Corsi di Laurea Triennale in Ingegneria Navale e Industriale

Geometria

Docente: Prof. Daniele Zuddas

Indice

Lezione 1: Richiami su insiemi, numeri, funzioni 2
Lezione 2: Richiami sulle funzioni 6
Lezione 3: Campi e spazi vettoriali numerici 8
Lezione 4: Numeri complessi 12
Lezione 5: Numeri complessi in forma polare 15
Lezione 6: Polinomi 20
Lezione 7: Zeri di polinomi 22
Lezione 8: Spazi vettoriali numerici complessi 25
Lezione 9: Matrici invertibili 28
Lezione 10: Sistemi lineari 32
Lezione 11: Risoluzione di un sistema lineare 38
Lezione 12: Determinante 41
Lezione 13: Matrice inversa 45
Lezione 14: Spazi vettoriali 47
Lezione 15: Sottospazi vettoriali 51
Lezione 16: Coordinate 55
Lezione 17: Dimensione 57
Lezione 18: Equazioni di sottospazi vettoriali 60
Lezione 19: Applicazioni lineari 63
Lezione 20: Nucleo e immagine 65
Lezione 21: Isomorfismi di spazi vettoriali 68
Lezione 22: Calcolo della matrice inversa 70
Lezione 23: Matrice del cambio di base 74
Lezione 24: Determinante di un endomorfismo 77
Lezione 25: Diagonalizzazione 79
Lezione 26: Forme bilineari 84
Lezione 27: Spazi vettoriali Euclidei 87
Lezione 28: Spazi vettoriali Euclidei 90
Lezione 29: Complemento ortogonale 94
Lezione 30: Geometria affine 98
Lezione 31: Geometria affine 100
Lezione 32: Elementi di Geometria Euclidea 104

Insiemi

Insieme. Collezione di *elementi* (oggetti) non ordinati e senza ripetizioni.
Un insieme si può assegnare elencandone gli elementi tra parentesi graffe

$$\{1\}, \quad \{3, -7, 42\}, \quad \{2, 12, 132, -1, 0\}$$

Esiste un insieme senza elementi, detto *insieme vuoto* e denotato con $\emptyset$
In molti casi un insieme si assegna specificando una proprietà $\mathcal{P}$ che ne caratterizza gli elementi. In simboli

$$X = \{x \mid x \text{ soddisfa } \mathcal{P}\}$$

X è quindi l'insieme di tutti e soli gli elementi che hanno la proprietà $\mathcal{P}$, qualunque essa sia.

Esempi. $\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$ insieme dei *numeri naturali*. Non è possibile elencarli tutti essendo infiniti.
$P = \{2n \mid n \in \mathbb{N}\}$ insieme dei numeri naturali pari.
$D = \{2n + 1 \mid n \in \mathbb{N}\}$ insieme dei numeri naturali dispari.

Per affermare che un elemento x appartiene all'insieme X si scrive

$$x \in X$$

e si legge "x appartiene a X". Il simbolo di non appartenenza è $\notin$.
Per esempio $2 \in \{5, 3, 10, 2, 0\}$ ma $1 \notin \{5, 3, 10, 2, 0\}$.
Per affermare che un insieme X è contenuto in un insieme Y si scrive

$$X \subset Y$$

e si legge "X è sottoinsieme di Y" o anche "X è contenuto in Y".
Precisamente questo significa che ogni elemento di X appartiene a Y.
Non confondere inclusione (di sottoinsiemi) e appartenenza (di elementi).
Per esempio

$$\{0, 3\} \subset \{5, 3, 10, 2, 0\}, \ 3 \in \{5, 3\}, \ \{3\} \subset \{5, 3\}, \ \{3\} \notin \{5, 3\}$$

$$\{0, 3\} \notin \{10, 2, 0\}, \ \{0, 3\} \in \{10, 2, 0, \{3, 0\}\}.$$

Se $X \subset Y$ scriviamo anche $Y \supset X$ (si legge Y contiene X).

Due insiemi X e Y sono uguali (sono lo stesso insieme) se hanno gli stessi elementi: $X = Y$ se e solo se $X \subset Y$ e $Y \subset X$. Ad esempio

$$\{1, 2, 3\} = \{2, 3, 1\} \quad \text{ma} \quad \{1, 2\} \neq \{3, 2\}.$$

N.B. $X \subset Y$ non esclude $X = Y$, pertanto $X \subset X$. Si scrive anche $X \subseteq Y$.
Invece la scrittura $X \subsetneq Y$ significa $X \subset Y$ e $X \neq Y$ (sottoinsieme proprio).

Operazioni tra insiemi

Intersezione. $X \cap Y := \{x \mid x \in X \text{ e } x \in Y\}$
Esempio: $\{1,2,3,5\} \cap \{10,9,1,3,5\} = \{1,3,5\}$.
Si ha $X \cap Y \subset X$ e $X \cap Y \subset Y$.

X e Y sono detti disgiunti se $X \cap Y = \emptyset$, ovvero se X e Y non hanno nessun elemento in comune.

Unione. $X \cup Y := \{x \mid x \in X \text{ o } x \in Y\}$
Esempio: $\{1,2,3,5\} \cup \{10,9,1,3,5\} = \{1,2,3,5,10,9\}$.
Si ha $X \subset X \cup Y$ e $Y \subset X \cup Y$.

