$ hanno la stessa parità, ovvero se e solo se $\mathrm{rnk}(\phi)$ è pari.
Notiamo che se $\dim V = 1$ tutti i prodotti scalari non degeneri sono isometrici: se $V = \mathrm{Span}(\underline{v})$, $\phi, \psi \in PS(V)$, $\phi(\underline{v},\underline{v}) = \alpha \neq 0$, $\psi(\underline{v},\underline{v}) = \beta \neq 0$, allora $F : V \to V$ lineare tale che $F(\underline{v}) = \lambda\underline{v}$ con λ una radice quadrata di α/β è una isometria tra (V,ϕ) e (V,ψ). Infatti,
$\psi(F(\underline{v}),F(\underline{v})) = \psi(\lambda\underline{v},\lambda\underline{v}) = \lambda^2\beta = \alpha = \phi(\underline{v},\underline{v})$.
Allora, la dimensione del radicale (o il rango) di un prodotto scalare complesso è un invariante completo per isometria sulla classe degli spazi vettoriali di dimensione fissata n.

Osserviamo che questo risultato segue anche da un ragionamento analogo a quello fatto per i prodotti scalari reali definiti.
Se ϕ è un prodotto scalare complesso, data una base ortogonale $\mathcal{B} = \{\underline{v}_1,\ldots,\underline{v}_n\}$ di (V,ϕ), a meno di riordinare i vettori, possiamo supporre che per i primi $k = \mathrm{rnk}(\phi)$ vettori $q_\phi(\underline{v}_i) \neq 0$, mentre i rimanenti siano isotropi e diano una base di $Rad(\phi)$.
Per $i = 1\_k$, sostituiamo $\underline{v}_i$ con $\underline{v}'_i = t_i\underline{v}_i$ scegliendo $t_i \in \mathbb{C}$ tale che $t_i^2 = \frac{1}{q_\phi(\underline{v}_i)}$.
Osserviamo che $\phi(\underline{v}'_i,\underline{v}'_i) = t_i^2\phi(\underline{v}_i,\underline{v}_i) = 1$, mentre $\phi(\underline{v}'_i,\underline{v}'_j) = 0$ se $i \neq j$.
Otteniamo una nuova base ortogonale $\mathcal{B}'$ di (V,ϕ) tale che

$$M_{\mathcal{B}'}(\phi) = diag(I_k, 0).$$

Una tale base si dice *base ortogonale normalizzata* (e $\underline{v}'_i$ si dice ottenuto normalizzando $\underline{v}_i$).

L'esistenza di basi ortogonali normalizzate e il fatto che la matrice associata ad un prodotto scalare in una base ortogonale normalizzata dipende solo dalla dimensione dello spazio vettoriale e dal rango del prodotto scalare, ovvero la discussione fatta precedentemente usando la teoria di Witt, danno la classificazione completa dei prodotti scalari complessi a meno di isometrie.

Teorema ($\mathbb{C}$)
Se $\mathbb{K} = \mathbb{C}$, la coppia $(\dim V, \mathrm{rnk}(\phi))$ è un invariante completo per isometria: (V, ϕ) e (W, ψ) sono isometrici se e solo se $\dim V = \dim W$ e $\mathrm{rnk}(\phi) = \mathrm{rnk}(\psi)$.

Vediamo come ottenere la classificazione per i prodotti scalari reali seguendo lo stesso ordine di idee.
Se ϕ è un prodotto scalare reale, data una base ortogonale $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ di (V, ϕ), a meno di riordinare i vettori, possiamo supporre che i primi p siano tali che $q_\phi(\underline{v}_i) > 0$, i successivi q siano tali che $q_\phi(\underline{v}_j) < 0$, mentre i rimanenti siano isotropi e diano una base di $Rad(\phi)$ (notiamo che $p + q = \mathrm{rnk}(\phi)$).
Per $i = 1 \_ p$, possiamo operare come prima: normalizziamo $\underline{v}_i$ sostituendolo con $\underline{v}'_i = t_i \underline{v}_i$ scegliendo $t_i \in \mathbb{R}$ tale che $t_i^2 = \frac{1}{q_\phi(\underline{v}_i)}$. Di nuovo, $\phi(\underline{v}'_i, \underline{v}'_i) = t_i^2 \phi(\underline{v}_i, \underline{v}_i) = 1$.
Per $j = p + 1 \_ p + q$ invece, normalizziamo $\underline{v}_j$ sostituendolo con $\underline{v}'_j = t_j \underline{v}_j$ scegliendo $t_j \in \mathbb{R}$ tale che $t_j^2 = -\frac{1}{q_\phi(\underline{v}_j)}$. In questo caso, $\phi(\underline{v}'_j, \underline{v}'_j) = t_j^2 \phi(\underline{v}_j, \underline{v}_j) = -1$.
Otteniamo una nuova base ortogonale $\mathcal{B}'$ di (V, ϕ) tale che

$$M_{\mathcal{B}'}(\phi) = diag(I_p, -I_q, 0).$$

Anche in questo caso, una tale base si dice *base ortogonale normalizzata*.

Per concludere la classificazione nel caso $\mathbb{K} = \mathbb{R}$, mostriamo che dato $\phi \in PS(V)$, per ogni base ortogonale normalizzata, il numero p di entrate positive e il numero q di entrate negative in $M_{\mathcal{B}}(\phi)$ non dipendono dalla scelta della base, ma sono un carattere intrinseco del prodotto scalare (e quindi invarianti per isometria).

Definizione:
Sia V uno spazio vettoriale su $\mathbb{R}$ e sia $\phi \in PS(V)$.
L'*indice di positività* di ϕ è dato da

$$i_+(\phi) = \max\{\dim W |\ W \subset V \text{ sottospazio tale che } \phi_{|_W} \text{ è definito positivo}\}.$$

L'*indice di negatività* di ϕ è dato da

$$i_-(\phi) = \max\{\dim W |\ W \subset V \text{ sottospazio tale che } \phi_{|_W} \text{ è definito negativo}\}.$$

La *segnatura* di ϕ è la terna

$$\sigma(\phi) = (i_+(\phi), i_-(\phi), \dim Rad(\phi))$$

la dimensione di $Rad(\phi)$ si dice anche *indice di nullità* di ϕ e si indica con $i_0(\phi)$.

Ad esempio, ϕ è definito positivo se e solo se $\sigma(\phi) = (n, 0, 0)$; è definito negativo se e solo se $\sigma(\phi) = (0, n, 0)$; è semidefinito positivo se e solo se $i_-(\phi) = 0$; è semidefinito negativo se e solo se $i_+(\phi) = 0$.
Poiché una isometria manda vettori con forma quadratica positiva (rispettivamente negativa, nulla), in vettori con forma quadratica positiva (rispettivamente negativa, nulla), gli indici di positività e negatività, e dunque la segnatura, sono invarianti per isometria.

Teorema (Sylvester)
Per ogni base ortogonale normalizzata $\mathcal{B}$ di (V, ϕ) il numero di entrate positive in $M_{\mathcal{B}}(\phi)$ è $i_+(\phi)$, il numero di entrate negative in $M_{\mathcal{B}}(\phi)$ è $i_-(\phi)$ e il numero di entrate nulle in $M_{\mathcal{B}}(\phi)$ è $i_0(\phi)$.

Dimostrazione
Sia $\mathcal{B} = \{\underline{v}_1, \dots \underline{v}_p, \underline{w}_1, \dots, \underline{w}_q, \underline{z}_1, \dots, \underline{z}_{n-p-q}\}$ una base ortogonale normalizzata di (V, ϕ), dove $n = \dim V$, $q_\phi(\underline{v}_i) = 1$, $i = 1\_p$, $q_\phi(\underline{w}_j) = -1$, $j = 1\_q$, $q_\phi(\underline{z}_k) = 0$, $k = 1\_n - p - q$.
Mostriamo che $p = i_+(\phi)$, $q = i_-(\phi)$ e $n - p - q = i_0(\phi)$.
Sappiamo già che $\underline{z}_1, \dots, \underline{z}_{n-p-q}$ danno una base di $Rad(\phi)$ e quindi abbiamo $n - p - q = i_0(\phi)$.
Consideriamo $W = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_p)$. Poiché $q_\phi(\alpha_1\underline{v}_1 + \dots + \alpha_p\underline{v}_p) = \sum_{i=1}^{p} \alpha_i^2$, per ogni $\alpha_1, \dots, \alpha_p \in \mathbb{R}$, $\phi_{|_W}$ è definito positivo. Quindi $i_+(\phi) \geq \dim W = p$.
Sia adesso $U \subset V$ un sottospazio di dimensione $i_+(\phi)$ tale che $\phi_{|_U}$ è definito positivo (si dice che U realizza l'indice di positività), e consideriamo il sottospazio $Z = \mathrm{Span}(\underline{w}_1, \dots, \underline{w}_q, \underline{z}_1, \dots, \underline{z}_{n-p-q})$. Analogamante a prima, $\phi_{|_Z}$ è semidefinito negativo. Se $\underline{v} \in U \cap Z$ abbiamo $q_\phi(\underline{v}) \geq 0$ in quanto $\underline{v} \in U$, e $q_\phi(\underline{v}) \leq 0$ in quanto $\underline{v} \in Z$, e quindi $q_\phi(\underline{v}) = 0$.
Ma l'unico vettore isotropo in U è il vettore nullo, quindi $\underline{v} = \underline{0}$ e U e Z sono in somma diretta. Dalla formula di Grassmann allora si ha

$$i_+(\phi) + q + n - p - q = \dim U + \dim Z \leq \dim V = n,$$

da cui $i_+(\phi) \leq p$.
Dunque, $i_+(\phi) = p$.
In modo analogo (usando $W' = \mathrm{Span}(\underline{w}_1, \dots, \underline{w}_q)$ su cui ϕ è definito negativo, U' che realizza l'indice di negatività e $Z' = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_p, \underline{z}_1, \dots, \underline{z}_{n-p-q})$ su cui ϕ è semidefinito positivo), si ha $i_-(\phi) = q$. ∎

Osserviamo che $i_+(\phi) + i_-(\phi) = \mathrm{rnk}(\phi)$ e $i_+(\phi) + i_-(\phi) + i_0(\phi) = \dim V$.

Il teorema di Sylvester, oltre a dare un modo esplicito per calcolare la segnatura (basta produrre una base ortogonale e contare i vettori a forma quadratica positiva/negativa/nulla), ci permette di dare la classificazione completa dei prodotti scalari reali a meno di isometrie in termini della segnatura.

Teorema ($\mathbb{R}$)
Se $\mathbb{K} = \mathbb{R}$, la segnatura è un invariante completo per isometria: (V, ϕ) e (W, ψ) sono isometrici se e solo se $\sigma(\phi) = \sigma(\psi)$.

Possiamo esprimere la segnatura di ϕ per mezzo degli invarianti trovati in precedenza $(\dim Rad(\phi), W(\phi), sgn(\phi))$, osservando che un piano iperbolico reale ha segnatura $(1, 1, 0)$ (essendo non degenere ma non anisotropo, la matrice associata in una base ortogonale normalizzata è $diag(1, -1)$) e che la segnatura è additiva sulle somme dirette ortogonali (poiché l'unione di basi ortogonali normalizzate degli addendi dà una base ortogonale normalizzata della somma). Allora data una decomposizione di Witt, $V = Rad(\phi) \oplus^{\perp} A \oplus^{\perp} H_1 \oplus^{\perp} \cdots \oplus^{\perp} H_w$, $\sigma(\phi) = \sigma(\phi_{|_A}) + w(1, 1, 0)$ e $\sigma(\phi_{|_A}) = (m, 0, 0)$ o $(0, m, 0)$ con $m = \dim A$. Ricordiamo che $w = W(\phi) - \dim Rad(\phi)$ e $\dim A = \dim V - \dim Rad(\phi) - 2w$. Otteniamo

$$i_+(\phi) = W(\phi) - \dim Rad(\phi) + \max\{0, sgn(\phi)(\dim V - 2W(\phi) + \dim Rad(\phi))\},$$

$$i_-(\phi) = W(\phi) - \dim Rad(\phi) + \max\{0, -sgn(\phi)(\dim V - 2W(\phi) + \dim Rad(\phi))\}.$$

Quindi,

$$\begin{aligned} sgn(\phi) = 1 &\Rightarrow i_+(\phi) = \dim V - W(\phi),\ i_-(\phi) = W(\phi) - \dim Rad(\phi), \\ sgn(\phi) = -1 &\Rightarrow i_+(\phi) = W(\phi) - \dim Rad(\phi),\ i_-(\phi) = \dim V - W(\phi), \\ sgn(\phi) = 0 &\Rightarrow i_+(\phi) = i_-(\phi) = W(\phi) - \dim Rad(\phi). \end{aligned}$$

Viceversa, possiamo ricavare indice di Witt e segno a partire dalla segnatura: riordinando i vettori di una base ortogonale normalizzata otteniamo una decomposizione di Witt con $\min\{i_+(\phi), i_-(\phi)\}$ piani iperbolici e una parte anisotropa definita positiva se $i_+(\phi) > i_-(\phi)$, definita negativa se $i_+(\phi) < i_-(\phi)$, per cui

$$W(\phi) = \dim Rad(\phi) + \min\{i_+(\phi), i_-(\phi)\},\ sgn(\phi) = sgn(i_+(\phi) - i_-(\phi))$$

Il teorema spettrale reale

Sia (V, ϕ) uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$ di caratteristica diversa da 2 munito del prodotto scalare ϕ non degenere.
Ricordiamo che l'omomorfismo di rappresentazione $F_\phi : V \to V^*$, $\underline{v} \mapsto \phi(\underline{v}, \cdot)$ è in questo caso un isomorfismo.

Dato $f \in \mathrm{End}(V)$, consideriamo $f^\top \in \mathrm{End}(V^*)$, $f^\top(g) = g \circ f$. Coniugando con l'isomorfismo F_ϕ otteneniamo $f^* = F_\phi^{-1} \circ f^\top \circ F_\phi \in \mathrm{End}(V)$ detto *l'endomorfismo aggiunto* di f (rispetto a ϕ).
Abbiamo il diagramma commutativo

$$\begin{array}{ccc} V & \xrightarrow{f^*} & V \\ {\scriptstyle F_\phi}\downarrow & \circlearrowleft & \downarrow{\scriptstyle F_\phi} \\ V^* & \xrightarrow{f^\top} & V^* \end{array}$$

da cui valutando $F_\phi \circ f^* = f^\top \circ F_\phi$ su $\underline{v} \in V$, $\phi(f^*(\underline{v}), \cdot) = \phi(\underline{v}, \cdot) \circ f \in V^*$, che valutato su $\underline{w} \in V$ dà $\phi(f^*(\underline{v}), \underline{w}) = \phi(\underline{v}, f(\underline{w}))$.
Otteniamo una caratterizzazione alternativa dell'aggiunto:
f^* è l'unico endomorfismo di V tale che, per ogni $\underline{v}, \underline{w} \in V$,

$$\phi(\underline{v}, f(\underline{w})) = \phi(f^*(\underline{v}), \underline{w}).$$

In termini matriciali, se $\mathcal{B}$ è una base di V, per ogni $\underline{v}, \underline{w} \in V$ si ha

$$\phi(\underline{v}, f(\underline{w})) = [\underline{v}]_{\mathcal{B}}^\top M_{\mathcal{B}}(\phi) M_{\mathcal{B}}^{\mathcal{B}}(f) [\underline{w}]_{\mathcal{B}} = \phi(f^*(\underline{v}), \underline{w}) = (M_{\mathcal{B}}^{\mathcal{B}}(f^*)[\underline{v}]_{\mathcal{B}})^\top M_{\mathcal{B}}(\phi)[\underline{w}]_{\mathcal{B}}$$

da cui $M_{\mathcal{B}}(\phi) M_{\mathcal{B}}^{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(f^*)^\top M_{\mathcal{B}}(\phi)$, ovvero

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

che di nuovo mostra l'esistenza e unicità dell'aggiunto.
In particolare, se $M_{\mathcal{B}}(\phi) = I_n$, $M_{\mathcal{B}}^{\mathcal{B}}(f^*) = M_{\mathcal{B}}^{\mathcal{B}}(f)^\top$.

Osserviamo infine che per ogni $f, g \in \mathrm{End}(V)$, $\lambda \in \mathbb{K}$ abbiamo:

- $(f+g)^* = f^* + g^*$, $(\lambda f)^* = \lambda f^*$,
- $(f^*)^* = f$,
- $(id_V)^* = id_V$,
- $(fg)^* = g^* f^*$,
- se f è invertibile, $(f^{-1})^* = (f^*)^{-1}$

(tutte evidenti dalle proprietà della trasposta ecceto la seconda, che però discende immediatamente dalla caratterizzazione alternativa o da quella matriciale). Ovvero, l'aggiunzione $^* : \mathrm{End}(V) \to \mathrm{End}(V)$ è lineare con polinomio minimo $t^2 - 1$ (quindi diagonalizzabile con autospazi relativi a 1 e -1).

Diciamo quindi che $f \in \mathrm{End}(V)$ è *autoaggiunto* se $f^* = f$, *anti-autoaggiunto* se $f^* = -f$.

Ricordiamo l'isomorfismo canonico

$$\chi : \mathrm{End}(V) \to Bil(V \times V^*, \mathbb{K}), \quad \chi(f)(\underline{v}, g) = g(f(\underline{v}))$$

per ogni $f \in \mathrm{End}(V)$, $\underline{v} \in V$, $g \in V^*$.
Tramite l'isomorfismo di rappresentazione F_ϕ, definiamo l'isomorfismo

$$b_\phi : Bil(V \times V^*, \mathbb{K}) \to Bil(V), \quad b_\phi(\psi)(\underline{v}, \underline{w}) = \psi(\underline{v}, F_\phi(\underline{w}))$$

per ogni $\psi \in Bil(V \times V^*, \mathbb{K})$, $\underline{v}, \underline{w} \in V$.
Ci sono diverse verifiche da fare:
• $b_\phi(\psi)$ è bilineare, infatti $\psi(\cdot, F_\phi(\underline{w}))$ e $\psi(\underline{v}, F_\phi(\cdot))$ sono chiaramente lineari, fissati $\underline{v}, \underline{w} \in V$;
• la lineartià di b_ϕ è ovvia;
• Il nucleo di b_ϕ è dato dalle $\psi \in Bil(V \times V^*, \mathbb{K})$ tali che $\psi(\underline{v}, F_\phi(\underline{w})) = 0$ per ogni $\underline{v}, \underline{w} \in V$, ma essendo F_ϕ surgettiva, al variare di $\underline{w} \in V$ $F_\phi(\underline{w})$ varia su tutti i funzionali, per cui $\psi(\underline{v}, g) = 0$ per ogni $\underline{v} \in V, g \in V^*$, ovvero $\psi = 0$;
• infine, b_ϕ è un isomorfismo poiché $\dim Bil(V \times V^*, \mathbb{K}) = \dim Bil(V) = n^2$.

Componendo otteniamo l'isomorfismo $\chi_\phi = b_\phi \circ \chi$, per cui

$$\chi_\phi : \mathrm{End}(V) \to Bil(V), \quad \chi_\phi(f)(\underline{v}, \underline{w}) = \phi(f(\underline{v}), \underline{w})$$

per ogni $f \in \mathrm{End}(V)$, $\underline{v}, \underline{w} \in V$.
Abbiamo allora che $f \in \mathrm{End}(V)$ è autoaggiunto rispetto a ϕ se e solo se $\chi_\phi(f)$ è un prodotto scalare.
In effetti, χ_ϕ rispetta le decomposizioni $\mathrm{End}(V) = V_1(*) \oplus V_{-1}(*)$ (data dagli endomorfismi autoaggiunti e dagli endomorfismi anti-autoaggiunti) e $Bil(V) = PS(V) \oplus A(V)$ (data dai prodotti scalari e dalle applicazioni bilineari antisimmetriche).
La dimostrazione segue immediatamente dalla caratterizzazione dell'aggiunto in termini del prodotto scalare.
Vediamo come ottenerla usando la caratterizzazione matriciale: fissata una base $\mathcal{B}$ di V, possiamo supporre $V = \mathbb{K}^n$, $(\mathbb{K}^n)^* = M(1, n, \mathbb{K})$, $\phi = \Phi_M$ con M simmetrica invertibile, $f = L_A$. Allora per ogni $X \in \mathbb{K}^n$, $R \in M(1, n, \mathbb{K})$ $\chi(f)(R, X) = RMX$, e per ogni $X, Y \in \mathbb{K}^n$, $\chi_\phi(X, Y) = X^\top MAY$.
MA è simmetrica se e solo se $MA = (MA)^\top = A^\top M$ se e solo se $A = M^{-1}A^\top M$ e quest'ultima è la matrice di f^*.