Differenza insiemistica. $X - Y = X \setminus Y := \{x \mid x \in X \text{ e } x \notin Y\}$.
Esempi
$\{1,2,3,5\} - \{10,9,3,5\} = \{1,2\}$.
$\{10,9,3,5\} - \{1,2,3,5\} = \{10,9\}$.
Si ha $X - Y \subset X$, $X - X = \emptyset$, $X - \emptyset = X$ e $(X - Y) \cap Y = \emptyset$.
Se $X \cap Y = \emptyset$ allora $X - Y = X$.

Coppie ordinate. Una coppia ordinata (x, y) di elementi (detti componenti) è un insieme in cui conta l'ordine e sono ammesse ripetizioni. Ad esempio $(2,3)$ è diverso da $(3,2)$ ed è diverso da $\{2,3\}$. Anche $(1,1)$ è una coppia ordinata, in questo caso con le due componenti uguali.

n-uple ordinate. Più in generale si possono fare n-uple ordinate

$$(x_1, x_2, \ldots, x_n),$$

cioè insiemi di $n \geqslant 1$ componenti in cui conta l'ordine e sono ammesse ripetizioni. Le 2-uple sono le coppie, le 3-uple sono dette terne, le 4-uple quaterne.

Prodotto cartesiano. $X \times Y = \{(x,y) \mid x \in X \text{ e } y \in Y\}$ è l'insieme di tutte le coppie ordinate in cui la prima componente appartiene al primo insieme e la seconda al secondo insieme.

$$\{1,2,3\} \times \{3,5\} = \{(1,3),(1,5),(2,3),(2,5),(3,3),(3,5)\}$$

Oss. Se X ha m elementi e Y ha n elementi allora $X \times Y$ ha mn elementi.

Più in generale si può definire il prodotto cartesiano di un numero n arbitrario di insiemi $X_1, X_2, \ldots, X_n$

$$X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \ldots, x_n) \mid x_1 \in X_1, x_2 \in X_2, \ldots, x_n \in X_n\}$$

cioè l'insieme di tutte le n-uple ordinate in cui la prima componente appartiene al primo insieme, la seconda al secondo e così via, in tutti i modi possibili.

Si pone $X^n = X \times \cdots \times X$, prodotto di X con sé stesso n volte. Quindi $X^1 = X$, $X^2 = X \times X$, $X^3 = X \times X \times X$, ... Ad esempio

$$\{1,2\}^2 = \{(1,1),(1,2),(2,1),(2,2)\}.$$

Quantificatori. In Matematica e in logica si usano i seguenti quantificatori.

- $\forall$ "per ogni" (quantificatore universale)
- $\exists$ "esiste (almeno uno)" (quantificatore esistenziale)
- $\exists!$ "esiste unico" (quantificatore unico)
- $\nexists$ "non esiste"'

Numeri

Numeri naturali. L'insieme $\mathbb{N}$ dei numeri naturali l'abbiamo già incontrato. In questo insieme è definita l'addizione e la moltiplicazione, ma non sempre si può fare la sottrazione.

Numeri interi. $\mathbb{Z} = \{\dots, -4, -2, -1, 0, 1, 2, 3, 4, \dots\}$ è l'insieme dei *numeri interi*. Su $\mathbb{Z}$ sono definite l'addizione, la moltiplicazione e la sottrazione.

Numeri razionali. $\mathbb{Q} = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z}, b \neq 0 \right\}$ è l'insieme dei *numeri razionali*, quozienti di numeri interi con $\frac{a}{b} = \frac{a'}{b'}$ se e solo se $ab' = a'b$.

Su $\mathbb{Q}$ sono definite addizione, moltiplicazione, sottrazione e divisione per un numero diverso da 0.

Numeri reali. $\mathbb{R}$ è l'insieme dei *numeri reali*, ad esempio

$$-3, 0, 5, \frac{1}{2}, -\frac{17}{6}, \sqrt{2}, \frac{\sqrt[3]{17}}{\sqrt{5}}, \pi, e, \dots$$

Come su $\mathbb{Q}$, anche su $\mathbb{R}$ sono definite addizione, moltiplicazione, sottrazione e divisione per un numero diverso da 0. Si ha

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Funzioni

Dati due insiemi X e Y, una *funzione* o *applicazione* o *mappa*

$$f : X \to Y$$

è costituita da tre insiemi X, Y e $G_f \subset X \times Y$ che soddisfa la proprietà seguente: per ogni $x \in X$ esiste un unico $y \in Y$ t.c. $(x, y) \in G_f$ e si scrive

$$y = f(x).$$

X si chiama *dominio*, Y *codominio* e G_f si chiama *grafico* di f. Pertanto $(x, f(x)) \in G_f$ e possiamo anche scrivere

$$G_f = \{(x, f(x)) \mid x \in X\}.$$

L'elemento $f(x) \in Y$ si dice *immagine* di $x \in X$ tramite f.

Oss. Per una funzione dominio, codominio e grafico non vanno "trovati" essendo dati a priori. Il grafico rappresenta la "legge" che permette di associare ad ogni elemento del dominio il corrispondente elemento del codominio.

Per definire una funzione occorre specificare dominio, codominio e grafico o equivalentemente assegnare ad ogni elemento del dominio la sua immagine nel codominio. Spesso una funzone è definita mediante una formula.

Esempio.

$$\begin{aligned} h : \{1, 2, 3\} &\to \{1, 2\} \\ 1 &\mapsto 2 \\ 2 &\mapsto 1 \\ 3 &\mapsto 2 \end{aligned}$$

è la funzione con grafico $G_h = \{(1, 2), (2, 1), (3, 2)\}$.