Specializziamo la discussione al caso degli *spazi Euclidei*:
(V, ϕ) si dice spazio Euclideo se V è uno spazio vettoriale su $\mathbb{R}$ e $\phi \in PS(V)$ è definito positivo.
Una base $\mathcal{B}$ di V ortogonale normalizzata per ϕ in questo caso si dice *base ortonormale*. È caratterizzata dal fatto che $M_\mathcal{B}(\phi) = I_n$.
L'isomorfismo di passaggio alle coordinate in una base ortonormale dà una

isometria tra (V, ϕ) e $(\mathbb{R}^n, \Phi_{I_n})$.
Notiamo che le coordinate in una base ortonormale $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ sono date da $[\underline{v}]_\mathcal{B} = \begin{pmatrix} \phi(\underline{v}, \underline{v}_1) \\ \vdots \\ \phi(\underline{v}, \underline{v}_n) \end{pmatrix}$. Infatti, scrivendo $\underline{v} = \sum_{i=1}^{n} \alpha_i \underline{v}_i$ con $\alpha_i \in \mathbb{R}$, si ha $\phi(\underline{v}, \underline{v}_j) = \phi(\alpha_j \underline{v}_j, \underline{v}_j) = \alpha_j \phi(\underline{v}_j, \underline{v}_j) = \alpha_j$.

Dato $f \in \mathrm{End}(V)$, passando in coordinate rispetto ad una base ortonormale $\mathcal{B}$, se $A = M_\mathcal{B}^\mathcal{B}(f)$ e $A^* = M_\mathcal{B}^\mathcal{B}(f^*)$, $A^* = A^\top$, per cui f è autoaggiunto se e solo se $A \in S(n, \mathbb{R})$ è simmetrica.

Definiamo il *gruppo ortogonale reale*

$$O(n, \mathbb{R}) = \{P \in GL(n, \mathbb{R}) |\ P^{-1} = P^\top\}$$

i cui elementi sui dicono *matrici ortogonali*. Notiamo infatti che se P_1, P_2 sono ortogonali, allora $(P_1 P_2)^{-1} = P_2^{-1} P_1^{-1} = P_2^\top P_1^\top = (P_1 P_2)^\top$, per cui $P_1 P_2$ è ortogonale. Inoltre, P_1^{-1} è ortogonale e I_n è ortogonale, quindi $O(n, \mathbb{R})$ è un sottogruppo di $GL(n, \mathbb{R})$.

Se $\mathcal{B}, \mathcal{B}'$ sono due basi ortonormali di V, allora la matrice di cambio di base $P = M_\mathcal{B}^{\mathcal{B}'}(id_V)$ soddisfa a $I_n = P^\top I_n P$, ovvero P è ortogonale.
Viceversa, se $\mathcal{B}$ è una base ortonormale di V e la matrice di cambio di base $M_\mathcal{B}^{\mathcal{B}'}(id_V)$ è ortogonale, allora la base $\mathcal{B}'$ è una base ortonormale di V.

Un endomorfismo $f \in \mathrm{End}(V)$ si dice *ortogonalmente diagonalizzabile* se esiste una base di V ortonormale per ϕ e di autovettori per f.
In termini matriciali, $A \in M(n, \mathbb{R})$ si dice *ortogonalmente diagonalizzabile* se esiste $P \in O(n, \mathbb{R})$ tale che $P^{-1} A P = P^\top A P = D$ è diagonale.
Questo deriva dal fatto che la base canonica è ortonormale per il prodotto scalare standard Φ_{I_n} su $\mathbb{R}^n$ e P è la matrice di cambio di base dalla base canonica ad una base ortonormale che diagonalizza $f = L_A$.

Proposizione
Condizione necesaria affinché $f \in \mathrm{End}(V)$ sia ortogonalmente diagonalizzabile è che f sia autoaggiunta, $f = f^*$.
Condizione necesaria affinché $A \in M(n, \mathbb{R})$ sia ortogonalmente diagonalizzabile è che A sia simmetrica, $A = A^\top$.

Dimostrazione
È chiaro che i due enunciati sono equivalenti passando in un sistema di coordinate date da una base ortonormale.
Dimostriamo il secondo: se $A = PDP^{-1} = PDP^\top$ con D diagonale; allora $A^\top = PD^\top P^\top = PDP^\top = A$. ∎

Possiamo enunciare il teorema spettrale reale che asserisce il viceversa:

Teorema Spettrale Reale

Sia (V, ϕ) uno spazio Euclideo e $f \in \mathrm{End}(V)$ autoaggiunto (rispetto a ϕ). Allora f è ortogonalmente diagonalizzabile.

Dimostrazione
Concludiamo la dimostrazione assumendo il seguente Lemma, che dimostreremo subito dopo:

Lemma
Nelle stesse ipotesi del teorema spettrale reale, il polinomio caratteristico di f è completamente fattorizzabile in $\mathbb{R}[t]$.

Procediamo per induzione su $n = \dim V$. Per $n = 1$ non c'è niente da dimostrare.
Supponiamo $n > 1$. Per il lemma, esiste un autovalore $\lambda \in \mathbb{R}$ di f e fissiamo $\underline{v} \in V$ autovettore di f relativo a λ. Poiché ϕ è definito positivo, $\underline{v}$ non è isotropo e, a meno di normalizzare, possiamo supporre $\phi(\underline{v}, \underline{v}) = 1$. Inoltre, abbiamo la decomposizione $V = \mathrm{Span}(\underline{v}) \oplus \underline{v}^{\perp}$. Notiamo che $\phi_{|_{\underline{v}^{\perp}}}$ rimane definito positivo. Mostriamo che $\underline{v}^{\perp}$ è f-invariante, e quindi $f_{|_{\underline{v}^{\perp}}}$ è un endomorfismo di $\underline{v}^{\perp}$ autoaggiunto rispetto a $\phi_{|_{\underline{v}^{\perp}}}$.
Infatti, dato $\underline{w} \in \underline{v}^{\perp}$, $\phi(\underline{v}, f(\underline{w})) = \phi(f(\underline{v}), \underline{w}) = \phi(\lambda \underline{v}, \underline{w}) = \lambda\phi(\underline{v}, \underline{w}) = 0$, quindi $f(\underline{w}) \in \underline{v}^{\perp}$.
Per ipotesi induttiva, esiste una base $\{\underline{v}_2, \ldots, \underline{v}_n\}$ ortonormale per $(\underline{v}^{\perp}, \phi_{|_{\underline{v}^{\perp}}})$ e di autovettori per $f_{|_{\underline{v}^{\perp}}}$. Allora $\{\underline{v}, \underline{v}_2, \ldots, \underline{v}_n\}$ è una base ortonormale per (V, ϕ) di autovettori per f.

Vediamo un'altra dimostrazione, sempre assumendo il lemma.
Essendo il polinomio caratteristico di f completamente fattorizzabile in $\mathbb{R}[t]$, f è triangolabile e, applicando il processo di ortogonalizzazione di Gram-Schmidt ad una base di V che triangola f, esiste una base $\mathcal{B}$ di V ortonormale per ϕ che triangola f. Consideriamo $M_{\mathcal{B}}^{\mathcal{B}}(f)$, la matrice di f in tale base: essa deve essere triangolare superiore ma anche simmetrica, essendo f autoaggiunto, e quindi è diagonale. ∎

Osserviamo che il teorema spettrale reale si può riformulare nel modo seguente:

Teorema Spettrale Reale
Sia (V, ϕ) uno spazio Euclideo e $f \in \mathrm{End}(V)$ un endomorfismo autoaggiunto. Allora $sp(f)$ è reale e V è la somma diretta ortogonale degli autospazi di f.

Il fatto che due autovettori relativi ad autovalori diversi sono ortogonali si può dimostrare direttamente: $\lambda, \mu \in sp(f)$, $\lambda \neq \mu$, $\underline{v} \in V_\lambda(f), \underline{w} \in V_\mu(f)$, allora $\lambda\phi(\underline{v}, \underline{w}) = \phi(\lambda \underline{v}, \underline{w}) = \phi(f(\underline{v}), \underline{w}) = \phi(\underline{v}, f(\underline{w})) = \phi(\underline{v}, \mu \underline{w}) = \mu\phi(\underline{v}, \underline{w})$. Dunque, $(\lambda - \mu)\phi(\underline{v}, \underline{w}) = 0$ per cui $\phi(\underline{v}, \underline{w}) = 0$.

Dimostriamo adesso il lemma. Usando un sistema di coordinate date da una base ortonormale, basta dimostrare che una matrice simmetrica reale ha poli-

nomio caratteristico completamente fattorizzabile in $\mathbb{R}[t]$.

Usiamo il meccanismo della complessificazione: data $A \in S(n,\mathbb{R})$, usando l'inclusione $M(n,\mathbb{R}) \subset M(n,\mathbb{C})$, pensiamo A come matrice complessa, e scriviamo $A_{\mathbb{C}}$ per indicare A come elemento di $M(n,\mathbb{C})$.
Osserviamo che $p_A = p_{A_{\mathbb{C}}}$ è completamente fattorizzabile in $\mathbb{C}[t]$, per cui dobbiamo dimostrare che $A_{\mathbb{C}}$ ha spettro reale.
Per farlo, complessifichiamo il prodotto scalare standard Φ_{I_n} su $\mathbb{R}^n$, ponendo per $\underline{z},\underline{w} \in \mathbb{C}^n$, $\Phi_{\mathbb{C}}(\underline{z},\underline{w}) = \underline{z}^\top \overline{\underline{w}}$.
Otteniamo $\Phi_{\mathbb{C}} : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ che non è un prodotto scalare, ma un esempio di *prodotto Hermitiano*:

- $\Phi_{\mathbb{C}}$ è $\mathbb{C}$-lineare nel primo argomento,
- $\Phi_{\mathbb{C}}(\underline{z},\underline{w}) = \overline{\Phi_{\mathbb{C}}(\underline{w},\underline{z})}$ per ogni $\underline{z},\underline{w} \in \mathbb{C}^n$,

per cui, $\Phi_{\mathbb{C}}$ è anti-$\mathbb{C}$-lineare nel secondo argomento, ovvero per ogni $\underline{z} \in \mathbb{C}^n$, $\Phi_{\mathbb{C}}(\underline{z},\cdot)$ è additiva e anti-omogenea (cioè $\Phi_{\mathbb{C}}(\underline{z},\lambda\underline{w}) = \overline{\lambda}\Phi_{\mathbb{C}}(\underline{z},\underline{w})\ \forall\ \underline{w} \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$).

Notiamo che se $\underline{z},\underline{w}$ sono reali, allora $\Phi_{\mathbb{C}}(\underline{z},\underline{w}) = \Phi_{I_n}(\underline{z},\underline{w})$, per cui $\Phi_{\mathbb{C}}$ estende a $\mathbb{C}^n$ il prodotto scalare standard su $\mathbb{R}^n$.
Inoltre, per ogni $\underline{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$, $\Phi_{\mathbb{C}}(\underline{z},\underline{z}) = \sum_{i=1}^{n} |z_i|^2$, per cui $\Phi_{\mathbb{C}}$ rimane "definito positivo", cioè se $\underline{z} \neq \underline{0}$, $\Phi_{\mathbb{C}}(\underline{z},\underline{z}) > 0$ (in generale, per un prodotto Hermitiano la forma quadratica è a valori reali).

Sia $\underline{z} \in \mathbb{C}^n$ un autovettore di $A_{\mathbb{C}}$ relativo all'autovalore $\lambda \in \mathbb{C}$, $A_{\mathbb{C}}\underline{z} = A\underline{z} = \lambda\underline{z}$.
Allora

$$\Phi_{\mathbb{C}}(\underline{z}, A\underline{z}) = \Phi_{\mathbb{C}}(\underline{z}, \lambda\underline{z}) = \overline{\lambda}\Phi_{\mathbb{C}}(\underline{z},\underline{z}),$$

mentre

$$\Phi_{\mathbb{C}}(A\underline{z}, \underline{z}) = \Phi_{\mathbb{C}}(\lambda\underline{z}, \underline{z}) = \lambda\Phi_{\mathbb{C}}(\underline{z},\underline{z}).$$

Ma $\Phi_{\mathbb{C}}(\underline{z}, A\underline{z}) = \underline{z}^\top \overline{A\underline{z}} = \underline{z}^\top A \overline{\underline{z}} = \underline{z}^\top A^\top \overline{\underline{z}} = (A\underline{z})^\top \overline{\underline{z}} = \Phi_{\mathbb{C}}(A\underline{z},\underline{z})$.
Quindi $(\lambda - \overline{\lambda})\Phi_{\mathbb{C}}(\underline{z},\underline{z}) = 0$, ed esendo $\underline{z} \neq \underline{0}$, $\lambda = \overline{\lambda}$, cioè $\lambda \in \mathbb{R}$. ❏

Il teorema spettrale reale dà anche il seguente corollario, sulla simultanea "diagonalizzazione" di prodotti scalari:

Corollario
Sia (V,ϕ) uno spazio Euclideo. Per ogni prodotto scalare ψ su V, esiste una base di V simultaneamente ortonormale per ϕ e ortogonale per ψ.

Dimostrazione
Mostriamo una cosa più forte, ovvero che il corollario è equivalente al teorema spettrale reale.
Mostriamo che il teorema spettrale reale implica il corollario.

Scriviamo ψ in modo unico come $\psi = \chi_\phi(f)$ con $f \in \mathrm{End}(V)$ autoaggiunto.
Se $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ è una base ortonormale per ϕ che diagonalizza f, poniamo $f(\underline{v}_i) = \lambda_i \underline{v}_i$, $\lambda_i \in \mathbb{R}$, allora se $i \neq j$, $\psi(\underline{v}_i, \underline{v}_j) = \phi(f(\underline{v}_i), \underline{v}_j) = \lambda_i \phi(\underline{v}_i, \underline{v}_j) = 0$ e quindi $\mathcal{B}$ è una base ortogonale per ψ.
Viceversa, mostriamo che il corollario implica il teorema spettrale reale.
Dato $f \in \mathrm{End}(V)$ autoaggiunto, consideriamo il prodotto scalare $\psi = \chi_\phi(f)$.
Se $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ è una base ortonormale per ϕ e ortogonale per ψ, ovvero $\psi(\underline{v}_i, \underline{v}_j) = 0$ se $i \neq j$, mostriamo che $\mathcal{B}$ è una base di autovettori per f. Per ogni $i = 1 \_ n$, scriviamo $f(\underline{v}_i) = \sum_{j=1}^{n} \alpha_{ij} \underline{v}_j$, $\alpha_{ij} \in \mathbb{R}$, allora se $i \neq j$, $\alpha_{ij} = \phi(f(\underline{v}_i), \underline{v}_j) = \psi(\underline{v}_i, \underline{v}_j) = 0$, e quindi $f(\underline{v}_i) = \alpha_{ii} \underline{v}_i$ come voluto. ❏

Applicando il teorema spettrale reale a $\mathbb{R}^n$ munito del prodotto scalare standard Φ_{I_n}, otteniamo la versione matriciale del teorema spettrale reale e del corollario:

Teorema Spettrale Reale Matriciale
I seguenti fatti sono tra loro equivalenti e veri:

1. Ogni matrice simmetrica reale è ortogonalmente diagonalizzabile: per ogni $A \in S(n, \mathbb{R})$ esiste $P \in O(n, \mathbb{R})$ tale che $P^{-1}AP$ è diagonale;
2. Per ogni prodotto scalare Φ_M su $\mathbb{R}^n$, $M \in S(n, \mathbb{R})$, esiste $P \in O(n, \mathbb{R})$ tale che $P^\top AP$ è diagonale.

❏

In effetti, la versione astratta e la versione matriciale del teorema spettrale reale sono equivalenti. Infatti, la base canonica di $\mathbb{R}^n$ è ortonormale per il prodotto scalare standard Φ_{I_n}. La matrice P^{-1} della versione matriciale è, a seconda dei casi, la matrice del cambio di base dalla base canonica ad una base ortonormale per Φ_{I_n} che diagonalizza l'endomorfismo $f = L_A$, oppure ad una base ortogonale per Φ_M. Fatta questa osservazione, da una parte la versione matriciale è un caso particolare del teorema astratto; dall'altra, passando in un sistema di coordinate date da una base ortonormale per ϕ, per cui ϕ diventa Φ_{I_n}, f diventa A e ψ diventa Φ_M, allora la versione astratta viene tradotta fedelmente per mezzo di quella matriciale.
Osserviamo che nel caso matriciale, l'equivalenza dei due enunciati è evidente: $P^{-1} = P^\top$, ed entrambi gli enunciati considerano un'arbitraria matrice simmetrica (poco importa se in un caso viene interpretata come endomorfismo e nell'altro come prodotto scalare).

Il teorema spettrale Hermitiano

Sia V uno spazio vettoriale su $\mathbb{C}$.
Un *prodotto Hermitiano* su V è una $\phi : V \times V \to \mathbb{C}$ lineare nel primo argomento e tale che $\phi(\underline{v}, \underline{w}) = \overline{\phi(\underline{w}, \underline{v})}$ per ogni $\underline{v}, \_w \in V$.
Segue che ϕ è additiva anche nel secondo argomento, ma non è lineare nel secondo argomento: per ogni $\underline{v}, \underline{w}_1, \underline{w}_2, \underline{w} \in V$, $\lambda \in \mathbb{K}$,

$$\phi(\underline{v}, \underline{w}_1 + \underline{w}_2) = \overline{\phi(\underline{w}_1 + \underline{w}_2, \underline{v})} = \overline{\phi(\underline{w}_1, \underline{v})} + \overline{\phi(\underline{w}_2, \underline{v})} = \phi(\underline{v}, \underline{w}_1) + \phi(\underline{v}, \underline{w}_2),$$

$$\phi(\underline{v}, \lambda \underline{w}) = \overline{\phi(\lambda \underline{w}, \underline{v})} = \overline{\lambda \phi(\underline{w}, \underline{v})} = \overline{\lambda} \phi(\underline{v}, \underline{w}).$$

Si dice che ϕ è *anti-lineare* nel secondo argomento. Le applicazioni da $V \times V$ a $\mathbb{C}$ che sono lineari nel primo argomento e anti-lineari nel secondo sono dette *sesquilineari*. I prodotti Hermitiani sono dunque sesquilineari e "coniugio-simmetrici".
Dalla proprietà di coniugio-simmetria segue che $\phi(\underline{v}, \underline{v}) \in \mathbb{R}$ per ogni $\underline{v} \in V$, ovvero che la forma quadratica associata a ϕ, $q_\phi(\underline{v}) = \phi(\underline{v}, \underline{v})$, è a valori reali.

Analogamente ai prodotti scalari, fissata una base $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ di V un prodotto Hermitiano ϕ si rappresenta tramite la matrice quadrata $H \in M(n, \mathbb{C})$, indicata con $M_{\mathcal{B}}(\phi)$, il cui elemento di posto (i, j) è $\underset{i}{\underline{H}}^{j} = \phi(\underline{v}_i, \underline{v}_j)$, per cui $\phi(\underline{v}, \underline{w}) = [\underline{v}]_{\mathcal{B}}^{\top} H \overline{[\underline{w}]_{\mathcal{B}}}$ per ogni $\underline{v}, \underline{w} \in V$. Notiamo che la proprietà di coniugio-simmetria si traduce in $H = \overline{H}^{\top}$. Tali matrici si dicono Hermitiane e formano un sottospazio di $M(n, \mathbb{C})$.
Cambiando base, la matrice H del prodotto Hermitiano cambia in $M^{\top} H \overline{M}$, dove M è la matrice del cambio di base.