Esempio.

$$\begin{aligned} g : \mathbb{N} &\to \mathbb{Z} \\ g(n) &= 3n - n^2 \end{aligned}$$

Per ogni insieme X si definisce la *funzione identica*

$$\begin{aligned} \mathrm{id}_X : X &\to X \\ \mathrm{id}_X(x) &= x \end{aligned}$$

Funzione composta. Date $f : X \to Y$ e $g : Y \to Z$

$$g \circ f : X \to Z \qquad (\textit{composizione di } f \textit{ e } g)$$
$$(g \circ f)(x) = g(f(x))$$

Attenzione all'ordine, non si possono in generale scambiare!
Si ha: $f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f$.

Immagine. Data $f : X \to Y$ si chiama *immagine* di f il sottoinsieme

$$\mathrm{Im}(f) = \{f(x) \mid x \in X\} \subset Y$$

i cui elementi sono tutte e sole le immagini degli elementi di X.

Funzioni iniettive. $f : X \to Y$ è *iniettiva* se

$$\forall\, x_1, x_2 \in X,\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

Equivalentemente:
f iniettiva $\Leftrightarrow$ elementi con la stessa immagine sono uguali.
f iniettiva $\Leftrightarrow$ elementi distinti vanno in elementi distinti.

Esempi.

$$h : \{1,2,3\} \to \{1,2\}$$
$$1 \mapsto 2$$
$$2 \mapsto 1$$
$$3 \mapsto 2$$

non è iniettiva perché $h(1) = h(3) = 2$.

$$k : \{1,2,3\} \to \{1,2,3,4\}$$
$$1 \mapsto 3$$
$$2 \mapsto 1$$
$$3 \mapsto 2$$

è iniettiva

Funzioni suriettive. $f : X \to Y$ è *suriettiva* se

$$\mathrm{Im}(f) = Y.$$

Equivalentemente:
f suriettiva $\Leftrightarrow$ ogni elemento del codominio è immagine di almeno un elemento del dominio.

Negli esempi precedenti, h è suriettiva mentre k no.

Funzioni biiettive. $f : X \to Y$ è *biiettiva* (o *biunivoca*) se è iniettiva e suriettiva. Per esempio h e k non sono biiettive mentre lo è la funzione

$$p : \{1,2,3\} \to \{1,2,3\}$$
$$1 \mapsto 1$$
$$2 \mapsto 3$$
$$3 \mapsto 2$$

Funzioni invertibili. Una funzione $f : X \to Y$ è *invertibile* se esiste una funzione $h : Y \to X$ t.c. $h \circ f = \mathrm{id}_X$ e $f \circ h = \mathrm{id}_Y$. Se tale h esiste allora è unica e si denota con f^{-1} e si ha quindi $f^{-1} \circ f = \mathrm{id}_X$ e $f \circ f^{-1} = \mathrm{id}_Y$. Si ha che una funzione è invertibile se e solo se è biiettiva. In particolare per dimostrare che una funzione è biiettiva è sufficiente esibire l'inversa. Nel nostro esempio si ha $p^{-1} = p$.

Oss. Per ogni insieme X si ha $\mathrm{id}_X^{-1} = \mathrm{id}_X$, cioè l'identità è biiettiva e coincide con la sua inversa.

Cardinalità. Dato un insieme finito X (con un numero finito di elementi) si chiama *cardinalità* di X il numero dei suoi elementi e si denota $|X|$ o anche $\#(X)$. Se X è infinito poniamo $|X| = \infty$. Ad esempio $|\emptyset| = 0$, $|\{3\}| = 1$, $\{1, 2, 9, 12\} = 4$, $|\mathbb{N}| = \infty$.

Siano ora X e Y insiemi finiti. Si ha

$$|X \cup Y| = |X| + |Y| - |X \cap Y|$$

infatti nel contare gli elementi dell'unione se sommiamo la cardinalità di X e quella di Y stiamo contando due volte gli elementi dell'intersezione la cui cardinalità va sottratta per avere il risultato corretto.

Oss. Supponiamo X e Y insiemi finiti.

1) $\exists\ f : X \to Y$ iniettiva $\Leftrightarrow |X| \leqslant |Y|$.

2) $\exists\ f : X \to Y$ suriettiva $\Leftrightarrow |X| \geqslant |Y|$.

3) $\exists\ f : X \to Y$ biiettiva $\Leftrightarrow |X| = |Y|$.

4) Se $|X| = |Y|$ allora $f : X \to Y$ iniettiva $\Leftrightarrow$ suriettiva $\Leftrightarrow$ biiettiva.

Esempio.

$$\begin{aligned} f : \{1, 2, 3, 4\} &\to \{1, 2, 3, 4\} \\ 1 &\mapsto 3 \\ 2 &\mapsto 1 \\ 3 &\mapsto 4 \\ 4 &\mapsto 2 \end{aligned}$$

biiettiva in quanto suriettiva. Calcoliamo f^{-1}, leggendo da destra a sinistra:

$$\begin{aligned} f^{-1} : \{1, 2, 3, 4\} &\to \{1, 2, 3, 4\} \\ 1 &\mapsto 2 \\ 2 &\mapsto 4 \\ 3 &\mapsto 1 \\ 4 &\mapsto 3 \end{aligned}$$

Mappa d'inclusione. Se $A \subseteq X$ definiamo la *mappa d'inclusione*

$$\begin{aligned} i_A &: A \to X \\ i_A(a) &= a, \quad \text{per ogni } a \in A \end{aligned}$$

Oss. $i_A : A \to X$ è iniettiva.

Campi

$\mathbb{R}$ è munito delle operazioni $+$ e $\cdot$ che soddisfano le seguenti $\forall a, b, c \in \mathbb{R}$:

(*1*)	$(a+b)+c = a+(b+c)$	Proprietà associativa per l'addizione
(*2*)	$a+0 = 0+a = a$	0 è elemento neutro per l'addizione
(*3*)	$a-a = 0$	$-a$ è l'inverso additivo (opposto) di a
(*4*)	$a+b = b+a$	Proprietà commutativa per l'addizione
(*5*)	$(ab)c = a(bc)$	Proprietà associativa per la moltiplicazione
(*6*)	$a \cdot 1 = 1 \cdot a = a$	1 è elemento neutro per la moltiplicazione
(*7*)	$a \cdot a^{-1} = 1$	a^{-1} è l'inverso moltiplicativo di $a \neq 0$
(*8*)	$ab = ba$	Proprietà commutativa per la moltiplicazione
(*9*)	$a(b+c) = ab+ac$	Proprietà distributiva della moltiplicazione rispetto all'addizione

Un *campo* è un insieme non vuoto $\mathbb{K}$ munito di due operazioni denotate formalmente $+$ e $\cdot$ con due elementi $0 \neq 1$, che soddisfano le proprietà precedenti. Pertanto $\mathbb{R}$ è un campo, detto *campo dei numeri reali*. Si osservi che la proprietà distributiva è l'unica dove intervengono entrambe le operazioni $+$ e $\cdot$ e ne stabilisce un legame.

Esistono molti esempi di campi. In questo corso useremo il campo reale $\mathbb{R}$ e il campo complesso $\mathbb{C}$ (che sarà definito più avanti). Si osservi che anche $\mathbb{Q}$ è un campo rispetto alle usuali operazioni di addizione e moltiplicazione (*campo dei numeri razionali*).

L'insieme dei numeri interi $\mathbb{Z}$ non è un campo perché non soddisfa la proprietà (*7*): una tale struttura algebrica è detta *anello commutativo*, che non studieremo in questo corso.

Spazi vettoriali numerici reali

Partiamo da un esempio importante:

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

insieme di tutte le coppie ordinate (x, y) di numeri reali, dette *vettori*.

Consideriamo due operazioni su $\mathbb{R}^2$.

Addizione. $\forall\, (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ definiamo

$$(x_1, y_1) + (x_2, y_2) \stackrel{\text{def}}{=} (x_1 + x_2, y_1 + y_2).$$

Si osservi che il $+$ a primo membro è l'operazione che stiamo definendo, mentre quelli a secondo membro sono l'usuale addizione di numeri reali. In sostanza per sommare due vettori di $\mathbb{R}^2$ si sommano le componenti corrispondenti, ottenendo così un terzo vettore di $\mathbb{R}^2$.

Esempio. $(3, -2) + (5, 4) = (8, 2)$.

Oss. L'addizione di vettori di $\mathbb{R}^2$ è associativa e commutativa, dato che lo è l'addizione di numeri reali.

Poniamo

$$0_{\mathbb{R}^2} := (0, 0)$$
$$-(x, y) := (-x, -y) \quad \forall\, (x, y) \in \mathbb{R}^2$$

si ha che $0_{\mathbb{R}^2}$ è elemento neutro per l'addizione

$$(x, y) + 0_{\mathbb{R}^2} = (x, y) \quad \forall\, (x, y) \in \mathbb{R}^2$$

e $-(x, y)$ è opposto additivo di (x, y)

$$(x, y) - (x, y) = 0_{\mathbb{R}^2}.$$

Moltiplicazione scalare. $\forall\, \alpha \in \mathbb{R}$ e $\forall\, (x, y) \in \mathbb{R}^2$ definiamo

$$\alpha(x, y) \stackrel{\text{def}}{=} (\alpha x, \alpha y).$$

Si osservi che a primo membro figura l'operazione che stiamo definendo, il prodotto di un numero per un vettore, mentre a secondo membro utilizziamo il prodotto di numeri. I numeri li chiameremo anche *scalari*.

Esempio. $3(2, -5) = (6, -15)$.
$-2(1, 1) + 3(2, 0) + (-3, 1) = (1, -1)$.

Un vettore $(x, y) \in \mathbb{R}^2$ si può rappresentare graficamente anche come una freccia nel piano cartesiano con origine in $0_{\mathbb{R}^2}$ e terminale nel punto di coordinate (x, y), fornendo un modello algebrico per i vettori geometrici piani usati anche in Fisica. La somma definita sopra coincide con quella ottenuta dalla regola del parallelogramma.

Le operazioni sulle coppie si generalizzano a

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ volte}} = \Big\{\underbrace{(x_1, \ldots, x_n)}_{\text{vettore}} \;\Big|\; \underbrace{x_1, \ldots, x_n}_{\text{coordinate}} \in \mathbb{R}\Big\},$$

insieme di tutte le n-uple ordinate di numeri reali, $n \in \mathbb{N}$ fissato. Chiamiamo *vettori* gli elementi di $\mathbb{R}^n$. x_i sono dette *coordinate* o *componenti*.