Se $V = \mathbb{C}^n$, i prodotti Hermitiani sono tutti e soli del tipo $\phi = \Phi_M$ con $M \in M(n, \mathbb{C})$ Hermitiana, dove $\Phi_M(\underline{x}, \underline{y}) = \underline{x}^{\top} M \overline{\underline{y}}$ per ogni $\underline{x}, \underline{y} \in \mathbb{C}^n$. Abbiamo già incontrato il prodoto Hermitiano standard Φ_{I_n} nella dimostrazione del teorema spettrale reale.

Lo studio dei prodotti Hermitiani ricalca parola per parola quello dei prodotti scalari reali: sono ben definite le nozioni di radicale, cono isotropo, rango, non degenericità, ortogonalità, isometria ed esistono basi ortogonali (anche normalizzate). In particolare, ha senso dire se un prodotto Hermitiano è definito positivo/negativo e possiamo definire la segnatura per un prodotto Hermitiano che risulta essere un invariante completo per l'isometria di spazi vetoriali comlpessi muniti di un prodotto Hermitiano (lasciamo al lettore la formulazione delle definizioni e la verifica delle dimostrazioni).

Restringiamoci al caso in cui V sia munito di un prodotto Hermitiano definito positivo (anche in questo caso si parla di spazio Euclideo).
L'analogo delle basi ortonormali sono le basi *unitarie*; l'analogo del gruppo ortogonale $O(V, \phi)$ è il gruppo *unitario* $U(V, \phi)$; il gruppo unitario matriciale

classico è $U(n,\mathbb{C}) = U(\mathbb{C}^n, \Phi_{I_n}) = \{P \in GL(n,\mathbb{C})|\ P^{-1} = \overline{P}^\top\}$.
Osserviamo che $O(n,\mathbb{R}) \subset U(n,\mathbb{C})$ e che le matrici di cambio di base tra basi unitarie sono unitarie.

Dato $f \in \mathrm{End}(V)$, l'aggiunto di f, $f^* \in \mathrm{End}(V)$, è definito dalla formula $\phi(\underline{v}, f(\underline{w})) = \phi(f^*(\underline{v}), \underline{w})$ per ogni $\underline{v}, \underline{w} \in V$. Nel caso $V = \mathbb{C}^n$, $\phi = \Phi_{I_n}$, $f = L_A$, allora f^* è dato dalla matrice $\overline{A}^\top$. In generale, data $A \in M(n,\mathbb{C})$, $\overline{A}^\top$ si dice *matrice aggiunta* di A e si indica con A^*. Lo stesso vale per le matrici di f e f^* in una base unitaria $\mathcal{B}$: $M_{\mathcal{B}}^{\mathcal{B}}(f^*) = \overline{M_{\mathcal{B}}^{\mathcal{B}}(f^*)}^\top$.

Analogamente al caso reale, vogliamo caratterizzare gli endomorfismi e le matrici *unitariamente diagonalizzabili*, cioè diagonalizzati da una base unitaria i primi, diagonalizzate tramite matrice unitaria le seconde.

Diciamo che un endomorfismo $f \in \mathrm{End}(V)$ è *normale* se commuta con il suo aggiunto: $ff^* = f^*f$.
Analogamente, $A \in M(n,\mathbb{C})$ è normale se $AA^* = A^*A$.

Ci sono tre inportanti classi di endomorfismi normali:
- gli endomorfimi *autoaggiunti*, per cui $f^* = f$;
- gli endomorfimi *anti-autoaggiunti*, per cui $f^* = -f$;
- gli endomorfimi *unitari*, per cui $f^* = f^{-1}$;
Gli analoghi matriciali sono:
- le matrici *Hermitiane*, per cui $A^* = A$;
- le matrici *anti-Hermitiane*, per cui $A^* = -A$;
- le matrici *unitarie*, per cui $A^* = A^{-1}$.

Proposizione
Condizione necessaria perché $f \in \mathrm{End}(V)$ o $A \in M(n,\mathbb{C})$ sia unitariamente diagonalizzabile è che sia normale.

Dimostrazione
Passando in coordinate rispetto ad una base unitaria che diagonalizza f, se la matrice di f in tale base è la matrice diagonale D, allora la matrice di f^* in tale base è $D^* = \overline{D}^\top$ che è ancora diagonale. Si conclude osservando che due matrici diagonali commutano.
In modo analogo, se per $U \in U(n,\mathbb{C})$ si ha che $U^{-1}AU = D$ è diagonale, allora passando alle matrici aggiunte $D^* = \overline{U^{-1}AU}^\top = \overline{U}^\top \overline{A}^\top \overline{U^{-1}}^\top = U^{-1}A^*U$.
Quindi, osservando che D^* è diagonale e commuta con D,
$AA^* = (UDU^{-1})(UD^*U^{-1}) = UDD^*U^{-1} = UD^*DU^{-1} = A^*A$. ∎

Analogamente al caso reale, il teorema spettrale Hermitiano ci mostra che vale anche il viceversa:

Teorema Spettrale Hermitiano
Sia V uno spazio vettoriale su $\mathbb{C}$ di dimensione finita munito di un prodotto

Hermitiano definito positivo. Se $f \in \mathrm{End}(V)$ è normale, allora è unitariamente diagonalizzabile.

Dimostrazione
Mostriamo che se $\lambda \in \mathbb{C}$ è un autovalore di f, allora $\overline{\lambda}$ è un autovalore di f^* e i relativi autospazi coincidono, $V_\lambda(f) = V_{\overline{\lambda}}(f^*)$.
Per la simmetria del problema, basta mostrare che $V_\lambda(f) \subset V_{\overline{\lambda}}(f^*)$.
Sia dunque $\underline{v} \in V$ un autovettore per f relativo a λ. Poiché ϕ è definito positivo, $f^*(\underline{v}) = \overline{\lambda}\underline{v}$ se e solo se $q_\phi(f^*(\underline{v}) - \overline{\lambda}\underline{v}) = 0$.
Poiché f e f^* commutano, $V_\lambda(f)$ è f^*-invariante, quindi $f^*(\underline{v}) - \overline{\lambda}\underline{v} \in V_\lambda(f)$.
D'altra parte, per ogni $\underline{w} \in V_\lambda(f)$,

$$\phi(f^*(\underline{v}), \underline{w}) = \phi(\underline{v}, f(\underline{w})) = \phi(\underline{v}, \lambda\underline{w}) = \overline{\lambda}\phi(\underline{v}, \underline{w}) = \phi(\overline{\lambda}\underline{v}, \underline{w}),$$

per cui $\phi(f^*(\underline{v}) - \overline{\lambda}\underline{v}, \underline{w}) = 0$.
La dimostrazione si conclude per induzione sulla dimensione di V analogamente alla dimostrazione nel caso reale: si fissa un autovettore $\underline{v}$ di f tale che $\phi(\underline{v}, \underline{v}) = 1$ e si mostra che le ipotesi valgono per le restrizioni a $\underline{v}^\perp$; $\underline{v}$ unito ad una base unitaria di $\underline{v}^\perp$ che diagonalizza la restrizione di f è una base unitaria di V che diagonalizza f.
L'unica cosa da vedere è che $\underline{v}^\perp$ è f-invariante e f^*-invariante. Per quanto discusso sopra, e per la simmetria del problema, basta vedere che $\underline{v}^\perp$ è f^*-invariante.
Sia allora $\underline{w} \in \underline{v}^\perp$ e mostriamo che $f^*(\underline{w}) \perp \underline{v}$:
$\phi(\underline{v}, f^*(\underline{w})) = \phi(f(\underline{v}), \underline{w}) = f(\lambda\underline{v}, \underline{w}) = \lambda\phi(\underline{v}, \underline{w}) = 0$ come voluto. ❏

Anche in questo caso abbiamo una dimostrazione alternativa del teorema spettrale Hermitiano, che si basa sul fatto che anche per i prodotti Hermitiani definiti positivi possiamo usare il processo di ortogonalizzazione di Gram-Schmidt per produrre una base unitaria da una base qualsiasi, senza alterarne la bandiera. La dimostrazione funziona allo stesso modo in cui funzionava per i prodotti scalari reali. Come conseguenza, ogni endomorfismo complesso è triangolabile in una base unitaria.
Dato $f \in \mathrm{End}(V)$ normale, sia allora $\mathcal{B}$ una base unitaria che triangola f e sia $T = M_{\mathcal{B}}^{\mathcal{B}}(f)$. In tale base, $M_{\mathcal{B}}^{\mathcal{B}}(f^*) = T^* = \overline{T}^\top$. Da $ff^* = f^*f$ otteniamo $TT^* = T^*T$. Se la prima riga di T è $(z_1\, z_2 \dots z_n)$, $z_i \in \mathbb{C}$, allora l'elemento di posto (1,1) di TT^* è $\sum_{i=1}^{n} |z_i|^2$, mentre l'elemento di posto (1,1) di T^*T è $|z_1|^2$, per cui $z_i = 0 \ \forall\ i = 2 \dots n$. Allo stesso modo si ragiona per le altre righe di T, trovando che $T_i^j = 0$ se $i \neq j$, ovvero, T è diagonale e la base unitaria $\mathcal{B}$ diagonalizza f.

Come corollario, otteniamo una caratterizzazione spettrale delle tre classi di endomorfismi normali individuate sopra:

Corollario

Sia $f \in \mathrm{End}(V)$ normale. Allora,
- f è autoaggiunto se e solo se ha spettro reale.
- f è anti-autoaggiunto se e solo se ha spettro immaginario puro.
- f è unitario se e solo se ha spettro unitario (contenuto nella circonferenza unitaria dei numeri complessi di modulo 1).

Dimostrazione
Essendo normale, f è unitariamente diagonalizzabile, e quindi esiste una base unitaria in cui è rappresentato dalla matrice diagonale D (che ha sulla diagonale gli autovalori λ_i di f). In questa base, f^* è rappresentato dalla matrice diagonale $\overline{D}^\top = \overline{D}$.
Si ha allora:
- $f^* = f \iff \overline{D} = D \iff D$ è reale ($\lambda_i \in \mathbb{R}$);
- $f^* = -f \iff \overline{D} = -D \iff D$ è immaginaria pura ($\lambda_i \in i\mathbb{R}$);
- $ff^* = id_V \iff D\overline{D} = I_n \iff D$ è unitaria ($\lambda_i\overline{\lambda_i} = 1$). ❏

Osserviamo che segue dal teorema che gli autospazi di un endomorfismo f normale sono in somma diretta ortogonale.
È facile vedere ditrettamente che se $\underline{v} \in V_\lambda(f)$, $\underline{w} \in V_\mu(f) = V_{\overline{\mu}}(f^*)$, con $\lambda \neq \mu$, allora $\underline{v} \perp \underline{w}$:

$$\begin{aligned}\lambda\phi(\underline{v},\underline{w}) &= \phi(\lambda\underline{v},\underline{w}) = \phi(f(\underline{v}),\underline{w}) = \\ &= \phi(\underline{v}, f^*(\underline{w})) = \phi(\underline{v},\overline{\mu}\underline{w}) = \mu\phi(\underline{v},\underline{w}).\end{aligned}$$

Quindi, $(\lambda-\mu)\phi(\underline{v},\underline{w}) = 0$, e dunque $\phi(\underline{v},\underline{w}) = 0$.

Applicando il teorema spettrale Hermitiano a $\mathbb{C}^n$ munito del prodotto Hermitiano standard Φ_{I_n}, otteniamo la versione matriciale del teorema spettrale Hermitiano:

Teorema Spettrale Hermitiano Matriciale
Ogni matrice normale complessa è unitariamente diagonalizzabile: per ogni $A \in M(n,\mathbb{C})$ tale che $AA^* = A^*A$, esiste $U \in U(n,\mathbb{C})$ tale che $U^{-1}AU$ è diagonale.

Abbiamo anche un analogo del corollario sulla simultaena diagonalizzazione di prodotti scalari (di cui diamo una dimostrazione diversa dal caso reale: in effetti entrambe le dimostrazioni funzionano in entrambi i casi).

Corollario
Sia (V,ϕ) con ϕ prodotto Hermitiano definito positivo.
Per ogni prodotto Hermitiano ψ su V, esiste una base di V simultaneamente unitaria per ϕ e ortogonale per ψ.

Dimostrazione
Sia $\mathcal{B}$ una base di V unitaria per ϕ e sia $H = M_{\mathcal{B}}(\psi)$. H è Hermitiana, quindi normale ed esiste $U \in U(n,\mathbb{C})$ tale che $U^{-1}HU = \overline{U}^T HU = D$ diagonale. In-

terpretando $\overline{U}$ (che è ancora unitaria) come matrice del cambio di base da $\mathcal{B}$ alla base di V $\mathcal{B}'$, $\mathcal{B}'$ è ancora una base unitaria per ϕ e ortogonale per ψ. ❑

Geometria Affine

Sia V uno spazio vettoriale sul campo $\mathbb{K}$.
Un insieme non vuoto $\mathbb{A}$ si dice *spazio affine* su V se esiste

$$+ : \mathbb{A} \times V \to \mathbb{A}, \quad (P, \underline{v}) \mapsto P + \underline{v}$$

tale che:

- per ogni $P \in \mathbb{A}$, $\underline{v}, \underline{w} \in V$, $(P + \underline{v}) + \underline{w} = P + (\underline{v} + \underline{w})$;
- per ogni $P, Q \in \mathbb{A}$, esiste un unico $\underline{v} \in V$ tale che $Q = P + \underline{v}$.

Gli elementi di uno spazio affine si dicono *punti* e l'applicazione $+$ si dice *somma punto/vettore*. Si pone $\dim A = \dim V$.

Mentre la prima proprietà della somma punto/vettore è una sorta di associatività, la seconda permette di definire $vec : \mathbb{A} \times \mathbb{A} \to V$ la cui restrizione $vec|_{\{P\}\times\mathbb{A}} : \{P\} \times \mathbb{A} \to V$ è biunivoca per ogni $P \in \mathbb{A}$. Infatti, dato $\underline{v} \in V$, $vec(P, P + \underline{v}) = \underline{v}$; se poi $Q, Q' \in \mathbb{A}$ sono tali che $vec(P, Q) = vec(P, Q') = \underline{v}$, allora $Q = P + \underline{v} = Q'$.
L'immagine tramite vec della coppia (P, Q) si dice il *vettore che congiunge P a Q* e come notazione useremo $vec(P, Q) = \overrightarrow{PQ}$ (detta notazione vettoriale) o $vec(P, Q) = Q - P$ (detta notazione aritmetica). Quindi, per ogni $P, Q \in \mathbb{A}$, $Q = P + \overrightarrow{PQ} = P + Q - P$.

Osserviamo che $\mathbb{A} = V$ è uno spazio affine su V, dove la somma punto/vettore è l'usuale somma. In questo caso, dati $P, Q \in V$, $vec(P, Q) = \overrightarrow{PQ} = Q - P$ è l'usuale differenza tra vettori, non solo una notazione. Lo spazio affine così ottenuto si indica con $\mathbb{A}(V)$. Nel caso di $V = \mathbb{K}^n$, si scrive $\mathbb{A}(\mathbb{K}^n) = \mathbb{A}^n$, detto *spazio affine standard*.

Vediamo le prime proprietà di questa struttura.

• Per ogni $P \in \mathbb{A}$, $\overrightarrow{PP} = \underline{0}$, ovvero $P + \underline{0} = P$.
Infatti,
$P + \underline{0} = (P + \overrightarrow{PP}) + \underline{0} = P + (\overrightarrow{PP} + \underline{0}) = P + \overrightarrow{PP} = P$.
Notiamo che questo implica che dato $\underline{v} \in V$, se esiste $P \in \mathbb{A}$ tale che $P + \underline{v} = P$, allora $\underline{v} = \underline{0}$.

• $\forall\ P, Q \in \mathbb{A}$, $\overrightarrow{QP} = -\overrightarrow{PQ}$.
Infatti,
$Q + (-\overrightarrow{PQ}) = P + \overrightarrow{PQ} + (-\overrightarrow{PQ}) = P + (\overrightarrow{PQ} - \overrightarrow{PQ}) = P + \underline{0} = P$.

• Per ogni $P, Q, R \in \mathbb{A}$, $\overrightarrow{PQ} + \overrightarrow{QR} + \overrightarrow{RP} = \underline{0}$.
Infatti,
$P + (\overrightarrow{PQ} + \overrightarrow{QR} + \overrightarrow{RP}) = (P + \overrightarrow{PQ}) + (\overrightarrow{QR} + \overrightarrow{RP}) = Q + (\overrightarrow{QR} + \overrightarrow{RP}) =$

$= (Q + \overrightarrow{QR}) + \overrightarrow{RP} = R + \overrightarrow{RP} = P.$

Questa proprietà viene detta *proprietà del triangolo.*

Usando la notazione aritmetica:
• $P - P = \underline{0}$,
• $P - Q = -(Q - P)$,
• $(Q - P) + (R - Q) + (P - R) = \underline{0}$,
coerentemente con le usuali regole aritmetiche.

Notiamo che la proprietà del triangolo può essere riscritta, nelle due notazioni, come: $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$, $(Q - P) + (R - Q) = R - P$.

Fissato $P \in \mathbb{A}$, restringendo $+$ a $\{P\} \times V$ otteniamo $S_P : V \to \mathbb{A}$, $S_P(\underline{v}) = P + \underline{v}$, biunivoca con inversa $F_P : \mathbb{A} \to V$, $F_P(Q) = \overrightarrow{PQ}$ (F_P è la restrizione di vec a $\{P\} \times \mathbb{A}$, composta con l'inclusione $\mathbb{A} \hookrightarrow \{P\} \times \mathbb{A}$).
Possiamo definire su $\mathbb{A}$ una struttura di spazio vettoriale su $\mathbb{K}$ con "origine" in $P = S_P(\underline{0})$, $(\mathbb{A}_P, +_P, \cdot_P)$, in modo che $S_P : V \to \mathbb{A}_P$ sia un isomorfismo:
per ogni $Q_1, Q_2 \in \mathbb{A}, \lambda \in \mathbb{K}$,

$$Q_1 +_P Q_2 = S_P(F_P(Q_1) + F_P(Q_2)) = P + \overrightarrow{PQ_1} + \overrightarrow{PQ_2} = Q_1 + \overrightarrow{PQ_2} = Q_2 + \overrightarrow{PQ_1},$$

$$\lambda \cdot_P Q_1 = S_P(\lambda F_P(Q_1)) = P + \lambda\overrightarrow{PQ_1}.$$

Estendendo la notazione aritmetica scriviamo:

$$Q_1 +_P Q_2 = P + Q_1 - P + Q_2 - P = Q_1 + Q_2 - P = Q_2 + Q_1 - P,$$

$$\lambda \cdot_P Q_1 = P + \lambda(Q_1 - P) = \lambda Q_1 + (1 - \lambda)P.$$

In questo senso, possiamo pensare a $\underline{v} = vec(P, Q)$ come al vettore $\underline{v}$ "applicato" in P o "uscente" da P.

Fissato $\underline{v} \in V$, restringendo $+$ a $\mathbb{A} \times \{\underline{v}\}$ otteniamo $\tau_{\underline{v}} : \mathbb{A} \to \mathbb{A}$, $\tau_{\underline{v}}(P) = P + \underline{v}$, biunivoca, detta *traslazione di vettore* $\underline{v}$.
Per ogni $\underline{v}, \underline{w} \in V$, $\tau_{\underline{v}} \circ \tau_{\underline{w}} = \tau_{\underline{w}} \circ \tau_{\underline{v}} = \tau_{\underline{v}+\underline{w}}$, $\tau_{\underline{0}} = id_{\mathbb{A}}$, $\tau_{\underline{v}}^{-1} = \tau_{-\underline{v}}$. Ovvero, l'insiemte $T(\mathbb{A})$ delle traslazioni di $\mathbb{A}$ dotato della composizione è un gruppo abeliano isomorfo a $(V, +)$, sottogruppo di $S(\mathbb{A})$.