Addizione. $\forall\,(x_1,\ldots,x_n),(y_1,\ldots,y_n)\in\mathbb{R}^n$ definiamo

$$(x_1,\ldots,x_n)+(y_1,\ldots,y_n)\stackrel{\text{def}}{=}(x_1+y_1,\ldots,x_n+y_n).$$

Moltiplicazione scalare. $\forall\,\alpha\in\mathbb{R}$, $\forall\,(x_1,\ldots,x_n)\in\mathbb{R}^n$ definiamo

$$\alpha(x_1,\ldots,x_n)\stackrel{\text{def}}{=}(\alpha x_1,\ldots,\alpha x_n).$$

Def. $\mathbb{R}^n$ munito delle operazioni di addizione e moltiplicazione scalare definite sopra è detto *spazio vettoriale numerico reale di dimensione n*.

Vettore nullo. È il vettore avente tutte le componenti nulle

$$0_{\mathbb{R}^n}\stackrel{\text{def}}{=}(0,\ldots,0)\in\mathbb{R}^n.$$

Spesso indicheremo il vettore nullo di $\mathbb{R}^n$ semplicemente con 0.

Vettore opposto. Per ogni $(x_1,\ldots,x_n)\in\mathbb{R}^n$ definiamo il *vettore opposto* ottenuto cambiando segno a tutte le componenti

$$-(x_1,\ldots,x_n)\stackrel{\text{def}}{=}(-x_1,\ldots,-x_n)$$

Naturalmente le operazioni di somma e moltiplicazione scalare si possono combinare in formule più complesse.

Esempio. $(3,-2,0)+2(1,1,4)=(5,0,8)$.

Spesso è preferibile scrivere i vettori di $\mathbb{R}^n$ come *vettori colonna* anziché come *vettori riga* come fatto finora, ovvero

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\quad\text{anziché}\quad(x_1,\ldots,x_n).$$

Le operazioni definite sopra si applicano anche ai vettori colonna, senza modifiche sostanziali.

Esempio. Svolgiamo il seguente calcolo in $\mathbb{R}^4$:

$$-3\begin{pmatrix}2\\3\\1\\0\end{pmatrix}+2\begin{pmatrix}5\\1\\-1\\4\end{pmatrix}=\begin{pmatrix}4\\-7\\-5\\8\end{pmatrix}$$

Più in generale si possono fare *combinazioni lineari* di vettori, cioè espressioni del tipo

$$\alpha_1 v_1+\cdots+\alpha_k v_k$$

dove $k\in\mathbb{N}$, $\alpha_1,\ldots,\alpha_k\in\mathbb{R}$ e $v_1,\ldots,v_k\in\mathbb{R}^n$ (ciascun v_i è un vettore di $\mathbb{R}^n$, quindi una n-upla ordinata di numeri reali). Gli scalari $\alpha_1,\ldots,\alpha_k$ sono detti *coefficienti* della combinazione lineare dei vettori $v_1,\ldots,v_k$.

Esempio.

$$\begin{pmatrix} 0 \\ -3 \\ 1 \\ 7 \end{pmatrix} - \begin{pmatrix} 5 \\ 1 \\ -1 \\ 4 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 0 \\ -3 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ -10 \\ 3 \end{pmatrix}$$

quindi il vettore a secondo membro è combinazione lineare dei tre vettori a primo membro, con coefficienti rispettivamente $1, -1, 4$.

Oss. $\mathbb{R}^3$ è il modello di spazio tridimensionale nel quale viviamo le esperienze di tutti i giorni. $\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$ è il modello di spazio-tempo quadridimensionale nel quale evolve lo spazio fisico nel tempo.

Numeri complessi

Per ogni numero reale $x \in \mathbb{R}$ si ha $x^2 \geqslant 0$. L'obiettivo di questa lezione è costruire un *campo* nel quale esista un elemento i t.c. $i^2 = -1$.

Nel seguito poniamo $\mathbb{C} := \mathbb{R}^2$. Abbiamo visto su $\mathbb{C}$, ovvero su $\mathbb{R}^2$, l'addizione e la moltiplicazione scalare

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \qquad \forall\, (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$$
$$\alpha(x, y) = (\alpha x, \alpha y), \qquad \forall\, \alpha \in \mathbb{R},\ \forall\, (x, y) \in \mathbb{R}^2$$

Per definizione poniamo

$$i \stackrel{\text{def}}{=} (0, 1) \in \mathbb{C} \qquad \textit{(unità immaginaria).}$$

Le coppie del tipo $(x, 0)$ le consideriamo come numeri reali, scriviamo quindi x anziché $(x, 0)$, $\forall\, x \in \mathbb{R}$.

Si ha quindi

$$(x, y) = (x, 0) + y(0, 1) = x + yi \quad \forall\, (x, y) \in \mathbb{R}^2.$$

Pertanto l'addizione e la moltiplicazione scalare si scrivono come

$$(x_1 + y_1 i) + (x_2 + y_2 i) = (x_1 + x_2) + (y_1 + y_2)i$$
$$\alpha(x + yi) = \alpha x + \alpha y i$$

Teor. *Su $\mathbb{C}$ esiste un'unica struttura di campo t.c.*

1) *l'addizione è quella considerata sopra;*
2) *la moltiplicazione estende la moltiplicazione scalare;*
3) $i^2 = -1$.

Dim. Unicità Supponiamo che una tale struttura di campo esista. Per ogni $x_1 + y_1 i$, $x_2 + y_2 i \in \mathbb{C}$ si ha:

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

Dunque se la moltiplicazione su $\mathbb{C}$ esiste allora deve essere data da questa formula. D'altra parte l'addizione è già stata definita, e questo dimostra l'unicità della struttura di campo.

Esistenza La dimostrazione dell'unicità suggerisce in che modo va definita la moltiplicazione:

$$(x_1 + y_1 i)(x_2 + y_2 i) \stackrel{\text{def}}{=} x_1 x_2 - y_1 y_2 + (x_1 y_2 + y_1 x_2)i.$$

Questa definizione estende la moltiplicazione scalare per un numero reale, come si vede subito ponendo $y_1 = 0$.

Resta da verificare che valgono le proprietà della definizione di campo viste nella lezione precedente. Per l'addizione sappiamo già che le proprietà sono soddisfatte. La verifica per la moltiplicazione è svolta di seguito.

Proprietà associativa.

$$\begin{aligned}&(x_1+y_1i)\big((x_2+y_2i)(x_3+y_3i)\big)=\\&(x_1+y_1i)\big((x_2x_3-y_2y_3)+(x_2y_3+y_2x_3)i\big)=\\&=x_1x_2x_3-x_1y_2y_3-y_1x_2y_3-y_1y_2x_3+\\&+(x_1x_2y_3+x_1y_2x_3+y_1x_2x_3-y_1y_2y_3)i\end{aligned}$$

Con un calcolo simile, lasciato per Esercizio, si verifica che

$$\big((x_1+y_1i)(x_2+y_2i)\big)(x_3+y_3i)$$

è uguale all'espressione precedente, verificando la proprietà associativa.

Elemento neutro moltiplicativo. Ponendo $x_1=1$ e $y_1=0$, si ha subito

$$1(x+yi)=1x+1yi=x+yi.$$

$i^2=-1$. È immediato ponendo $x_1=x_2=0$ e $y_1=y_2=1$.

Inverso. Osserviamo che

$$(x+yi)(x-yi)=x^2+y^2\in\mathbb{R}.$$

Quindi $x+yi\neq 0\Rightarrow x\neq 0$ oppure $y\neq 0\Rightarrow x^2+y^2\neq 0\Rightarrow$

$$(x+yi)\frac{(x-yi)}{x^2+y^2}=1.$$

Dunque

$$(x+yi)^{-1}=\frac{x-yi}{x^2+y^2}=\frac{x}{x^2+y^2}-\frac{y}{x^2+y^2}i$$

Proprietà commutativa. Ciascun addendo non cambia se si scambiano gli indici 1 e 2.

Proprietà distributiva. Si tratta di svolgere il calcolo e verificare che vale questa proprietà. Esercizio. □

Def. $\mathbb{C}$ con questa struttura di campo è detto *campo complesso*. Gli elementi di $\mathbb{C}$ sono detti *numeri complessi*.

Oss. Per definizione, ogni numero complesso $z\in\mathbb{C}$ si scrive in modo unico come $z=x+yi$, con $x,y\in\mathbb{R}$.

Oss. Per ogni $x\in\mathbb{R}$ si ha $x=x+0i\in\mathbb{C}$ e quindi possiamo considerare

$$\mathbb{R}\subset\mathbb{C}.$$

N.B. Generalmente continueremo ad utilizzare la notazione $\mathbb{R}^2$ per indicare l'insieme delle coppie ordinate di numeri reali, quando non saremo interessati alla struttura moltiplicativa.

Def. Per ogni $z=x+yi\in\mathbb{C}$, con $x,y\in\mathbb{R}$, definiamo

- la *parte reale* di z come $\mathrm{Re}(z)=x\in\mathbb{R}$
- la *parte immaginaria* di z come $\mathrm{Im}(z)=y\in\mathbb{R}$
- il *modulo* di z come $|z|=\sqrt{x^2+y^2}=\sqrt{\mathrm{Re}(z)^2+\mathrm{Im}(z)^2}\geqslant 0$

- il *coniugato* di z come $\bar{z} = x - yi = \mathrm{Re}(z) - \mathrm{Im}(z)\, i \in \mathbb{C}$.

Oss. $z \in \mathbb{C} \Rightarrow z = \mathrm{Re}(z) + \mathrm{Im}(z)\, i$.
$z\bar{z} = x^2 + y^2 = |z|^2 \geqslant 0$.
$|z| = \sqrt{z\bar{z}}$
$|z| = 0 \Leftrightarrow z = 0$.
$\bar{\bar{z}} = z$.
Per $z \neq 0$ si ha $z^{-1} = \dfrac{1}{z} = \dfrac{\bar{z}}{z\bar{z}} = \dfrac{\bar{z}}{|z|^2}$.
$z \in \mathbb{R} \Leftrightarrow z = \bar{z}$.