Dato $P \in \mathbb{A}$, $(\tau_{\underline{v}} \circ S_P)(\underline{u}) = \tau_{\underline{v}}(P + \underline{u}) = P + \underline{u} + \underline{v} = S_{P+\underline{v}}(\underline{u})$ $\forall \underline{u} \in V$, quindi $\tau_{\underline{v}} = S_{P+\underline{v}} \circ S_P^{-1}$ è un isomorfismo da $\mathbb{A}_P$ a $\mathbb{A}_{P+\underline{v}}$.
Osserviamo che, ponendo $Q = P + \underline{v}$, abbiamo $S_Q \circ S_P^{-1} = \tau_{\overrightarrow{PQ}}$

Avendo la possibilità di definire molteplici strutture di spazio vettoriale su uno spazio affine, possiamo estendere agli spazi affini tutte le definizioni viste per gli spazi vettoriali, a patto che non dipendano dalla struttura usata per definirle. In particolare, vogliamo estendere i concetti di combinazione lineare, sottospazio vettoriale, applicazione lineare, base.

Iniziamo con estendere la nozione di combinazione lineare.

In generale, la stessa combinazione lineare di punti di $\mathbb{A}$ vista nei vari spazi vettoriali $(\mathbb{A}_P, +_P, \cdot_P)$ dà risultati differenti. Ad esempio, fissati due punti $Q_1, Q_2 \in \mathbb{A}$, si può ottenere qualsiasi punto di $\mathbb{A}$ come loro somma in un qualche $\mathbb{A}_P$: se $Q \in \mathbb{A}$, $Q_1 +_{Q_1+_Q Q_2} Q_2 = Q$.
Esistono però delle particolari combinazioni lineari il cui risultato è lo stesso indipendentemente da quale struttura di spazio vettoriale $(\mathbb{A}_P, +_P, \cdot_P)$ venga usata per eseguirle.
Tali combinazioni lineari si dicono *combinazioni affini* e sono caratterizate dalla seguente proprietà:

Proposizione
Dati $Q_1, \dots, Q_m \in \mathbb{A}$ e $\lambda_1, \dots, \lambda_m \in \mathbb{K}$, la combinazione lineare in $(\mathbb{A}_P, +_P, \cdot_P)$, $\lambda_1 \cdot_P Q_1 +_P \cdots +_P \lambda_m \cdot_P Q_m$ non dipende da $P \iff \sum_{i=1}^{m} \lambda_i = 1$.

Dimostrazione
Fissati $P_0 \neq P_1 \in \mathbb{A}$, $\overrightarrow{P_0P_1} \neq \underline{0}$, $P_1 = P_0 + \overrightarrow{P_0P_1}$ e usando la proprietà del triangolo, $\overrightarrow{P_1Q_i} = \overrightarrow{P_1P_0} + \overrightarrow{P_0Q_i}$. Allora
$P_1 + \sum_{i=1}^{m} \lambda_i \overrightarrow{P_1Q_i} = P_0 + \overrightarrow{P_0P_1} + \sum_{i=1}^{m} \lambda_i (\overrightarrow{P_1P_0} + \overrightarrow{P_0Q_i}) = P_0 + \sum_{i=1}^{m} \lambda_i \overrightarrow{P_0Q_i} \iff$
$\overrightarrow{P_0P_1} + \sum_{i=1}^{m} \lambda_i \overrightarrow{P_1P_0} = \underline{0} \iff (1 - \sum_{i=1}^{m} \lambda_i)\overrightarrow{P_0P_1} = \underline{0} \iff \sum_{i=1}^{m} \lambda_i = 1.$ ❏

La combinazione affine dei punti $Q_1, \dots, Q_m \in \mathbb{A}$ e coefficienti $\lambda_1, \dots, \lambda_m \in \mathbb{K}$, $\sum_{i=1}^{m} \lambda_i = 1$, si indica, estendendo la notazione aritmetica, con $\lambda_1 Q_1 + \cdots + \lambda_m Q_m$.

Di nuovo, la notazione è compatibile con le usuali regole aritmetiche di somma e prodotto.
Ad esempio, se $Q = \lambda_1 Q_1 + \cdots + \lambda_m Q_m$, $\sum_{i=1}^{m} \lambda_i = 1$ con $\lambda_1 \neq 0$, allora $Q_1 = \frac{1}{\lambda_1} Q - \frac{\lambda_2}{\lambda_1} Q_2 - \cdots - \frac{\lambda_m}{\lambda_1} Q_m$ (che è una combinazione affine poiché $\frac{1}{\lambda_1} - \sum_{i=2}^{m} \frac{\lambda_i}{\lambda_1} = 1$; notare che invece non ha senso scrivere $\lambda_1 Q_1 = Q - \lambda_2 Q_2 - \cdots - \lambda_m Q_m$, a meno che $\lambda_1 = 1$).
Infatti, usando la struttura di $\mathbb{A}_{Q_1}$ per eseguire le combinazioni affini,
$Q = Q_1 + \sum_{i=2}^{m} \lambda_i \overrightarrow{Q_1Q_i}$, per cui $\overrightarrow{Q_1Q} = \sum_{i=2}^{m} \lambda_i \overrightarrow{Q_1Q_i}$, e quindi
$Q_1 + \frac{1}{\lambda_1} \overrightarrow{Q_1Q} - \sum_{i=2}^{m} \frac{\lambda_i}{\lambda_1} \overrightarrow{Q_1Q_i} = Q_1 + \frac{1}{\lambda_1} \sum_{i=2}^{m} \lambda_i \overrightarrow{QQ_1} - \sum_{i=2}^{m} \frac{\lambda_i}{\lambda_1} \overrightarrow{QQ_i} = Q_1 + \underline{0} = Q_1.$

Osserviamo che le definizioni di somma e prodotto per scalari in $\mathbb{A}_P$, sono in effetti date da combinazioni affini, e che in generale, una combinazione lineare in $\mathbb{A}_P$ dei punti $Q_1, \dots, Q_m \in \mathbb{A}$ è una combinazione affine di $P, Q_1, \dots, Q_m$:

$$\lambda_1 \cdot_P Q_1 +_P \cdots +_P \lambda_m \cdot_P Q_m = \lambda_1 Q_1 + \cdots + \lambda_m Q_m + (1 - \sum_{i=1}^{m} \lambda_i) P.$$

Infine, osserviamo che è sempre possibile inserire in una combinazione affine un punto con coefficiente nullo senza alterare il risultato.

Estendiamo adesso la nozione di sottospazio vettoriale.

Definizione:
$E \subset \mathbb{A}$ si dice *sottospazio affine* se è chiuso per combinazioni affini dei suoi punti. Notare che ammettiamo anche il caso $E = \emptyset$.

Notiamo che l'intersezione di un numero arbitratio di sottospazi affini è un sottospazio affine (possibilmente vuoto).

Se E è un sottospazio affine non vuoto, fissato $P \in E$, poiché le combinazioni lienari in $\mathbb{A}_P$ di punti di E sono combinazioni affini di punti di E e di P, E è chiuso per combinazioni lienari in $\mathbb{A}_P$ ed è quindi un sottospazio vettoriale di $\mathbb{A}_P$.
Viceversa, un sottospazio vettoriale F di un qualche $\mathbb{A}_Q$ è un sottospazio affine (ed è quindi un sottospazio vettoriale di ogni $\mathbb{A}_P$ con $P \in F$). Infatti, usando la struttura di $\mathbb{A}_Q$ per eseguire le combinazioni affini, queste diventano combinazioni lineari in $\mathbb{A}_Q$.

Se E è un sottospazio affine non vuoto, fissato $P \in E$, usiamo l'isomorfismo $S_P : V \to \mathbb{A}_P$, $\underline{v} \mapsto P + \underline{v}$ ed otteniamo $S_P^{-1}(E) = \{\overrightarrow{PQ} |\ Q \in E\}$ sottospazio vettoriale di V.

Proposizione
Nelle ipotesi sopra, $S_P^{-1}(E)$ non dipende da P.

Dimostrazione
Dati $P_1, P_2 \in E$, $S_{P_1}^{-1}(E) = S_{P_2}^{-1}(E) \iff S_{P_1}(S_{P_2}^{-1}(E)) = E$. Ricordando che $S_{P_1} \circ S_{P_2}^{-1} = \tau_{P_1 - P_2}$, basta dimostrare che $\tau_{P_1-P_2}(E) = E$ e, ragionando in modo analogo scambiando P_1 con P_2 e usando $\tau_{P_2-P_1} = (\tau_{P_1-P_2})^{-1}$, basta dimostrare che $\tau_{P_1-P_2}(E) \subset E$.
Se dunque $Q \in E$, $\tau_{P_1-P_2}(Q) = Q + P_1 - P_2 \in E$ essendo una combinazione affine di punti di E. ∎

Il sottospazio vettoriale $S_P^{-1}(E) \subset V$ è detto *giacitura* di E (o *direzione* di E o *spazio tangente* ad E) ed è indicato con $Giac(E)$. Per completezza, poniamo $Giac(\emptyset) = \emptyset$.
Se E è non vuoto, poniamo $\dim E = \dim Giac(E)$.

Viceversa, dato $W \subset V$ un sottospazio vettoriale, e dato $P \in \mathbb{A}$, allora il sottospazio vettoriale di $\mathbb{A}_P$ $S_P(W) = \{P + \underline{w} |\ \underline{w} \in W\} = P + W$ è un sottospazio affine di giacitura W "passante" per P.
In effetti $P + W$ è l'unico sottospazio affine di giacitura W passante per P: due tali sottospazi affini coincidono con $S_P(W)$. Notiamo che se $E, F \subset \mathbb{A}$ sono sottospazi affini tali che $E \cap F \neq \emptyset$ e $Giac(E) = Giac(F) = W$, allora $E = F$,

in quanto entrambi coincidono con $P+W$, con $P \in E \cap F$.

Osserviamo che per ogni $P_1 \in E$,
$P_1 + W = S_{P_1}(W) = (\tau_{P_1-P} \circ S_P)(W) = \tau_{P_1-P}(E) = E$.
Quindi, dotato della restrizione di $+$ a $E \times Giac(E)$, E è spazio affine su $Giac(E)$.

Dato $X \subset \mathbb{A}$ non vuoto, il *sottospazio affine generato da* X è il più piccolo sottospazio affine di $\mathbb{A}$ che contiene X

$$\mathrm{Span}_a(X) = \bigcap_{\substack{X \subset E \subset \mathbb{A} \\ \text{ssp.affine}}} E.$$

È immediato verificare che $\mathrm{Span}_a(X) = \mathrm{Span}_{\mathbb{A}_P}(X)$ per ogni $P \in X$ e che $\mathrm{Span}_a(X) = \{\text{combinazioni affini (finite) di punti (distinti) di X}\}$.

Viceversa, se $E \subset \mathbb{A}$ è un sottospazio affine non vuoto, fissato $P \in E$, sia X un insieme di generatori per E considerato come sottospazio vettoriale di $\mathbb{A}_P$. Allora $E = \mathrm{Span}_a(X)$.

Ricapitolando, $E \subset \mathbb{A}$ è un sottospazio affine non vuoto se e solo se:
- esiste $P \in \mathbb{A}$ tale che E è uno spazio vettoriale in $\mathbb{A}_P$;
- oppure, esistono $W \subset V$ sottospazio e $P \in \mathbb{A}$ tale che $E = P + W$;
- oppure, esiste $X \subset \mathbb{A}$ tale che $E = \mathrm{Span}_a(X)$.

Inoltre:
- $E = P + Giac(E)$ per ogni $P \in E$.
- Possiamo descrivere la giacitura in modo più intrinseco indipendente da P:
$Giac(E) = \{\overrightarrow{PQ} |\ P, Q \in E\}$.
Infatti $\{\overrightarrow{PQ} |\ P, Q \in E\} = \bigcup_{P \in E} S_P^{-1}(E) = \bigcup_{P \in E} Giac(E) = Giac(E)$.

Siano $E, F \subset \mathbb{A}$ sottospazi affini non vuoti.
Abbiamo già osservato che se $Giac(E) = Giac(F) = W$ allora $E \cap F = \emptyset$ oppure $E = F = P + W$ con $P \in E \cap F$.
Inoltre, $E \cap F$ è un sottospazio affine e se $E \subset F$ allora $Giac(E) \subset Giac(F)$.

Se $E \cap F \neq \emptyset$, allora $Giac(E \cap F) = Giac(E) \cap Giac(F)$.
Infatti, dato $P \in E \cap F$, scriviamo $E = P + Giac(E)$, $F = P + Giac(F)$. Allora $Q \in E \cap F \iff Q = P + \underline{v} = P + \underline{w}$ con $\underline{v} \in Giac(E), \underline{w} \in Giac(F)$ (che implica $\underline{v} = \underline{w}$) $\iff Q = P + \underline{u}$ con $\underline{u} \in Giac(E) \cap Giac(F)$. Quindi, $E \cap F = P + Giac(E) \cap Giac(F)$.

Dati $P \in E, Q \in F$ allora
$E \cap F = \emptyset \iff \overrightarrow{PQ} \notin Giac(E) + Giac(F)$.
Infatti, se fosse $\overrightarrow{PQ} = \underline{v} + \underline{w}$ con $\underline{v} \in Giac(E)$, $\underline{w} \in Giac(F)$, allora $E \ni P + \underline{v} = P + (\overrightarrow{PQ} - \underline{w}) = (P + \overrightarrow{PQ}) - \underline{w} = Q - \underline{w} \in F$ ↯.
Viceversa, se fosse $R \in E \cap F$, $\overrightarrow{PR} \in Giac(E)$, $\overrightarrow{RQ} \in Giac(F)$, e quindi

$\overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ} \in Giac(E) + Giac(F)$ ↯.

Osserviamo che se $Giac(E) + Giac(F) = V$, allora dati $P \in E, Q \in F$ necessariamente $\overrightarrow{PQ} \in Giac(E) + Giac(F)$ e quindi $E \cap F \neq \emptyset$.

In generale, $E \cup F$ non è un sottospazio affine (lo è se e solo se $E \subset F$ o $F \subset E$), per cui definiamo la loro *somma affine* come $E + F = \mathrm{Span}_a(E \cup F)$.

Mostriamo che $Giac(E + F) = Giac(E) + Giac(F) + \mathrm{Span}(\overrightarrow{PQ})$, con $P \in E$, $Q \in F$.
Infatti, poiché $E, F \subset E + F$, allora $Giac(E), Giac(F) \subset Giac(E + F)$. Inoltre $P, Q \in E + F$, per cui $\overrightarrow{PQ} \in Giac(E + F)$. Quindi, essendo $Giac(E + F)$ un sottospazio vettoriale, $Giac(E + F) \supset Giac(E) + Giac(F) + \mathrm{Span}(\overrightarrow{PQ})$.
Viceversa, scriviamo $E = P + Giac(E)$, $F = Q + Giac(F)$ e scegliamo due punti R, R' di $E + F$.
Scriviamo
$R = \sum_{i=1}^{h} \lambda_i(P + \underline{v}_i) + \sum_{j=1}^{k} \mu_j(Q + \underline{w}_j)$, con $\underline{v}_i \in Giac(E)$, $\underline{w}_j \in Giac(F)$, $\lambda_i, \mu_j \in \mathbb{K}$, $\sum_{i=1}^{h} \lambda_i + \sum_{j=1}^{k} \mu_j = 1$,
$R' = \sum_{i=1}^{h'} \lambda'_i(P + \underline{v}'_i) + \sum_{j=1}^{k'} \mu'_j(Q + \underline{w}'_j)$, con $\underline{v}'_i \in Giac(E)$, $\underline{w}'_j \in Giac(F)$, $\lambda'_i, \mu'_j \in \mathbb{K}$, $\sum_{i=1}^{h} \lambda'_i + \sum_{j=1}^{k} \mu'_j = 1$.
Osserviamo che $\sum_{i=1}^{h'} \lambda'_i - \sum_{i=1}^{h} \lambda_i = \sum_{i=1}^{h} \mu_i - \sum_{i=1}^{h'} \mu'_i = \alpha$.
Allora $R - R' = \alpha(Q - P) + \underline{v} + \underline{w}$, dove $\underline{v} = \sum_{i=1}^{h} \lambda_i \underline{v}_i - \sum_{i=1}^{h'} \lambda'_i \underline{v}'_i \in Giac(E)$, $\underline{w} = \sum_{i=1}^{h} \mu_i \underline{w}_i - \sum_{i=1}^{h'} \mu'_i \underline{w}'_i \in Giac(F)$. Questo mostra l'altra inclusione.

Osserviamo che se $E \cap F \neq \emptyset$, allora $Giac(E + F) = Giac(E) + Giac(F)$, altrimemti $Giac(E + F) = (Giac(E) + Giac(F)) \oplus \mathrm{Span}(\overrightarrow{PQ})$ con $P \in E$, $Q \in F$.
Otteniamo quindi l'analogo della formula di Grassmann per sottospazi affini non vuoti:
- se $E \cap F \neq \emptyset$, allora vale la usuale formula di Grassmann

$$\dim(E + F) = \dim E + \dim F - \dim E \cap F,$$

come si può vedere anche considerando che $E, F, E \cap F$ e $E + F$ sono sottospazi vettoriali di $\mathbb{A}_P$ con $P \in E \cap F$.
- se $E \cap F = \emptyset$, vale la formula di Grassmann modificata:

$$\dim(E + F) = \dim E + \dim F - \dim(Giac(E) \cap Giac(F)) + 1.$$

Estendiamo adesso la nozione di applicazione lineare.

Definizione:
Siano V, W spazi vettoriali su $\mathbb{K}$, $\mathbb{A}$ spazio affine su V, $\mathbb{B}$ spazio affine su W. $f : \mathbb{A} \to \mathbb{B}$ si dice *applicazione affine* se conserva le combinazioni affini di punti: per ogni $Q_1, \dots, Q_k \in \mathbb{A}$, $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tali che $\sum \lambda_i = 1$,

$$f(\sum_{i=1}^{k} \lambda_i Q_i) = \sum_{i=1}^{k} \lambda_i f(Q_i).$$

Fissato $P \in \mathbb{A}$, consideriamo f come applicazione $f_P : \mathbb{A}_P \to \mathbb{B}_{f(P)}$. Mostriamo che f_P è lineare.
Dati $Q_1, Q_2 \in \mathbb{A}_P$,
$f_P(Q_1 +_P Q_2) = f(Q_1 + Q_2 - P) = f(Q_1) + f(Q_2) - f(P) = f_P(Q_1) +_f (P) f_P(Q_2)$
mostra che f_P è additiva.
Dati $Q \in \mathbb{A}_P$, $\mu \in \mathbb{K}$,
$f_P(\mu \cdot_P Q) = f(\mu Q + (1-\mu)P) = \mu f(Q) + (1-\mu) f(P) = \mu \cdot_{f(P)} f_P(Q)$
mostra che f_p è omogenea.

Viceversa, e con la stessa dimostrazione, una $F : \mathbb{A} \to \mathbb{B}$ per cui esistono $P \in \mathbb{A}$, $Q \in \mathbb{B}$, tali che F è lineare se considerata come applicazione $F : \mathbb{A}_P \to \mathbb{B}_Q$, allora F è una applicazione affine (e $F(P) = Q$).

Usando gli isomorfismi $S_P : V \to \mathbb{A}_P$, $S_{f(P)} : W \to \mathbb{B}_{f(P)}$ otteniamo una applicazione lineare $df_P = S^{-1}_{f(P)} \circ f_P \circ S_P : V \to W$.

Proposizione
Nelle ipotesi sopra, df_P non dipende da P.

Dimostrazione
Se $P_1, P_2 \in \mathbb{A}$, vogliamo vedere che $S^{-1}_{f(P_1)} \circ f \circ S_{P_1} = S^{-1}_{f(P_2)} \circ f \circ S_{P_2}$, ovvero che $S_{f(P_1)} S^{-1}_{f(P_2)} \circ f \circ S_{P_2} \circ S^{-1}_{P_1} = f$. Ricordando che $S_{f(P_1)} S^{-1}_{f(P_2)} = \tau_{f(P_1)-f(P_2)}$, $S_{P_2} \circ S^{-1}_{P_1} = \tau_{P_2 - P_1}$, per ogni $Q \in \mathbb{A}$ si ha $(\tau_{f(P_1)-f(P_2)} \circ f \circ \tau_{P_2-P_1})(Q) = f(Q + P_2 - P_1) + f(P_1) - f(P_2) = f(Q) + f(P_2) - f(P_1) + f(P_1) - f(P_2) = f(Q)$, come voluto. ∎

L'applicazione $df_P \in \mathrm{Hom}(V, W)$ si dice *differenziale* di f (o *applicazione tangente* di f) e si indica con df. Abbiamo, per ogni $P \in \mathbb{A}$, il diagramma commutativo

$$\begin{array}{ccc} \mathbb{A}_P & \xrightarrow{f_P} & \mathbb{B}_{f(P)} \\ {\scriptstyle S_P}\uparrow & \circlearrowleft & \uparrow{\scriptstyle S_{f(P)}} \\ V & \xrightarrow{df} & W \end{array}$$

per cui da $f \circ S_P = S_{f(P)} \circ df$ otteniamo $df(\underline{v}) + f(P) = f(P + \underline{v})$ per ogni $\underline{v} \in V$.
Posto $\underline{v} = \overrightarrow{PQ}$ si ha $df(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$, ovvero $f(Q) = f(P) + df(Q - P)$ per ogni $P, Q \in \mathbb{A}$.

Viceversa, data $F : V \to W$ lineare, e dati $P \in \mathbb{A}$, $Q \in \mathbb{B}$, allora l'applicazione $f = S_Q \circ F \circ S_P^{-1} : \mathbb{A} \to \mathbb{B}$ è affine (essendo lineare se pensata come applicazione da $\mathbb{A}_P$ a $\mathbb{B}_Q$) e $df = F$.

Riassumendo, $f : \mathbb{A} \to \mathbb{B}$ è una applicazione affine se e solo se:
- esiste $P \in \mathbb{A}$ tale che $f_P = f : \mathbb{A}_P \to \mathbb{B}_{f(P)}$, è lineare;
- oppure, esistono $df \in \mathrm{Hom}(V, W)$, $P \in \mathbb{A}$ tali che $f = S_{f(P)} \circ df \circ S_P^{-1}$.

Notiamo che f è iniettiva/surgettiva se e solo se df lo è. Quindi f è biunivoca se e solo se df è un isomorfismo di spazi vettoriali.
Una $f : \mathbb{A} \to \mathbb{B}$ affine si dice *isomorfismo affine* se è biunivoca e la sua inversa è affine.
È immediato verificare che la composizione di applicazioni affini è affine e il differenziale della composizione è la composizione dei rispettivi differenziali; che $d(id_{\mathbb{A}}) = id_V$; e infine che se una applicazione affine è invertibile, l'inversa è automaticamente affine e il differenziale dell'inversa è l'inversa del differenziale.

Nel caso di una applicazione affine $f : \mathbb{A}(V) \to \mathbb{A}(W)$, scelto $P = \underline{0}$ abbiamo $f(\underline{v}) = f(\underline{0}) + df(\underline{v})$ per ogni $\underline{v} \in \mathbb{A}(V) = V$, ovvero $f = \tau_{f(\underline{0})} \circ df$ è composizione, in modo unico, di una traslazione di $\mathbb{A}(W) = W$ con una applicazione lineare da V a W.
Nel caso di una applicazione affine $f : \mathbb{A}^n \to \mathbb{A}^m$, esitono unici $A \in M(m, n, \mathbb{K})$, $\underline{b} \in \mathbb{K}^m$ tali che $f(\underline{x}) = A\underline{x} + \underline{b}$ per ogni $\underline{x} \in \mathbb{A}^n$ ($df = L_A$, $\underline{b} = f(\underline{0})$).

È inoltre immediato vedere che, data $f : \mathbb{A} \to \mathbb{B}$ affine, se $E \subset \mathbb{A}$ è un sottospazio affine allora $f(E)$ è un sottospazio affine. In particolare abbiamo che, se $E = P + U$ con $P \in \mathbb{A}$ e $U \subset V$ sottospazio, allora $f(E) = f(P) + df(U)$, per cui $Giac(f(E)) = df(Giac(E))$. Notiamo che $\dim f(E) \leq \dim E$ e che se f è iniettiva, allora manda sottospazi affini in sottospazi affini della stessa dimensione.
Le $f : \mathbb{A} \to \mathbb{A}$ affini biunivoche si dicono *affinità* e l'insieme delle affinità di $\mathbb{A}$ si indica con $Aff(\mathbb{A})$, che dotato della composizione è un gruppo.

Estendiamo adesso i concetti di lineare indipendenza e di base.

Definizione:
$Q_0, \ldots, Q_m \in \mathbb{A}$ si dicono *affinemente indipendenti* se $Q_1, \ldots, Q_m$ sono linearmente idipendenti in $\mathbb{A}_{Q_0}$. Equivalentemente, se e solo se $\overrightarrow{Q_0Q_1}, \ldots, \overrightarrow{Q_0Q_m}$ sono linearmente indipendenti in V.

Mostriamo che la definizione non dipende dall'ordine dei punti: dati $\lambda_i \in \mathbb{K}$,

$$\lambda_1 \cdot_{Q_0} Q_1 +_{Q_0} \cdots +_{Q_0} \lambda_m \cdot_{Q_0} Q_m = \lambda_1 Q_1 + \cdots + \lambda_m Q_m + (1 - \sum_{i=1}^{m} \lambda_i) Q_0 =$$

$$= (1 - \sum_{i=1}^{m} \lambda_i) Q_0 + \lambda_1 Q_1 + \cdots + \lambda_{m-1} Q_{m-1} + \lambda_m Q_m =$$

$$= \lambda_0 \cdot_{Q_m} Q_0 +_{Q_m} \lambda_1 \cdot_{Q_m} Q_1 +_{Q_m} \cdots +_{Q_m} \lambda_{m-1} \cdot_{Q_m} Q_{m-1}, \text{ dove } \lambda_0 = (1 - \sum_{i=1}^{m} + \lambda_i).$$

Infatti, $\lambda_m = 1 - (\lambda_0 + \cdots + \lambda_{m-1})$.
Quindi supponiamo che $Q_0, \ldots, Q_{m-1}$ siano linearmente indipendenti in $\mathbb{A}_{Q_m}$.
Da $\lambda_1 \cdot_{Q_0} Q_1 +_{Q_0} \cdots +_{Q_0} \lambda_m \cdot_{Q_0} Q_m = Q_0$ si ha $\lambda_0 = 1$ e $\lambda_1 = \cdots = \lambda_{m-1} = 0$, da cui anche $\lambda_m = 0$. Quindi $Q_1, \ldots, Q_m$ sono linearmente indipendenti in $\mathbb{A}_{Q_0}$
Analogamente si ragiona supponendo $Q_1, \ldots, Q_m$ linearmente indipendenti in $\mathbb{A}_{Q_0}$ e si ottiene $Q_0, \ldots, Q_{m-1}$ linearmente indipendenti in $\mathbb{A}_{Q_m}$.
Formule analoghe valgono per le combinazioni lineari negli altri $\mathbb{A}_{Q_i}$.

Osserviamo che se $Q_0, \ldots, Q_m \in \mathbb{A}$ sono affinemente indipendenti, allora $m \leq \dim A = \dim V$, e $m = \dim A$ se e solo se $Q_1, \ldots, Q_m$ sono una base di $\mathbb{A}_{Q_0}$ (ovvero, $\overrightarrow{Q_0Q_1}, \ldots, \overrightarrow{Q_0Q_m}$ sono una base di V).

Definizione:
Se $\dim A = n$, un insieme ordinato di $n+1$ punti di $\mathbb{A}$ affinemente indipendenti $\{P_0, \ldots, P_n\}$ si dice *riferimento affine* di $\mathbb{A}$.

Le proprietà seguenti sono immediate e derivano direttamente dagli analoghi vettoriali:
- $\mathcal{R} = \{P_0, \ldots, P_n\} \subset \mathbb{A}$ è un riferimento affine di $\mathbb{A}$ se e solo se ogni $Q \in \mathbb{A}$ si scrive in modo unico come combinazione affine di elementi di $\mathcal{R}$: esistono unici $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$, $\sum_{i=0}^{n} \lambda_i = 1$, tali che $Q = \lambda_0 P_0 + \cdots + \lambda_n P_n$. $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ si dicono le *coordinate affini* di Q nel riferimento affine $\mathcal{R}$ e si indicano con $[Q]_{\mathcal{R}}$.
- Una applicazione affine $f : \mathbb{A} \to \mathbb{B}$ è univocamente determinata dai suoi valori su un riferimento affine di $\mathbb{A}$: se $\mathcal{R} = \{P_0, \ldots, P_n\} \subset \mathbb{A}$ è un riferimento affine di $\mathbb{A}$ e sono dati $Q_0, \ldots, Q_n \in \mathbb{B}$, allora esiste una unica $f : \mathbb{A} \to \mathbb{B}$ affine tale che $f(P_i) = Q_i$ per $i = 0\_n$. Esplicitamente, se $P = \lambda_0 P_0 + \cdots + \lambda_n P_n$, $\lambda_i \in \mathbb{K}$, $\sum_{i=0}^{n} \lambda_i = 1$, allora $f(P) = \lambda_0 Q_0 + \cdots + \lambda_n Q_n$.
- $f : \mathbb{A} \to \mathbb{B}$ è un isomorfismo affine se e solo se manda riferimenti affini di $\mathbb{A}$ in riferimenti affini di $\mathbb{B}$.
Per $\mathbb{A}^n$, $\{\underline{0}, \underline{e}_1, \ldots, \underline{e}_n\}$ è il riferimento affine *canonico*. Se $\dim A = n$, allora fissato un riferimento affine di $\mathbb{A}$ $\mathcal{R} = \{P_0, \ldots, P_n\} \subset \mathbb{A}$, l'unica applicazione affine $f : \mathbb{A} \to \mathbb{A}^n$ che manda $P_0 \mapsto \underline{0}$, $P_i \mapsto \underline{e}_i$ per $i = 1\_n$, è un isomorfismo affine, detto isomorfismo di passaggio alle coordinate affini rispetto al riferimento $\mathcal{R}$. Ovviamente, per ogni $Q \in \mathbb{A}$, $f(Q) = [Q]_{\mathcal{R}}$, overo se $Q = \lambda_0 P_0 + \cdots + \lambda_n P_n$, $f(Q) = \lambda_1 \underline{e}_1 + \cdots + \lambda_n \underline{e}_n$.

Dati $P, Q \in \mathbb{A}$ distinti, $r(P,Q) = \mathrm{Span}_a(P,Q)$ è la *retta affine* per P e Q. Notiamo che $\{P, Q\}$ è un riferimento affine per $r(P,Q)$ (come pure una qualsiasi coppia di punti distinti in $r(P,Q)$).
Due rette affini con la stessa giacitura si dicono *parallele*. In generale, due sottospazi affini $E, F \subset \mathbb{A}$ si dicono paralleli se $Giac(E) \subset Giac(F)$ oppure $Giac(F) \subset Giac(E)$.
Dato $R \in r(P,Q)$, R si scrive in modo unico come $R = \lambda Q + (1-\lambda)P$ ($\overrightarrow{PR} = \lambda \overrightarrow{PQ}$). λ si dice *rapporto semplice* della terna ordinata (P, Q, R) e

si indica con $[P;Q;R]$. Notiamo che $[P;Q;R] = 0 \iff R = P$, $[P;Q;R] = 1 \iff R = Q$.
Osserviamo che una applicazione affine iniettiva manda rette affini in rette affini, preserva il parallelismo e preserva il rapporto semplice di terne di punti allineati (dove $Q_1, \ldots, Q_m \in \mathbb{A}$, $m \geq 3$, distinti si dicono *allineati* se appartengono alla stessa retta affine).

Cosideriamo il gruppo delle affinità di $\mathbb{A}$, $Aff(\mathbb{A})$. Elenchiamo alcune proprietà che sono verificate da ogni $f \in Aff(\mathbb{A})$.

1. f è biunivoca;
2. f manda sottospazi affini in sottospazi affini della stessa dimensione preservando il rapporto semplice;
3. f manda rette affini in rette affini preservando il rapporto semplice;
4. f manda rette affini in rette affini ed esiste una retta affine r tale che $f_{|_r}$ preserva il rapporto semplice;
5. f manda rette affini in rette affini.

Chiaramente $2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5$.
Un problema interessante è quello di determinare un sottoinsieme minimale di queste proprietà che sia sufficiente affinché $f \in Aff(\mathbb{A})$ (e la risposta dipenderà dalle proprietà algebriche del campo $\mathbb{K}$).

In questa direzione, vale il seguente teorema, detto teorema fondamentale della geometria affine:

Teorema
Sia V uno spazio vettoriale di dimensione almeno 2 sul campo $\mathbb{K}$ di caratteristica diversa da 2, e sia $\mathbb{A}$ uno spazio affine su V. Sia $f : \mathbb{A} \to \mathbb{A}$ biunivoca. Allora

(a). Se f manda rette affini in rette affini, allora f manda sottospazi affini in sottospazi affini preservando la dimensione ($5 \Rightarrow 2$).

(b). Se f manda rette affini in rette affini ed esiste una retta affine r tale che $f_{|_r}$ preserva il rapporto semplice, allora $f \in Aff(\mathbb{A})$ ($4 \Rightarrow f \in Aff(\mathbb{A})$).

(c). Se $\mathbb{K} = \mathbb{R}$ e f manda rette affini in rette affini, allora $f \in Aff(\mathbb{A})$ ($5 \Rightarrow f \in Aff(\mathbb{A})$ su $\mathbb{R}$).

Osservazione: la parte finale di questo capitolo non è materia di esame.

Premettiamo il seguente:

Lemma
Se $\mathbb{K}$ ha caratteristica diversa da 2, $E, F \subset \mathbb{A}$ sottospazi affini tali che $E \cap F \neq \emptyset$

e $\dim E > 0$, allora $E + F = \bigcup_{E \ni P \neq Q \in F} r(P, Q)$.

Dimostrazione
Notiamo che l'inclusione $E + F \supset \bigcup_{E \ni P \neq Q \in F} r(P, Q)$ è sempre vera, essendo $E + F$ chiuso per combinazioni affini dei suoi punti.
Fissiamo $P_0 \in E \cap F$, e consideriamo E e F come sottospazi vettoriali di $\mathbb{A}_{P_0}$.
Notiamo che $P_0 \in r(P, P_0)$ per ogni $P \in E$, $P \neq P_0$.
Se $S \in E + F$, $S \neq P_0$, allora scriviamo $S = A +_{P_0} B = A + B - P_0$ con $A \in E$, $B \in F$. Otteniamo $P = 2 \cdot_{P_0} A = 2A - P_0 \in E$, $Q = 2 \cdot_{P_0} B = 2B - P_0 \in F$ e la retta $r(P, Q)$ contiene S, in quanto $S = \frac{1}{2}P + \frac{1}{2}Q$. ❑

Dimostriamo adesso il teorema.

Dimostrazione
Possiamo dimostrare (a) per induzione sulla dimensione di un sottospazio affine $E \subset \mathbb{A}$. Se $\dim E = 1$, E è una retta affine e per ipotesi $f(E)$ è una retta affine.
Se $\dim E > 1$, allora scelto $P_0 \in E$, consideriamo E come sottospazio vettoriale di $\mathbb{A}_{P_0}$ e scriviamo $E = W \oplus U$ con W, U sottospazi vettoriali di $\mathbb{A}_{P_0}$, $\dim U = 1$.
Per induzione $f(W)$ è un sottospazio affine di dimensione $\dim E - 1$, e $f(U)$ è una retta affine che interseca $f(W)$ in un punto, essendo f biunivoca. Allora, per il lemma, E è l'unione delle rette che congiungono un punto di W e uno di U e quindi $f(W)$ è l'unione delle rette che congiungono un punto di $f(W)$ e uno di $f(U)$, ovvero $f(E) = f(W) + f(U)$ che è un sottospazio affine di dimensione $\dim f(W) + \dim f(U) - \dim f(W) \cap f(U) = \dim W + \dim U = \dim E$.

Osserviamo che se due rette $r, r' \subset \mathbb{A}$ hanno la stessa giacitura, allora anche $f(r), f(r')$ hanno la stessa giacitura; ovvero f preserva il parallelismo tra rette. Infatti, se $r \neq r'$, allora $r \cap r' = \emptyset$ e quindi $\dim(r + r') = 2$ ($r + r'$ è un piano affine). Ne segue che $H = f(r + r')$ ha dimensione 2 e contiene le rette $f(r), f(r')$ che sono disgiunte, essendo f biunivoca. Se avessero giaciture diverse, allora $Giac(f(r)) + Giac(f(r')) = Giac(H)$ e quindi $f(r), f(r')$ si intersecherebbero in H ↯.

Fissiamo $P_0 \in \mathbb{A}$, e sia $P \in \mathbb{A}$, $P \neq P_0$. Allora f manda la retta $r(P_0, P)$ nella retta $r(f(P_0), f(P))$. Per $\lambda \in \mathbb{K}$, poniamo $P_\lambda = \lambda P + (1-\lambda)P_0 = P_0 + \lambda(P - P_0)$.
allora $f(P_\lambda) = f(P_0) + \mu(f(P) - f(P_0))$ per un certo $\mu \in \mathbb{K}$ che dipende da P e da λ. Ponendo $\mu = \varphi(P, \lambda)$, otteniamo $\varphi : (\mathbb{A} \setminus \{P_0\}) \times \mathbb{K} \to \mathbb{K}$ tale che per ogni $P \in A$, $P \neq P_0$, $\varphi(P, 0) = 0$, $\varphi(P, 1) = 1$ e la restrizione di φ a $\{P\} \times \mathbb{K}$ è biunivoca.

Analizziamo la dipendenza di φ dalla scelta del punto P.
Sia Q un altro punto di $\mathbb{A}$ tale che P_0, P, Q siano affinemente indipendenti.
I tre punti generano un piano affine $H \subset \mathbb{A}$ che contiene le due rette $r(P_0, P)$ e $r(P_0, Q)$ incidenti in P_0.
Per $\lambda \in \mathbb{K}$, consideriamo i punti $P_\lambda = P_0 + \lambda(P - P_0)$, $Q_\lambda = P_0 + \lambda(Q - P_0)$.

Se $\lambda \neq 0$, le due rette $r(P,Q)$ e $r(P_\lambda, Q_\lambda)$ hanno entrambe giacitura $\mathrm{Span}(Q-P)$ ($Q_\lambda - P_\lambda = \lambda(Q-P)$), e quindi sono rette parallele nel piano H. Quindi, le rette $r(f(P), f(Q))$, $r(f(P_\lambda), f(Q_\lambda))$ sono parallele nel piano $f(H)$.
Dunque, $f(Q_\lambda) - f(P_\lambda) = \varphi(Q,\lambda)(f(Q) - f(P_0)) - \varphi(P,\lambda)(f(P) - f(P_0))$ deve essere un multiplo di $f(Q) - f(P) = (f(Q) - f(P_0)) - (f(P) - f(P_0))$.
Essendo $f(Q) - f(P_0)$, $f(P) - f(P_0)$ linearmente indipendenti, segue che $\varphi(P,\lambda) = \varphi(Q,\lambda)$.
Osserviamo che questo vale anche se sostituiamo P con qualsiasi altro punto della retta $r(P_0, P)$ diverso da P_0.
Abbiamo così dimostrato che φ non dipende dalla scelta del punto $P \neq P_0$. Scriveremo semplicemente $\varphi(\lambda)$ al posto di $\varphi(P,\lambda)$ e considereremo φ come applicazione $\varphi : \mathbb{K} \to \mathbb{K}$.

Siano $P_0, P, Q \in \mathbb{A}$ affinemente indipendenti come sopra.
Vogliamo mostrare che per ogni $\lambda \in \mathbb{K}$,

$$f(Q + \lambda(P - P_0)) = f(Q) + \varphi(\lambda)(f(P) - f(P_0).$$

Nel piano H generato da P_0, P, Q, il punto $Q + \lambda(P - P_0)$ è l'intersezione della retta $r(Q, Q + (P - P_0))$ (parallela alla retta $r(P_0, P)$) e la retta passante per $P_\lambda = P_0 + \lambda(P - P_0)$ parallela alla retta $r(P_0, Q)$.
Considerando la configurazione immagine tramite f di queste due coppie di rette parallele, si realizza che, per ogni $\lambda \in \mathbb{K}$, il punto $f(Q + \lambda(P - P_0))$ è l'intersezione della retta $\{f(Q) + s(f(P) - f(P_0))|\ s \in \mathbb{K}\}$ e della retta $\{f(P_0) + \varphi(\lambda)(f(P) - f(P_0)) + r(f(Q) - f(P_0))|\ r \in \mathbb{K}\}$. Poiché i vettori $f(P) - f(P_0)$ e $f(Q) - f(P_0)$ sono linearmente indipendenti, il punto di intersezione si ha per $s = \varphi(t)$ e $r = 1$, da cui otteniamo che, come voluto, $f(Q + \lambda(P - P_0)) = f(Q) + \varphi(\lambda)(f(P) - f(P_0))$.