Con una verifica diretta si dimostra che $\forall\, z_1, z_2, z \in \mathbb{C}$

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$$
$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$$
$$(\bar{z})^{-1} = \overline{(z^{-1})}$$

Si ha pertanto

$$|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1 \bar{z}_1 z_2 \bar{z}_2 = |z_1|^2 |z_2|^2$$

da cui estraendo la radice quadrata

$$|z_1 z_2| = |z_1||z_2|.$$

Esempio. $1 + 2i + 5 - 3i = 6 - i$

$(1 + 2i)(5 - 3i) = 5 - 3i + 10i - 6i^2 = 11 + 7i$

$\overline{1 + 2i} = 1 - 2i, \quad |1 - 2i| = \sqrt{5}$

$\mathrm{Re}(1 - 2i) = 1$, $\mathrm{Im}(1 - 2i) = -2$
$(1 - 2i)^{-1} = \dfrac{1 + 2i}{|1 - 2i|^2} = \dfrac{1}{5} + \dfrac{2}{5} i.$
$|i| = 1$

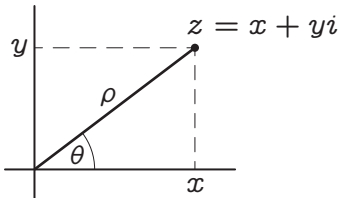
$\bar{i} = -i, \quad i^{-1} = -i,$

$i^3 = i^2 i = -i, \quad i^4 = (i^2)^2 = (-1)^2 = 1.$

Forma polare di un numero complesso

Dato $z = x + yi \in \mathbb{C}$, $z \neq 0$, con $x, y \in \mathbb{R}$, poniamo

$$\rho := |z|, \quad \arg z = \theta := \begin{cases} \arccos \dfrac{x}{|z|}, & \text{se } y \geq 0 \\ -\arccos \dfrac{x}{|z|}, & \text{se } y < 0 \end{cases}$$



Quindi ρ è il modulo di z cioè la distanza tra z e 0, mentre θ si chiama *argomento o anomalia* di z. L'argomento è definito soltanto a meno di multipli interi di 2π e in genere scegliamo il *valore principale*

$$\arg z \in]-\pi, \pi] \quad \text{oppure} \quad \arg z \in [0, 2\pi[.$$

Si ha

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases}$$

Quindi $z \in \mathbb{C}$ si può rappresentare in *forma polare* (o *trigonometrica*)

$$z = \rho(\cos \theta + i \sin \theta)$$

con $\rho = |z| \geqslant 0$ e $\theta = \arg z$ (estendiamo la forma polare anche a $z = 0$ lasciando l'argomento indefinito).

Se $\operatorname{Im} z = y = 0$ allora $z = x$ è un numero reale e si ha

$$|z| = |x|, \quad \arg z = \begin{cases} 0 & \text{se } x > 0 \\ \pi & \text{se } x < 0. \end{cases}$$

Se $\operatorname{Re} z = x = 0$ allora $z = yi$ è un numero immaginario e si ha

$$|z| = |y|, \quad \arg z = \begin{cases} \dfrac{\pi}{2} & \text{se } y > 0 \\ -\dfrac{\pi}{2} & \text{se } y < 0. \end{cases}$$

Se $\operatorname{Re} z = x \neq 0$ si ha anche

$$\arg z = \begin{cases} \arctan \dfrac{y}{x} & \text{se } x > 0 \\ \arctan \dfrac{y}{x} + \pi & \text{se } x < 0. \end{cases}$$

Esempio. $z = \dfrac{1}{\sqrt{2}} + \dfrac{1}{\sqrt{2}} i$, $|z| = 1$, $\arg z = \arccos \dfrac{1}{\sqrt{2}} = \dfrac{\pi}{4}$

$$\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} i = \cos \frac{\pi}{4} + i \sin \frac{\pi}{4}.$$

$w = 1 - \sqrt{3}\, i$, $|w| = 2$, $\arg z = -\arccos \dfrac{1}{2} = -\dfrac{\pi}{3}$

$$1 - \sqrt{3}\, i = 2\left(\cos\left(-\frac{\pi}{3}\right) + i \sin\left(-\frac{\pi}{3}\right)\right).$$

Uguaglianza. Dati due numeri complessi in forma polare

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$$
$$z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

$$z_1 = z_2 \;\Leftrightarrow\; \rho_1 = \rho_2 \text{ e } \exists\, k \in \mathbb{Z} \text{ t.c. } \theta_1 - \theta_2 = 2k\pi.$$

La forma polare non è utile per fare l'addizione ma risulta molto utile per fare la moltiplicazione e le potenze (e, come vedremo, anche le radici).

Prodotti. Dati due numeri complessi in forma polare

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$$
$$z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

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

In altre parole nel prodotto si moltiplicano i moduli e si sommano gli argomenti. Posto

$$z = \rho(\cos\theta + i\sin\theta) \neq 0,$$

si ha anche

$$\begin{aligned} z^{-1} &= \rho^{-1}(\cos\theta - i\sin\theta) = \\ &= \rho^{-1}\big(\cos(-\theta) + i\sin(-\theta)\big) \end{aligned}$$

cioè l'inverso si calcola invertendo il modulo e cambiando segno all'argomento. Si ha quindi

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\big(\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)\big).$$

Esempio.

$$\begin{aligned} 3\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) 5\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right) &= 15\left(\cos\left(\frac{\pi}{3}+\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{3}+\frac{\pi}{6}\right)\right) \\ &= 15\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) \\ &= 15i. \end{aligned}$$

Potenze. Iterando il ragionamento si ottengono formule per le potenze con esponente intero, per $z = \rho(\cos\theta + i\sin\theta) \neq 0$ e $n \in \mathbb{Z}$:

$$z^n = \rho^n\big(\cos(n\theta) + i\sin(n\theta)\big).$$

Esempio.

$$(1 - \sqrt{3}\,i)^3 = 2^3\left(\cos\left(-\frac{3\pi}{3}\right) + i\sin\left(-\frac{3\pi}{3}\right)\right) = -8.$$

Radici di un numero complesso

Caso reale.