Consideriamo ora un riferimento affine di $\mathbb{A}$ $\mathcal{R} = \{P_0, P_1, \ldots, P_n\}$.
Per ogni $i = 2\_\_n$ la retta $r(P_0, P_i)$ interseca in P_0 il sottospazio affine $\mathrm{Span}_a(P_0.P_1, \ldots, P_{i-1})$ ed essendo f biunivoca la stessa proprietà vale per le immagini $f(P_0), \ldots, f(P_n)$, quindi $f(\mathcal{R})$ è un riferimento affine di $\mathbb{A}$.
Ogni $P \in \mathbb{A}$ si scrive in modo unico come combinazione affine del riferimento $\mathcal{R}$, $P = P_0 + \sum_{j=1}^{n} a_j(P_j - P_0)$, e quindi, applicando induttivamente quanto visto prima, abbiamo che

$$\begin{aligned} f(P) &= f(P_0 + \sum_{j=1}^{n} a_j(P_J - P_0)) = \\ &= f(P_0 + \sum_{j=1}^{n-1} a_j(P_J - P_0)) + \varphi(a_n)(f(P_n) - f(P_0)) = \\ &= \cdots = \\ &= f(P_0) + \sum_{j=1}^{n} \varphi(a_j)(f(P_j) - f(P_0)). \end{aligned}$$

Riassumendo il risultato di questa analisi abbiamo il seguente:

Se $f : \mathbb{A} \to \mathbb{A}$ è biunivoca e manda rette affini in rette affini (e $\mathbb{K}$ non ha caratteristica 2), allora f è affine se e solo se $\varphi = id_{\mathbb{K}}$.

Possiamo ora dimostrare il punto (b).
Sia r una retta tale che $f_{|_r} : r \to f(r)$ preserva il rapporto semplice.
Possiamo svolgere la discussione precedente scegliendo $P_0, P \in r$. Allora, per ogni $Q \in r$, si ha $Q = P_0 + \lambda(P - P_0)$, con $\lambda = [P_0, P, Q]$, quindi, $f(P) = f(P_0) + \varphi(\lambda)(f(Q) - f(P_0)) = f(P_0) + \lambda(f(Q) - f(P_0))$.
Per l'arbitrarietà di Q otteniamo che per ogni $\lambda \in \mathbb{K}$, $\varphi(\lambda) = \lambda$ come voluto.

Mostriamo adesso che $\varphi : \mathbb{K} \to \mathbb{K}$ è un isomorfismo di campi.
Sappiamo già che $\varphi(0) = 0$ e $\varphi(1) = 1$.
Mostriamo che per ogni $t_1, t_2 \in \mathbb{K}$, $\varphi(t_1 + t_2) = \varphi(t_1) + \varphi(t_2)$.
L'idea è quella di rappresentare geometricamente la somma $t_1 + t_2$ per mezzo di una opportuna configurazione di rette parallele o incidenti e ottenere il risultato per mezzo della configurazione immagine tramite f. Siano P_0, P, Q come prima. Consideriamo le rette $r(P_0, P)$ e $r(P_0, Q)$ nel piano H generato da P_0, P, Q, incidenti nel punto P_0. La prima retta ha la parametrizzazione

$$P_t = P_0 + t(P - P_0), \quad t \in \mathbb{K}.$$

Consideriamo la retta $r(P_0, P) + (Q - P_0)$ parametrizzata da

$$R_t = Q + t(P - P_0) = P_0 + (Q - P_0) + t(P - P_0), \quad t \in \mathbb{K}.$$

Essa è parallela a $r(P_0, P)$ e incidente alla la retta $r(P_0, Q)$ nel punto Q.
Nella configurazione immagine tramite f abbiamo

$$f(R_t) = f(P_0) + (f(Q) - f(P_0)) + \varphi(t)(f(P) - f(P_0)).$$

Consideriamo il punto

$$R_{t_1+t_2} = P_0 + (Q - P_0) + (t_1 + t_2)(P - P_0) = P_{t_2} + (R_{t_1+t_2} - P_{t_2})$$

dove le due espressioni riflettono la presentazione del punto come l'intersezione di due specifiche rette: $R_{t_1+t_2} = (r(P_0, P) + (Q - P_0)) \cap r(P_{t_2}, R_{t_1+t_2})$.
D'altra parte, la seconda retta è parallela alla retta $r(P_0, R_{t_1})$. Questo comporta l'ulteriore espressione

$$R_{t_1+t_2} = P_0 + t_2(P - P_0) + (Q - P_0) + t_1(P - P_0).$$

Consideriamo la configurazione immagine tramite f. Poiché f preserva il parallelismo e l'incidenza e ricordando la definizione di φ, si ricava che esiste un $s \in \mathbb{K}$ tale che

$$\begin{aligned} f(R_{t_1+t_2}) &= f(P_0) + (f(Q) - f(P_0)) + \varphi(t_1 + t_2)(f(P) - f(P_0)) = \\ &= f(P_0) + \varphi(t_2)(f(P) - f(P_0)) + s(f(Q) - f(P_0)) + \varphi(t_1)(f(P) - f(P_0)). \end{aligned}$$

Poiché i vettori $f(P) - f(P_0)$ e $f(Q) - f(P_0)$ sono linearmente indipendenti, necessariamente $s = 1$ e $\varphi(t_1 + t_2) = \varphi(t_1) + \varphi(t_2)$.
Mostriamo ora che $\varphi(t_1 t_2) = \varphi(t_1)\varphi(t_2)$.
Con le stesse notazioni di prima, anche in questo caso rappresentiamo geometricamente il prodotto $t_1 t_2$ e concludiamo analizzando la configurazione immagine tramite f. Consideriamo la parametrizzazione standard della retta $r(P_0, Q)$ $Q_t = P_0 + t(Q - P_0)$. Le rette $r(P_{t_1}, Q)$ e $r(P_{t_1 t_2} Q_{t_2})$ sono parallele e possiamo esprimerle rispettivamente come

$$r(P_{t_1}, Q) = \{P_0 + (1 - s)(Q - P_0) + st_1(P - P_0) |\ s \in \mathbb{K}\},$$

$$r(P_{t_1 t_2}, Q_{t_2}) = \{P_0 + (1 - s)t_2(Q - P_0) + s(t_1 t_2)(P - P_0) |\ s \in \mathbb{K}\}.$$

Applicando la "relazione di Talete" ai due triangoli di vertici rispettivamente P_0, Q, P_{t_1} e $P_0, Q_{t_2}, P_{t_1 t_2}$, otteniamo $t_1/1 = t_1 t_2/t_2$.
Passando alle immagini tramite f, abbiamo le rette parallele

$$\{f(P_0) + (1 - s)(f(Q) - f(P_0)) + s\varphi(t_1)(f(P) - fP_0)) |\ s \in \mathbb{K}\},$$

$$\{f(P_0) + (1 - s)\varphi(t_2)(f(Q) - f(P_0)) + s\varphi(t_1 t_2)(f(P) - f(P_0) |\ s \in \mathbb{K}\}.$$

Possiamo infine applicare la relazione di Talete ai due triangoli di vertici rispettivamente $f(P_0), f(Q), f(P_{\varphi(t_1)})$ e $f(P_0), f(Q_{\varphi(t_2)}), f(P_{\varphi(t_1 t_2)})$ e ottenere $\varphi(t_1)/\phi(1) = \varphi(t_1 t_2)/\varphi(t_2)$ cioè $\varphi(t_1 t_2) = \varphi(t_1)\varphi(t_2)$.

Per concludere la dimostrazione di (c), mostriamo che se $\mathbb{K} = \mathbb{R}$, l'unico isomorfismo di campi $\varphi : \mathbb{R} \to \mathbb{R}$ è $id_{\mathbb{R}}$.
Per ogni $n \in \mathbb{Z}$, $n > 0$, $\varphi(1) = 1 \Rightarrow \varphi(n) = \varphi(1) + \cdots + \varphi(1) = n$. Inoltre, $1 = \varphi(n/n) = n\varphi(1/n) \Rightarrow \varphi(1/n) = 1/n$. Lo stesso vale se $n < 0$.
Quindi per ogni $m, n \in \mathbb{Z}$, $n \neq 0$, $\varphi(m/n) = m/n$, cioè $\phi(q) = q$ per ogni $q \in \mathbb{Q}$.
D'altra parte, per ogni $x \in \mathbb{R}$, $x \neq 0$, $\varphi(x^2) = \varphi(x)^2 > 0$ quindi, per ogni $x > 0$, $\varphi(x) > 0$ e φ è strettamente crescente in quanto se $x > y$, $x - y > 0$, per cui $\varphi(x) - \varphi(y) = \varphi(x - y) > 0$.
Se adesso esistesse un $x \in \mathbb{R}$ tale che $\varphi(x) \neq x$, allora $x < \varphi(x)$ oppure $x > \varphi(x)$. Nel primo caso, esiste un razionale q tale che $x < q < \varphi(x)$ da cui $\varphi(x) < \varphi(q) = q < \varphi(x)$ ↯. Il secondo caso è analogo al primo. ❑

Quadriche

Sia $\mathbb{K}$ un campo e sia $\mathbb{K}[t_1, \dots, t_n]$ l'anello dei polinomi in n indeterminate $t_1, \dots, t_n$ a coefficienti in $\mathbb{K}$.
Ogni $p \in \mathbb{K}[t_1, \dots t_n]$ definisce una funzione polinomiale (che indichiamo ancora con p) $p : \mathbb{K}^n \to \mathbb{K}$, $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto p(x_1, \dots, x_n)$, ed un luogo di zeri

$$Z(p) = p^{-1}(\{0\}) = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n | \ p(x_1, \dots, x_n) = 0\}.$$

Abbiamo già incontrato i luoghi di zeri dei polinomi di grado 1, cioè gli iperpiani affini di $\mathbb{K}^n$, e adesso vogliamo studiare i luoghi di zeri dei polinomi di grado 2. Inizieremo con delle considerazioni generali, valide per ogni campo e per ogni grado, per poi restringersi al caso dei polinomi quadratici (dove supporremo che $\mathbb{K}$ abbia caratteristica diversa da 2) ed infine ai casi $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$ (dove daremo una classificazione completa).

Osserviamo che $Z(0) = \mathbb{K}^n$ e che per ogni $p \in \mathbb{K}[t_1, \dots, t_n]$, $Z(p) = Z(\lambda p)$ per ogni $\lambda \in \mathbb{K}^*$. Per questo, ci possiamo limitare ai polinomi non nulli considerati a meno di multipli scalari non nulli.
Diciamo quindi che due polinomi $p_1, p_2 \in \mathbb{K}[t_1, \dots t_n] \setminus \{0\}$ sono *proporzionali* se esiste $\lambda \in \mathbb{K}$, $\lambda \neq 0$, tale che $p_1 = \lambda p_2$.
È immediato verificare che la proporzionalità è una relazione di equivalenza su $\mathbb{K}[t_1, \dots, t_n] \setminus \{0\}$ e che il luogo di zeri e il grado di un polinomio sono degli invarianti.

Dato $p \in \mathbb{K}[t_1, \dots, t_n]$, $\deg p = d$, la sua classe di proporzionalità $[p]$ si dice *ipersuperficie di grado* d e p si dice una sua *equazione*. Ogni ipersuperficie $S = [p]$ determina quindi un sottoinsieme di $\mathbb{K}^n$, $Z(p)$, detto il *supporto* della ipersuperficie, indicato con $Z(S)$.

L'insieme quoziente di $\mathbb{K}[t_1, \dots, t_n] \setminus \{0\}$ modulo la relazione di proporzionalità si indica con $\mathbb{PK}[t_1, \dots t_n]$ è detto *spazio proiettivo* su $\mathbb{K}[t_1, \dots, t_n]$.
Osserviamo che la costruzione dello spazio proiettivo ha senso per ogni spazio vettoriale V su $\mathbb{K}$: la relazione di proporzionalità su $V \setminus \{\underline{0}\}$ si definisce allo stesso modo, $\underline{v}_1 \equiv \underline{v}_2 \iff \underline{v}_1 = \lambda \underline{v}_2$ con $\lambda \in \mathbb{K}$, $\lambda \neq 0$, e lo spazio proiettivo su V è l'insieme quoziente $\mathbb{P}V = (V \setminus \{\underline{0}\})/_{\equiv}$. Per ogni $\underline{v} \in V$, $[\underline{v}] = \mathrm{Span}(\underline{v}) \setminus \{0\}$, per cui $\mathbb{P}V$ parametrizza l'insieme delle rette di V.

Notiamo che le ipersuperfici di grado 0 hanno supporti vuoti, mentre le ipersuperfici di grado 1 hanno per supporti gli iperpiani affini di $\mathbb{K}^n$ (e in generale, i supporti non sono sottospazi vettoriali). Conviene allora pensare $\mathbb{K}^n$ come spazio affine, $\mathbb{A}^n$, invece che come spazio vettoriale. Nel seguito useremo il simbolo $\mathbb{K}^n$ solo quando vorremo evidenziare la struttra vettoriale standard, altrimenti useremo il simbolo $\mathbb{A}^n$, per specificare che stiamo considerando la

struttura affine standard.

Due sottoinsiemi S, T di $\mathbb{A}^n$ si dicono *affinemente equivalenti*, $S \sim_a T$, se esiste $f \in Aff(\mathbb{A}^n)$ tale che $f(S) = T$. Questo definisce una relazione di equivaleza su $\wp(\mathbb{A}^n)$ che possiamo restringere ai supporti delle ipersuperfici: ha quindi senso studiare i supporti a meno di equivalenza affine.

Vediamo come sollevare la relazione di equivalenza affine su $\mathbb{K}[t_1, \dots, t_n] \setminus \{0\}$.

Notiamo che se pensiamo a $p \in \mathbb{K}[t_1, \dots, t_n]$ come funzione polinomiale, allora se $f \in Aff(\mathbb{A}^n)$ si ha

$$f(Z(p)) = Z(p \circ f^{-1}),$$

infatti, $\underline{x} \in f(Z(p)) \iff f^{-1}(\underline{x}) \in Z(p)$.
Possiamo però dare un senso anche alla "composizione" di un polinomio p e di una affinità f ottenendo il polinomio $p \circ f$: scriviamo $f(\underline{x}) = P\underline{x} + \underline{t}$ con $P \in GL(n, \mathbb{K})$, $\underline{t} \in \mathbb{K}^n$; definiamo i polinomi q_j, $j = 1 \_ n$, ponendo formalmente $\begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} = P \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} + \underline{t}$, e poniamo

$$(p \circ f)(t_1, \dots, t_n) = p(q_1(t_1, \dots, t_n), \dots, q_n(t_1, \dots, t_n)).$$

Ad esempio, se $p(t_1, t_2) = t_1^2 t_2 + a t_1 t_2^3 + b t_1 + c$ e $f(\binom{x}{y}) = \binom{3x+y+1}{2x-y-2}$,
allora $q_1(t_1, t_2) = 3t_1 + t_2 + 1$, $q_2(t_1, t_2) = 2t_1 - t_2 - 2$ e
$p \circ f(t_1, t_2) = q_1^2 q_2 + a q_1 q_2^3 + b q_1 + c =$
$= (3t_1 + t_2 + 1)^2 (2t_1 - t_2 - 2) + a(3t_1 + t_2 + 1)(2t_1 - t_2 - 2)^3 + b(3t_1 + t_2 + 1) + c.$

Definiamo quindi l'analogo dell'equivalenza affine su $\mathbb{K}[t_1, \dots, t_n]$:
$p_1 \sim_a p_2$ se e solo se esiste $f \in Aff(\mathbb{A}^n)$ tale che $p_2 = p_1 \circ f$ (il che implica $f(Z(p_2)) = Z(p_1)$).
Notiamo che i polinomi q_j hanno grado 1, quindi $\deg(p \circ f) \leq \deg p$, e, allo stesso modo, poiché p si ricava da $p \circ f$ componendo con f^{-1}, $\deg p \leq \deg(p \circ f)$. Quindi il grado è un invariante per equivalenza affine polinomiale.
Notiamo anche che l'equivalenza affine è compatibile con la proporzionalità, nel senso che $(\lambda p) \circ f = \lambda(p \circ f)$, e quindi è ben definita l'equivalenza affine anche per le ipersuperfici in $\mathbb{PK}[t_1, \dots, t_n]$. Se $C = [p]$ è una ipersuperficie e f una affinità di $\mathbb{A}^n$, poniamo $f(C) = [p \circ f^{-1}]$: al variare di f, si ottengono tutte le ipersuperfici affinemente equivalenti a C, ovvero la clase di equivalenza affine di C che si dice anche il *tipo affine* di C. Notiamo che la definizione è specificamente data in modo che per i supporti si abbia $Z(f(C)) = f(Z(C))$.

Per studiare i supporti delle ipersuperfici a meno di equivalenza affine, possiamo allora studiare le ipersuperfici a meno di equivalenza affine. Questo non è proprio equivalente, perché in generale non c'è una corrispondenza biunivoca tra supporti e ipersuperfici (come vedremo, ad esempio, nel caso $\mathbb{K} = \mathbb{R}$), ma sicuramente la classificazione delle ipersuperfici include quella dei supporti.

Restringeremo la nostra analisi alle ipersuperfici di grado 2 (in grado 0 e 1,

tutte le ipersuperfici sono affinemente equivalenti) dette *quadriche* (*coniche* se $n = 2$). Indichiamo con $Q(n, \mathbb{K})$ l'insieme delle quadriche in n variabili su $\mathbb{K}$.
Il motivo di considerare i polinomi di grado 2 è il seguente fatto: un polinomio $p \in \mathbb{K}[t_1, \dots, t_n]$ di grado 2 si scrive come $p = p_0 + p_1 + p_2$, con p_i omogeneo di grado $\deg p_i = i$; la funzione polinomiale data da p_1 rappresenta un funzionale lineare su $\mathbb{K}^n$, mentre la funzione polinomiale data da p_2 rappresenta la forma quadratica di un prodotto scalare su $\mathbb{K}^n$, ed abbiamo studiato entrambi questi oggetti approfonditamente.
Poiché per lo sviluppo della teoria dei prodotti scalari è stato essenziale supporre che il campo $\mathbb{K}$ non avesse caratteristica 2, facciamo questa ipotesi anche in questo caso: da ora in poi supporremo $char\mathbb{K} \neq 2$.

Più in dettaglio, dato $p \in \mathbb{K}[t_1, \dots, t_n]$, $\deg p \leq 2$, scriviamo

$$p(t_1, \dots, t_n) = \sum_{i=1}^{n} a_{ii} t_i^2 + \sum_{1 \leq i < j \leq n} a_{ij} t_i t_j + \sum_{i=1}^{n} b_i t_i + c,$$

con $a_{ij}, b_i, c \in \mathbb{K}$. Allora, posto

$$A = \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} & \cdots & \frac{1}{2}a_{1n} \\ \frac{1}{2}a_{12} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{2}a_{n-1n} \\ \frac{1}{2}a_{1n} & \cdots & \frac{1}{2}a_{n-1n} & a_{nn} \end{pmatrix}, \quad \underline{b} = \frac{1}{2}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

ovvero, $\underline{b} \in \mathbb{K}^n$ e $A \in S(n, \mathbb{K})$ tale che $\underset{i\text{—}}{\overset{i}{A}} = a_{ii}$ per $i = 1\text{—}n$, e $\underset{i\text{—}}{\overset{j}{A}} = \frac{1}{2}a_{ij}$ per $1 \leq i < j \leq n$, allora

$$p(t_1, \dots, t_n) = (t_1 \ t_2 \ \cdots \ t_n) A \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} + 2\underline{b}^\top \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} + c.$$

Ponendo $T = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}$ possiamo scrivere

$$p(t_1, \dots, t_n) = T^\top A T + 2\underline{b}^\top T + c$$

Notiamo che $\deg p = 2 \iff A \neq 0$.
Possiamo usare una notazione ancora più compatta ponendo $M \in S(n+1, \mathbb{K})$, $M = \begin{pmatrix} A & \underline{b} \\ \underline{b}^\top & c \end{pmatrix}$ e scrivendo

$$p(t_1, \dots, t_n) = (t_1 \ \cdots \ t_n \ 1) M \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ 1 \end{pmatrix},$$

ovvero ponendo $T' = \binom{T}{1}$,

$$p(t_1, \dots, t_n) = (T')^\top M T'.$$

Infatti $(T')^\top M T' = (T')^\top \begin{pmatrix} AT + \underline{b} \\ \underline{b}^\top T + c \end{pmatrix} = T^\top AT + T^\top \underline{b} + \underline{b}^\top T + c$ e basta osservare che $T^\top \underline{b} = \underline{b}^\top T$.
M si dice la *matrice che rappresenta* p e si indica con $\mathcal{M}(p)$; A si dice la *matrice che rappresenta la parte quadratica di* p e si indica con $\mathcal{A}(p)$.
Otteniamo $\mathcal{M} : \mathbb{K}_2[t_1, \dots, t_n] \to S(n+1, \mathbb{K})$ che è un isomorfismo di spazi vettoriali (la linearità segue dal fatto che i coefficienti di $\mathcal{M}(p)$ sono lineari nei coefficienti di p, l'iniettività e la surgettività sono ovvie; è anche ovvio che i due spazi vettoriali hanno la stessa dimensione).

Osserviamo che l'applicazione $J : \mathbb{A}^n \to \mathbb{A}^{n+1}$, $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$ è affine e iniettiva con immagine l'iperpiano affine $H = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} \in \mathbb{A}^{n+1} |\ x_{n+1} = 1 \}$.
Otteniamo un isomorfismo affine J che identifica $\mathbb{A}^n$ con H.
Tramite questa identificazione, vediamo che il luogo di zeri di p risulta essere l'intersezione di H con il cono isotropo del prodotto scalare su $\mathbb{K}^{n+1}$ dato dalla matrice M: $Z(p) = J^{-1}(H \cap CI(\Phi_M))$.

Notiamo che abbiamo un isomorfismo di gruppi tra $Aff(\mathbb{A}^n)$ e $Aff(H)$, dato da $f \mapsto J \circ f \circ J^{-1}$. Esplicitamente, se $f(\underline{x}) = P\underline{x} + \underline{t}$, con $P \in GL(n, \mathbb{K})$, $\underline{t} \in \mathbb{K}^n$,

$$(J \circ f \circ J^{-1})(\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}) = (J \circ f)(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = J(P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \underline{t}) =$$
$$= \begin{pmatrix} P\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \underline{t} \\ 1 \end{pmatrix} = \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}.$$

La matrice $\begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} \in GL(n+1, \mathbb{K})$ si dice la matrice *completa* dell'affinità f, o più semplicemente la matrice che *rappresenta* l'affinità f.

Abbiamo quindi che $Aff(H)$ è dato dal sottogruppo di $GL(n+1, \mathbb{K})$ delle matrici nella forma $\begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix}$ con $P \in GL(n, \mathbb{K})$, $\underline{t} \in \mathbb{K}^n$. Osserviamo che tale sottogruppo coincide con il sottogruppo delle matrici $G \in GL(n+1, \mathbb{K})$ tali che $G(H) = H$. Infatti, è chiaro che una affinità di H manda H in H. Viceversa, se $G \in GL(n+1, \mathbb{K})$ manda H in H, allora, poiché $\underline{e}_n \in H$, $G\underline{e}_n \in H$ e quindi l'ultima colonna di G è del tipo $\begin{pmatrix} b_1 \\ \vdots \\ b_n \\ 1 \end{pmatrix}$; ma per $i = 1\_n$, anche $\underline{e}_i + \underline{e}_n \in H$, per cui $G\underline{e}_i + G\underline{e}_n \in H$ e guardando l'ultima coordinata, la colonna i-ma di G

ha 0 nell'ultima riga. Quindi $G = \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix}$ con $P \in M(n, \mathbb{K})$, $\underline{t} \in \mathbb{K}^n$, e poiché G è invertibile, anche P lo è.

Notiamo che le traslazioni di $\mathbb{A}^n$ corrispondono al sottogruppo delle matrici nella forma $\begin{pmatrix} I_n & \underline{t} \\ 0 & 1 \end{pmatrix}$ con $\underline{t} \in \mathbb{K}^n$.

Osservazione: il cono isotropo di Φ_M è il luogo di zeri del polinomio omogeneo $(t_1 \ \cdots \ t_n \ t_{n+1})M \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ t_{n+1} \end{pmatrix} = T^\top AT + 2\underline{b}^\top T t_{n+1} + c t_{n+1}^2$ e quindi il polinomio $p = p_2 + p_1 + c$ può essere pensato come la "restrizione" ad H del polinomio omogeneo in $n+1$ indeterminate $p_2 + p_1 t_{n+1} + c t_{n+1}^2$.

È chiaro come cambia la matrice che rappresenta un polinomio quadratico quando il polinomio cambia per proporzionalità: $\mathcal{M}(\lambda p) = \lambda \mathcal{M}(p)$.
Meno ovvio è come cambia la matrice quando il polinomio cambia per equivalenza affine.
Data $f \in Aff(\mathbb{A}^n)$, scriviamo $f(\underline{x}) = P\underline{x} + \underline{t}$, $P \in GL(n, \mathbb{K})$, $\underline{t} \in \mathbb{K}^n$. Allora $p \circ f$ si ottiene valutando p su $\begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} = P \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} + \underline{t}$ e, osservando che $\begin{pmatrix} q_1 \\ \vdots \\ q_n \\ 1 \end{pmatrix} = \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ 1 \end{pmatrix}$, si ha che

$$p \circ f(t_1, \dots, t_n) = (\begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ 1 \end{pmatrix})^\top M \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ 1 \end{pmatrix}.$$

Quindi $\mathcal{M}(p \circ f) = \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix}^\top M \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix}$.
La matrice $\mathcal{M}(p)$ cambia quindi per congruenza, coerentemente con il fatto che la pensiamo come matrice di un prodotto scalare, ma non tutte le matrici invertibili sono ammesse, solo quelle che fissano H.
Più esplicitamente, se $\mathcal{M}(p) = \begin{pmatrix} A & \underline{b} \\ \underline{b}^\top & c \end{pmatrix}$, la nuova matrice è

$$\mathcal{M}(p \circ f) = \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix}^\top \begin{pmatrix} A & \underline{b} \\ \underline{b}^\top & c \end{pmatrix} \begin{pmatrix} P & \underline{t} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} P^\top & 0 \\ \underline{t}^\top & 1 \end{pmatrix} \begin{pmatrix} AP & A\underline{t} + \underline{b} \\ \underline{b}^\top P & \underline{b}^\top \underline{t} + c \end{pmatrix} =$$
$$= \begin{pmatrix} P^\top AP & P^\top A\underline{t} + P^\top \underline{b} \\ \underline{t}^\top AP + \underline{b}^\top P & \underline{t}^\top A\underline{t} + 2\underline{b}^\top \underline{t} + c \end{pmatrix}.$$

Notiamo che anche la matrice A che rappresenta la parte quadratica p_2 del polinomio p cambia per congruenza, e che questa volta sono ammesse tutte le matrici invertibili.

Notiamo subito che, visto che $\mathcal{M}(p)$ e $\mathcal{A}(p)$ cambiano per congruenza se p cambia per equivalenza affine, mentre cambiano per proporzionalità se p varia per proporzionalità, gli invarianti per isometria di $\mathcal{M}(p)$ e $\mathcal{A}(p)$ che sono invarianti anche per proporzionalità sono invarianti per l'equivalenza affine di quadriche. Ad esempio, $\operatorname{rnk}\mathcal{M}(p)$, $\operatorname{rnk}\mathcal{A}(p)$, $W(\mathcal{M}(p))$ e $W(\mathcal{A}(p))$ son invarianti per equivalenza affine di quadriche (osserviamo che nel caso $\mathbb{K} = \mathbb{R}$, la segnatura di un prodotto scalare non è invariante per proporzionalità: ad esempio si ha $\sigma(-\phi) = (i_-(\phi), i_+(\phi), i_0(\phi)))$.

Abbiamo quindi una bigezione tra $Q(n, \mathbb{K})$ (polinomi di grado 2 in n variabili su $\mathbb{K}$ a meno di proporzionalità) e il sottoinsieme di $\mathbb{P}S(n+1, \mathbb{K})$ (matrici simmetriche di ordine $n+1$ a coefficienti in $\mathbb{K}$ a meno di proporzionalità) dato dalle classi delle matrici M in cui la sottomatrice $A = \underset{n+1}{\underset{\longrightarrow}{M}}\overset{n+1}{|} \neq 0$.
Nel seguito rappresenteremo una ipersuperficie quadrica tramite la matrice simmetrica che rappresenta una sua equazione e trasferiremo lo studio della equivalenza affine di quadriche nello studio della relazione data sulle matrici simmetriche rappresentative dalla congruenza per mezzo di matrici in $Aff(H)$ (che chiameremo di nuovo equivalenza affine).
In particolare, se $\mathbb{K} = \mathbb{R}, \mathbb{C}$, vogliamo trovare invarianti completi per l'equivalenza affine di quadriche (o matrici simmetriche) e una equazione privilegiata (o una forma matriciale normale) in ogni classe di equivalenza affine.

Fissato $O \in \mathbb{A}^n$, la *simmetria centrale di centro* O è l'affinità $\sigma_O \in Aff(\mathbb{A}^n)$ data dalla mappa antipodale nello spazio vettoriale $\mathbb{A}^n_O$, $\sigma_O = -id_{\mathbb{A}^n_O}$. Esplicitamente, $\sigma_O(X) = 2O - X$ per ogni $X \in \mathbb{A}^n$, ovvero, σ_O corrisponde, come affinità di H, alla matrice $\begin{pmatrix} -I_n & 2O \\ 0 & 1 \end{pmatrix}$.
Data una quadrica $C \in Q(n, \mathbb{K})$, $O \in \mathbb{A}^n$ si dice un *centro* di C se $\sigma_O(C) = C$. Notiamo che, se il supporto di C non è vuoto, allora O è un vero e proprio centro di simmetria per il supporto, nel senso che $\sigma_O(Z(C)) = Z(C)$. Se C ha almeno un centro, diciamo che C è una quadrica *a centro*.
Se C si rappresenta con la matrice $M = \begin{pmatrix} A & \underline{b} \\ \underline{b}^\top & c \end{pmatrix}$, allora $\sigma_O(C)$ si rappresenta con la matrice $M' = \begin{pmatrix} A & -2AO - \underline{b} \\ -2O^\top A - \underline{b}^\top & 4O^\top(AO + \underline{b}) + c \end{pmatrix}$.
O è un centro di C se e solo se esiste $\lambda \in \mathbb{K}$, $\lambda \neq 0$, tale che $M' = \lambda M$. Poiché $A \neq 0$, da $A = \lambda A$ segue $\lambda = 1$. Quindi, $2AO + \underline{b} = -\underline{b}$, ovvero $AO = -\underline{b}$, che implica anche l'uguaglianza dei termini noti.
Quindi, C è a centro se e solo se il sistema lineare $A\underline{x} = -\underline{b}$ è risolubile, ovvero se e solo se $\operatorname{rnk} A = \operatorname{rnk}(A|\underline{b})$. In tal caso, i centri formano un sottospazio affine di $\mathbb{A}^n$. In particolare, se A è invertibile, esiste un unico centro per C.

Ne caso C sia a centro, usando la traslazione τ_{-O} con O un centro di C (τ_O corrisponde alla matrice $\begin{pmatrix} I_n & O \\ 0 & 1 \end{pmatrix}$; geometricamente, stiamo traslando il centro

nell'origine di $\mathbb{K}^n$), C è affinemente equivalente alla quadrica con matrice

$$\begin{pmatrix} A & AO+\underline{b} \\ O^\top A+\underline{b}^\top & O^\top AO+2\underline{b}^\top O+c \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & \underline{b}^\top O+c \end{pmatrix}.$$

Consideriamo il prodotto scalare Φ_A su $\mathbb{K}^n$. Poichè esiste una base ortogonale, allora esiste $P \in GL(n,\mathbb{K})$ tale che $P^\top AP = D$ è diagonale (e, poiché $A \neq 0$, possiamo richiedere che $\underline{D}^{1}_{1}| \neq 0$). Usando allora l'affinità data da P (che corrisponde alla matrice $\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}$), C è affinemente equivalente alla quadrica con matrice

$$\begin{pmatrix} D & 0 \\ 0 & \underline{b}^\top O+c \end{pmatrix}.$$

Poiché se f è un'affinità di $\mathbb{A}^n$, f manda un centro di C in un centro di $f(C)$ (ovvero "essere a centro" è una proprietà invariante per equivalenza affine), è vero anche il viceversa: se nella classe di equivalenza di C esiste una quadrica con marice diagonale (per cui $\underline{0}$ è un centro), allora C è a centro.
Infatti, se $f^{-1}(\underline{x}) = P\underline{x}+\underline{t}$, con $P \in GL(n,\mathbb{K})$, $\underline{t} \in \mathbb{K}^n$, e $Q = f^{-1}(Q_1) = PQ_1+\underline{t}$ è un centro di C, allora $A(PQ_1+\underline{t}) = -\underline{b}$. I centri di $f(C)$ sono le soluzioni del sistema lineare $P^\top AP\underline{x} = -P^\top A\underline{t} - P^\top \underline{b}$ ed essendo $APQ_1 = -A\underline{t} - \underline{b}$ si nota che Q_1 è soluzione.

Osserviamo che la forma matriciale a cui siamo arrivati è ottenuta usando solo l'equivalenza affine e possiamo ancora usare la proporzionalità.
Se il termine noto $\underline{b}^\top O+c \neq 0$, allora, dividendo la matrice per il termine noto, C si rappresenta con una matrice del tipo $\begin{pmatrix} D & 0 \\ 0 & 1 \end{pmatrix}$, altrimenti, C si rappresenta con una matrice del tipo $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ e, usando la proporzionalità, possiamo assumere che l'elemento di posto $(1,1)$ della matrice D valga 1.

Diamo un'interpretazione più geometrica.

Interpretando le matrici che rappresentano le affinità di $\mathbb{A}^n$ come matrici di cambio di base da una base $\mathcal{B} = \{\underline{v}_1,\dots,\underline{v}_{n+1}\}$ alla base canonica di $\mathbb{K}^{n+1}$ ($\underline{v}_i$ è la colonna i-ma della matrice rappresentativa) allora $\underline{v}_1,\dots,\underline{v}_n$ sono una base di $W_n = \text{Span}(\underline{e}_1,\dots,\underline{e}_n) = Giac(H)$ e $\underline{v}_{n+1} \in H$. Chiamiamo una tale base *adattata* ad H.
Viceversa, data una base $\mathcal{B} = \{\underline{v}_1,\dots,\underline{v}_{n+1}\}$ adattata ad H, la matrice con colonna i-ma $\underline{v}_i$ rappresenta una affinità di $\mathbb{A}^n$.

Se la quadrica C è rappresentata dalla matrice M, consideriamo il prodotto scalare Φ_M su $\mathbb{K}^{n+1}$ (la cui matrice nella base canonica, che è adattata ad H, è M).
Notiamo che $(\Phi_M)_{|_{W_n}} \neq 0$ (si rappresenta tramite la matrice non nulla A).
Allora, le matrici che rappresentano le quadriche affinemente equivalenti a C,

ottenute agendo su C con una affinità di $\mathbb{A}^n$, sono esattamente le matrici che rappresentano Φ_M nelle basi adattate ad H.
Cercare quindi una affinità di $\mathbb{A}^n$ che porti C in una quadrica con equazione "semplice" equivale a cercare una base di $\mathbb{K}^{n+1}$ adattata ad H in cui la matrice di Φ_M sia "semplice".

Vediamo per prima cosa quando è possibile ottenere una matrice diagonale, ovvero quando esiste una base adattata ad H ortogonale per Φ_M (sappiamo che questo succede se e solo se la quadrica è a centro): in tal caso, l'ultimo vettore di tale base deve essere ortogonale a W_n (qui e nel seguito gli ortogonali sono fatti tramite Φ_M).

Supponiamo allora che $W_n^\perp \cap H \neq \emptyset$ e fissiamo $\underline{v}_{n+1}$ in tale intersezione.
Allora, prendendo una base $\underline{v}_1, \dots, \underline{v}_n$ di W_n ortogonale per Φ_M, e unendo $\underline{v}_{n+1}$ otteniamo una base di $\mathbb{K}^{n+1}$ adattata ad H ortogonale per Φ_M, per cui la matrice associata a Φ_M in questa base è diagonale.
Notiamo che $W_n^\perp \cap H = \{\binom{\underline{x}}{1} \in \mathbb{K}^{n+1} |\ A\underline{x} + \underline{b} = \underline{0}\}$, quindi è non vuoto se e solo se la quardrica è a centro. Più precisamente, scrivendo $\underline{v}_{n+1} = \binom{\underline{x}_{n+1}}{1}$, $\underline{x}_{n+1}$ è un centro della quadrica.

Se invece $W_n^\perp \cap H = \emptyset$, allora $W_n^\perp \subset Giac(H) = W_n$. Infatti, essendo H un iperpiano affine in $\mathbb{A}^{n+1}$, ogni sottospazio affine di $\mathbb{A}^{n+1}$ con giacitura non contenuta in $Giac(H)$ interseca H.
Ne segue che $Rad(\Phi_M) \subset W_n^\perp \subset W_n$ e

$$\begin{aligned}\dim Rad((\Phi_M)_{|_{W_n}}) &= \dim W_n \cap W_n^\perp = \dim W_n^\perp = \\ &= 1 + \dim(W_n \cap Rad(\Phi_M)) = 1 + \dim Rad(\Phi_M).\end{aligned}$$

Esiste quindi $\underline{v}_n \in Rad((\Phi_M)_{|_{W_n}}) \setminus Rad(\Phi_M)$, e di conseguenza esiste anche $\underline{v}_{n+1} \in \mathbb{K}^{n+1}$ non ortogonale a $\underline{v}_n$.
Notiamo che $\underline{v}_n^\perp = W_n$, per cui $\underline{v}_{n+1} \notin W_n$, e quindi $U = \mathrm{Span}(\underline{v}_n, \underline{v}_{n+1})$ è un piano iperbolico. Quindi possiamo supporre $\underline{v}_{n+1}$ isotropo e, a meno di riscalarlo, $\underline{v}_{n+1} \in H$; inoltre, riscalando opportunamente $\underline{v}_n$, possiamo supporre che $\Phi_M(\underline{v}_n, \underline{v}_{n+1}) = 1/2$.
Notiamo infine che $U^\perp \subset W_n$ è un supplementare di U.
Scegliendo una base $\underline{v}_1, \dots, \underline{v}_{n-1}$ di $U^\perp$ ortogonale per Φ_M ed aggiungendo $\underline{v}_n$ e $\underline{v}_{n+1}$, si ottiene una base di $\mathbb{K}^{n+1}$ adattata ad H tale che la matrice di Φ_M in tale base è del tipo $diag(D', \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix})$, con $D' \neq 0$ diagonale (e, poiché $A \neq 0$, possiamo fare in modo che $\underline{D'}^{1}_{1} \neq 0$).

Otteniamo quindi che ogni quadrica è affinemente equivalente ad una quadrica la cui equazione è data da una matrice del tipo $diag(D, \alpha)$ o $diag(D, \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix})$ con $\alpha \in \mathbb{K}$ e D diagonale della taglia opportuna con coefficiente di posto $(1,1)$ non nullo. Usando anche la proporzionalità, si ottengono matrici del tipo $diag(D, 1)$,

$diag(D,0)$ o $diag(D, \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix})$, dove nel scondo e terzo caso, possiamo ancora usare la proporzionalità per rendere uguale a 1 il coefficiente di posto $(1,1)$ di D (nel terzo caso si deve anche riscalare opportunamente $\underline{v}_n$).

Rispettivamente, se $D = diag(\lambda_1, \dots, \lambda_h, 0, \dots, 0)$, con i $\lambda_i \in \mathbb{K}$ non nulli, otteniamo equazioni del tipo $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2 + 1$, $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2$, con $1 \leq h \leq n$, oppure del tipo $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2 + t_n$ con $1 \leq h \leq n-1$ ($\lambda_1 = 1$ nel secondo e terzo caso).

Osserviamo che nel caso a centro si ha

$$\operatorname{rnk} \mathcal{M}(p) = \operatorname{rnk} \mathcal{A}(p) \text{ o } \operatorname{rnk} \mathcal{M}(p) = \operatorname{rnk} \mathcal{A}(p) + 1,$$

mentre nel caso non a centro

$$\operatorname{rnk} \mathcal{M}(p) = \operatorname{rnk} \mathcal{A}(p) + 2.$$

Per l'indice di Witt, invece, nel caso a centro abbiamo

$$W(\mathcal{M}(p)) = W(\mathcal{A}(p)) \text{ o } W(\mathcal{M}(p)) = W(\mathcal{A}(p)) + 1,$$

mentre nel caso non a centro

$$W(\mathcal{M}(p)) = W(\mathcal{A}(p)).$$

Dalla discussione fatta, vediamo che la classificazione delle quadriche di $\mathbb{A}^n$ a meno di equivalenza affine è completa se riusciamo a classificare le possibili forme matriciali diagonali dei prodotti scalari su $\mathbb{K}^n$, ovvero se abbiamo invarianti completi per isometria su $PS(\mathbb{K}^n)$.

Una quadrica C rappresentata dalla matrice M si dice *non degenere* se il prodotto scalare Φ_M è non degenere, ovvero se $\det M \neq 0$. Altrimenti, la quadrica si dice *degenere*.

Le quadriche degeneri sono di due tipi:
- se la parte quadratica $\mathcal{A}(p)$ è invertibile allora siamo nel caso a centro e la forma normale ha termine noto nullo, ovvero abbiamo una equazione del tipo $\lambda_1 t_1^2 + \cdots + \lambda_n t_n^2$, il cui luogo di zeri è un *cono* di vertice $\underline{0}$. Intersecando con l'iperpiano affine E di equazione $t_n - k$, $k \in \mathbb{K}$, $k \neq 0$, otteniamo una quadrica non degenere Q in E e si dice che C è un cono su Q (il cono è dato dalle rette per il vertice e un punto di Q e dall'intersezione del cono con la giacitura di E).
- altrimenti sulla diagonale della matrice D abbiamo almeno uno 0 (che supponiamo nel posto (n,n)), ovvero abbiamo una equazione del tipo $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2$ o $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2 + 1$ con $h < n$, oppure una equazione del tipo $\lambda_1 t_1^2 + \cdots + \lambda_h t_h^2 + t_n$ con $h < n-1$. In entrambi i casi si trata di un *cilindro* (la coordinata $h+1$-ma può assumemere qualsiasi valore). Intersecando con gli iperpiani affini di equazione $t_{h+1} - k$, $k \in \mathbb{K}$, si ottiene sempre la stessa quadrica (detta *base* o *generatrice* del cilindro) che, se è degenere, è a sua volta un cono su una quadrica

non degenere o un cilindro con base una quadrica in dimensione $n-2$ e possiamo reiterare questa descrizione.
Quindi una quadrica è non degenere, oppure è una combinazione di coni e/o cilindri su una quadrica non degenere in dimensione minore.
Per la classificazione quindi, è sufficiente classificare le quadriche non degeneri (che essenzialmente è come dire che per classificare i prodotti scalari basta classificare i non degeneri).

Possiamo allora concludere la classificazione delle quadriche se $\mathbb{K} = \mathbb{C}$ o $\mathbb{K} = \mathbb{R}$ usando le forme matriciali date dalle basi ortogonali normalizzate. Le matrici e le equazioni risultanti si dicono in *forma normale*.

Per $\mathbb{K} = \mathbb{C}$, le possibili matrici in forma normale sono:

$$diag(I_k, 0_{n-k+1}) = \begin{pmatrix} I_k & 0 \\ 0 & 0_{n-k+1} \end{pmatrix}$$

che corrisponde all'equazione in forma normale $t_1^2 + \cdots + t_k^2$ ed è completamente determinata dal fatto che $k = \operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p)$;

$$diag(I_k, 0_{n-k}, 1) = \begin{pmatrix} I_k & 0 & 0 \\ 0 & 0_{n-k} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

che corrisponde all'equazione in forma normale $t_1^2 + \cdots + t_k^2 + 1$ ed è completamente determinata dal fatto che $k = \operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p) - 1$;

$$diag(I_k, 0_{n-k-1}, \left(\begin{smallmatrix} 0 & 1/2 \\ 1/2 & 0 \end{smallmatrix}\right)) = \begin{pmatrix} I_k & 0 & & 0 & \\ 0 & 0_{n-k-1} & & 0 & \\ & & 0 & & 1/2 \\ 0 & 0 & & 1/2 & 0 \end{pmatrix}$$

che corrisponde all'equazione in forma normale $t_1^2 + \cdots + t_k^2 + t_n$ ed è completamente determinata dal fatto che $k = \operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p) - 2$.

Otteniamo che la coppia $(\operatorname{rnk} \mathcal{A}(p), \operatorname{rnk} \mathcal{M}(p))$ è un invariante completo per l'equivalenza affine delle quadriche in $\mathbb{A}^n = \mathbb{A}(\mathbb{C}^n)$.

Osserviamo che, a meno di equivalenza affine, esistono un unico modello non degenere a centro (quello la cui forma normale è I_{n+1} e la cui equazione normale è $t_1^2 + \cdots + t_n^2 + 1$) ed un unico modello non degenere non a centro (quello la cui forma normale è $diag(I_{n-1}, \left(\begin{smallmatrix} 0 & 1/2 \\ 1/2 & 0 \end{smallmatrix}\right))$ e la cui equazione normale è $t_1^2 + \cdots + t_{n-1}^2 + t_n$). Le coniche affinemente equivalenti al primo si dicono *ellissoidi complessi*, quelle affinemente equivalenti al secondo *paraboloidi complessi*.
Ogni quadrica non degenere complessa è quindi un ellissoide o un paraboloide.

Per $\mathbb{K} = \mathbb{R}$, le possibili matrici in forma normale sono:

$$diag(I_p, -I_q, 0_{n-p-q+1}) = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_{n-p-q+1} \end{pmatrix}, \quad p > q$$

che corrisponde all'equazione $t_1^2 + \cdots + t_p^2 - t_{p+1}^2 - \cdots - t_{p+q}^2$ ed è completamente determinata dal fatto che $\operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p) = p + q$ e $W(\mathcal{A}(p)) = n - p$ (infatti, se $\operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p)$ siamo nel caso a centro con termine noto nullo e $\operatorname{rnk} \mathcal{A}(p)$ dice quanti zero ci sono sulla diagonale della parte quadratica, poi $W(\mathcal{A}(p))$ completa, a meno del segno, la parte quadratica);

$$diag(I_p, -I_q, 0_{n-p-q}, 1) = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & -I_q & 0 & 0 \\ 0 & 0 & 0_{n-p-q} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

che corrisponde all'equazione $t_1^2 + \cdots + t_p^2 - t_{p+1}^2 - \cdots - t_{p+q}^2 + 1$ ed è completamente determinata dal fatto che $\operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p) - 1 = p + q$, $W(\mathcal{A}(p)) = \min(p, q) + (n - p - q)$, $W(\mathcal{M}(p)) = W(\mathcal{A}(p))$ se $p \geq q$, $W(\mathcal{M}(p)) = W(\mathcal{A}(p)) + 1$ se $p < q$ (infatti, se $\operatorname{rnk} \mathcal{M}(p) = \operatorname{rnk} \mathcal{A}(p) + 1$ siamo nel caso a centro con termine noto non nullo e $\operatorname{rnk} \mathcal{A}(p)$ dice quanti zero ci sono sulla diagonale della parte quadratica, poi $W(\mathcal{A}(p))$ completa, a meno del segno, la parte quadratica e $W(\mathcal{M}(p))$ determina chi è il maggiore tra p e q);

$$diag(I_p, -I_q, 0_{n-p-q-1}, \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}) = \begin{pmatrix} I_p & 0 & 0 & 0 & 0 \\ 0 & -I_q & 0 & 0 & 0 \\ 0 & 0 & 0_{n-p-q-1} & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{1}{2} \\ 0 & 0 & \cdots & \frac{1}{2} & 0 \end{pmatrix}, \quad p > q$$

che corrisponde all'equazione $t_1^2 + \cdots + t_p^2 - t_{p+1}^2 - \cdots - t_{p+q}^2 + t_n$ ed è completamente determinata dal fatto che $\operatorname{rnk} \mathcal{A}(p) = \operatorname{rnk} \mathcal{M}(p) - 2 = p + q$, $W(\mathcal{A}(p)) = n - p$ (infatti, se $\operatorname{rnk} \mathcal{M}(p) = \operatorname{rnk} \mathcal{A}(p) + 2$ siamo nel caso non a centro e $\operatorname{rnk} \mathcal{A}(p)$ dice quanti zero ci sono sulla diagonale della parte quadratica, poi $W(\mathcal{A}(p))$ completa, a meno del segno, la parte quadratica).

Otteniamo che la quadrupla $(\operatorname{rnk} \mathcal{A}(p), \operatorname{rnk} \mathcal{M}(p), W(\mathcal{A}(p)), W(\mathcal{M}(p)))$ è un invariante completo per l'equivalenza affine delle quadriche in $\mathbb{A}^n = \mathbb{A}(\mathbb{R}^n)$.

Oserviamo che, a meno di equivalenza affine, esistono diversi modelli non degeneri:
- quelli a centro hanno equazione normale del tipo $t_1^2 + \cdots + t_p^2 - t_{p+1}^2 - \cdots - t_n^2 + 1$. Per $p = n$ il supporto è vuoto (corrisponde al caso in cui M è definita). Per $p = 0$ otteniamo la sfera unitaria in $\mathbb{R}^n$ e le quadriche in questa classe di equivalenza affine si dicono *ellissoidi reali*. Per $0 < p < n$ si ottengono quadriche con supporto non vuoto dette *iperboloidi reali*.

- quelli non a centro hanno equazione normale del tipo $t_1^2 + \cdots + t_p^2 - t_{p+1}^2 - \cdots - t_n^2 + t_n$. Si ottengono quadriche con supporto non vuoto dette *paraboloidi reali*.

Vediamo nel dettaglio il caso delle quadriche di $\mathbb{R}$.

Nel caso non degenere otteniamo i modelli $t_1^2 + 1$ ($\mathcal{M}(p)$ definita) che ha luogo di zeri vuoto e $-t_1^2 + 1$ ($\mathcal{M}(p)$ non definita) che ha luogo di zeri $\{\pm 1\}$; nel caso degenere otteniamo t_1^2 che ha per luogo di zeri $\{\underline{0}\}$. I supporti delle quadriche di $\mathbb{R}$ sono quindi dati dall'insieme vuoto, da una coppia di punti e da un punto "doppio".
Osserviamo che tutte le quadriche sono a centro che è sempre unico: per la quadrica non degenere non vuota è il punto medio del supporto, per il punto doppio è il punto stesso (ma anche la quadrica a supporto vuoto ha un centro!). Notiamo che questo è coerente con quanto è noto sulle equazioni polinomiali quadratiche reali: $p(t_1) = at_1^2 + bt_1 + c$, $a \neq 0$, ha $0, 1,$ o 2 radici reali. La matrice che rappresenta p è $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ che è degenere se $0 = \det M = ac - b^2/4 = -\Delta/4$, ed è definita se $\det M > 0$, cioè se $\Delta < 0$.

Vediamo nel dettaglio il caso delle quadriche di $\mathbb{R}^2$, dette coniche.

Per le coniche non degeneri $\operatorname{rnk} \mathcal{M}(p) = 3$ e si ottengono i seguenti modelli:
• $t_1^2 + t_2^2 + 1$ con luogo di zeri vuoto. La quaterna classificante è $(2, 3, 0, 0)$. Le coniche con questa quaterna sono a centro e hanno supporto vuoto (il centro è dato da un punto);
• $t_1^2 - t_2^2 + 1$ con luogo di zeri una iperbole equilatera. La quaterna classificante è $(2, 3, 1, 1)$. Le coniche con questa quaterna sono a centro hanno per supporto una iperbole (il centro è il centro dell'iperbole);
• $-t_1^2 - t_2^2 + 1$ con luogo di zeri la circonferenza unitaria. La quaterna classificante è $(2, 3, 0, 1)$. Le coniche con questa quaterna sono a centro e hanno per supporto una ellisse (il centro è il centro dell'ellisse);
• $t_1^2 + t_2$ con luogo di zeri una parabola. La quaterna classificante è $(1, 3, 1, 1)$. Le coniche con questa quaterna sono non a centro e hanno per supporto una parabola.

Per le coniche degeneri con $\operatorname{rnk} \mathcal{M}(p) = 2$ si ottengono i seguenti modelli:
• $t_1^2 + t_2^2$ con luogo di zeri $\{\underline{0}\}$. La quaterna classificante è $(2, 2, 0, 1)$. Le coniche con questa quaterna sono a centro e hanno supporto dato da un punto (il centro è il punto);
• $t_1^2 - t_2^2$ con luogo di zeri due rette incidenti. La quaterna classificante è $(2, 2, 1, 2)$. Le coniche con questa quaterna sono a centro e hanno supporto dato da due rette incidenti (il centro è il punto di incidenza);
• $t_1^2 + 1$ con luogo di zeri vuoto. La quaterna classificante è $(1, 2, 1, 1)$. Le coniche con questa quaterna sono a centro e hanno supporto vuoto (il luogo dei centri è una retta);

• $-t_1^2+1$ con luogo di zeri la coppia di rette di equazioni $t_1=\pm 1$. La quaterna classificante è $(1,2,1,2)$. Le coniche con questa quaterna sono a centro e hanno per supporto una coppia di rette parallele (il luogo dei centri è la parallela per il punto medio di un segmento che congiunge le due rette);

Per le coniche degeneri con $\operatorname{rnk}\mathcal{M}(p)=1$ si ottiene il seguente modello:
• t_1^2 con luogo di zeri una retta. La quaterna classificante è $(1,1,1,2)$. Le coniche con questa quaterna hanno per supporto una retta "doppia" (il luogo dei centri è la retta).

Per le coniche degeneri, il punto e la coppia di rette incidenti sono coni sul loro centro; la conica vuota è un cilindro con base la quadrica vuota di $\mathbb{R}$; la coppia di rette parallele è un cilindro caon base la quadrica non degenere di $\mathbb{R}$ data da due punti; la retta doppia è un cilindro con base la quadrica non degenere di $\mathbb{R}$ data da un punto doppio.

Vediamo nel dettaglio il caso delle quadriche di $\mathbb{R}^3$.
Per le quadriche non degeneri $\operatorname{rnk}\mathcal{M}(p)=4$ e si ottengono i seguenti modelli:
• $t_1^2+t_2^2+t_3^2+1$ con luogo di zeri vuoto. La quaterna classificante è $(3,4,0,0)$. Le coniche con questa quaterna sono a centro e hanno supporto vuoto (il centro è dato da un punto);
• $t_1^2+t_2^2-t_3^2+1$ con luogo di zeri ottenuto ruotando una iperbole attorno al suo asse contenente i fuochi. La quaterna classificante è $(3,4,1,1)$. Le coniche con questa quaterna sono a centro hanno per supporto un *iperboloide a due falde* (il centro è il centro dell'iperboloide);
• $t_1^2-t_2^2-t_3^2+1$ con luogo di zeri ottenuto ruotando una iperbole attorno al suo asse non contenente i fuochi. La quaterna classificante è $(3,4,1,2)$. Le coniche con questa quaterna sono a centro e hanno per supporto un *iperboloide ad una falda* (il centro è il centro dell'iperboloide);
• $-t_1^2-t_2^2-t_3^2+1$ con luogo di zeri la sfera unitaria. La quaterna classificante è $(3,4,0,1)$. Le coniche con questa quaterna sono a centro e hanno per supporto un *ellissoide* (il centro è il centro dell'ellissoide);
• $t_1^2+t_2^2+t_3$ con luogo di zeri ottenuto ruotando una parabola attorno al suo asse. La quaterna classificante è $(2,4,1,1)$. Le coniche con questa quaterna sono non a centro e hanno per supporto un *paraboloide ellittico*.
• $t_1^2-t_2^2+t_3$ con luogo di zeri non vuoto. La quaterna classificante è $(2,4,2,2)$. Le coniche con questa quaterna sono non a centro e hanno per supporto un *paraboloide iperbolico* o *sella*.

Per le quadriche degeneri con $\operatorname{rnk}\mathcal{M}(p)=3$ si ottengono i seguenti modelli:
• $t_1^2+t_2^2+t_3^2$ con luogo di zeri $\{\underline{0}\}$. La quaterna classificante è $(3,3,0,1)$. Le coniche con questa quaterna sono a centro e hanno per supporto un punto (il centro è dato dal punto);
• $t_1^2+t_2^2-t_3^2$ con luogo di zeri un cono di centro $\{\underline{0}\}$. La quaterna classificante è $(3,3,1,2)$. Le coniche con questa quaterna sono a centro e hanno per supporto un cono (il centro è il vertice del cono);

• $t_1^2 + t_2^2 + 1$ con luogo di zeri vuoto. La quaterna classificante è $(2, 3, 1, 1)$. Le coniche con questa quaterna sono a centro e hanno supporto vuoto (il centro è dato da un punto);
• $t_1^2 - t_2^2 + 1$ con luogo di zeri un cilindro con base una iperbole equilatera. La quaterna classificante è $(2, 3, 2, 2)$. Le coniche con questa quaterna sono a centro e hanno per supporto un cilindro con base una iperbole (il luogo dei centri è dato da una retta);
• $-t_1^2 - t_2^2 + 1$ con luogo di zeri un cilindro con base una circonferenza. La quaterna classificante è $(2, 3, 1, 2)$. Le coniche con questa quaterna sono a centro e hanno per supporto un cilindro con base una ellisse (il luogo dei centri è dato da una retta);
• $t_1^2 + t_3$ con luogo di zeri un cilindro con base una parabola. La quaterna classificante è $(1, 3, 2, 2)$. Le coniche con questa quaterna sono non a centro e hanno per supporto un cilindro con base una parabola.

Per le quadriche degeneri con $\operatorname{rnk} \mathcal{M}(p) = 2$ si ottengono i seguenti modelli:
• $t_1^2 + t_2^2$ con luogo di zeri l'asse t_3. La quaterna classificante è $(2, 2, 1, 2)$. Le coniche con questa quaterna sono a centro e hanno per supporto una retta (il luogo dei centri è dato dalla retta);
• $t_1^2 - t_2^2$ con luogo di zeri due piani incidenti. La quaterna classificante è $(2, 2, 2, 3)$. Le coniche con questa quaterna sono a centro e hanno per supporto una coppia di piani incidenti (il luogo dei centri è dato dall'inersezione dei due piani);
• $t_1^2 + 1$ con luogo di zeri vuoto. La quaterna classificante è $(1, 2, 2, 2)$. Le coniche con questa quaterna sono a centro e hanno supporto vuoto (il luogo dei centri è un piano);
• $-t_1^2 + 1$ con luogo di zeri due piani paralleli. La quaterna classificante è $(1, 2, 2, 3)$. Le coniche con questa quaterna sono a centro e hanno per supporto una coppia di piani paralleli (il luogo dei centri è dato dal piano parallelo per il punto medio di un segmento che congiunge i due piani).

Per le coniche degeneri con $\operatorname{rnk} \mathcal{M}(p) = 1$ si ottiene il seguente modello:
• t_1^2 con luogo di zeri un piano. La quaterna classificante è $(1, 1, 2, 3)$. Le coniche con questa quaterna sono a centro e hanno per supporto un piano "doppio" (il luogo dei centri è il piano).

Di nuovo, le quadriche degeneri sono coni o cilindri su coniche.