Def. Dato $n \in \mathbb{N}$, $n \geqslant 2$, la *radice n-esima* di numero reale $\rho \geqslant 0$ è l'*unico* numero reale $u \geqslant 0$ t.c. $u^n = \rho$. Si scrive $u = \sqrt[n]{\rho}$, o anche $\sqrt{\rho}$ per $n = 2$.

Quindi, ad es., si ha $\sqrt{4} = 2$, ma non è corretto scrivere $\sqrt{4} = -2$ e nemmeno $\sqrt{4} = \pm 2$.

Caso complesso. Dato $z = \rho(\cos\theta + i\sin\theta) \in \mathbb{C}$ e dato $n \in \mathbb{N}$, $n \geqslant 2$, vogliamo trovare tutte le *radici n-esime complesse* di z, cioè i numeri complessi $w \in \mathbb{C}$ t.c.

$$w^n = z.$$

Se $z = 0$ allora evidentemente $w = 0$ è l'unica radice n-esima di 0. Se $z \neq 0$, determiniamo tutte le radici in forma polare

$$\begin{gathered} w = r(\cos\psi + i\sin\psi) \\ w^n = r^n\big(\cos(n\psi) + i\sin(n\psi)\big) \;\Leftrightarrow \\ r^n = \rho, \quad n\psi = \theta + 2k\pi \quad (k \in \mathbb{Z}) \;\Leftrightarrow \\ r = \sqrt[n]{\rho}, \quad \psi_k = \frac{\theta + 2k\pi}{n} \quad (k \in \mathbb{Z}). \end{gathered}$$

Quindi il modulo di w è la radice n-esima del modulo di z mentre $\arg w$ assume più valori $\psi_0, \psi_1 \ldots, \psi_{n-1}$ in corrispondenza di $k = 0, 1, \ldots, n-1$ (risp.). Tuttavia $\psi_n = \psi_0 + 2\pi$, e così di seguito i valori successivi si ripetono a meno di multipli interi di 2π. Si ha quindi che le radici n-esime di $z \neq 0$ sono esattamente n:

$$\begin{aligned} w_0 &= \sqrt[n]{\rho}\left(\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right) \\ w_1 &= \sqrt[n]{\rho}\left(\cos\frac{\theta+2\pi}{n} + i\sin\frac{\theta+2\pi}{n}\right) \\ &\vdots \\ w_k &= \sqrt[n]{\rho}\left(\cos\frac{\theta+2k\pi}{n} + i\sin\frac{\theta+2k\pi}{n}\right) \\ &\vdots \\ w_{n-1} &= \sqrt[n]{\rho}\left(\cos\frac{\theta+2(n-1)\pi}{n} + i\sin\frac{\theta+2(n-1)\pi}{n}\right). \end{aligned}$$

Hanno lo stesso modulo $\sqrt[n]{\rho}$ e due argomenti consecutivi differiscono di

$$\psi_{k+1} - \psi_k = \frac{2\pi}{n}$$

valore indipendente da k. Pertanto le radici n-esime di $z \neq 0$ sono i vertici di un n-gono regolare di raggio $\sqrt[n]{\rho}$ e si trovano sulla circonferenza di centro l'origine e raggio $\sqrt[n]{\rho}$.

Radici dell'unità. Definiamo

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

come l'insieme dei punti $z = x + yi \in \mathbb{C} = \mathbb{R}^2$ t.c. $|z|^2 = x^2 + y^2 = 1$. In altre parole S^1 è la circonferenza unitaria del piano con centro l'origine.

Oss. $z \in S^1 \Rightarrow z^{-1} = \bar{z} \in S^1$. Infatti $|z|^2 = z\bar{z} = 1$.
Ogni punto $z \in S^1$ si può scrivere come

$$z = \cos\theta + i\sin\theta, \quad \theta \in [0, 2\pi].$$

Le radici n-esime di 1 sono date dalle formule precedenti per $\rho = 1$ e $\theta = 0$:

$$\begin{aligned} w_0 &= 1 \\ w_1 &= \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n} \\ &\vdots \\ w_k &= \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} \\ &\vdots \\ w_{n-1} &= \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n}. \end{aligned}$$

Quindi $w_k \in S^1$ per $k = 0, 1, \ldots, n-1$. Eseguiamo alcuni calcoli.

$\boxed{n=2}$ $w_0 = 1$, $w_1 = -1$.

$\boxed{n=3}$ $w_0 = 1$, $w_1 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$,
$w_2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i = \bar{w}_1 = w_1^{-1}$.
Vertici di un triangolo equilatero.

$\boxed{n=4}$ $w_0 = 1$, $w_1 = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$, $w_2 = \cos\pi + i\sin\pi = -1$,
$w_3 = \cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2} = -i = \bar{w}_1 = w_1^{-1}$.
Vertici di un quadrato.

Oss. 1 è radice n-esima di 1.
w_k radice n-esima di $1 \Rightarrow w_k^{-1}$ radice n-esima di 1. Infatti

$$\left(w_k^{-1}\right)^n = (w_k^n)^{-1} = 1.$$

w_j, w_k radici n-esime di $1 \Rightarrow w_j w_k$ radice n-esima di 1. Infatti

$$(w_j w_k)^n = w_j^n w_k^n = 1.$$

Quindi l'insieme delle radici n-esime di 1 ha n elementi ed è chiuso rispetto a moltiplicazione e inversione. Si dice che forma un *gruppo*.

Oss. Per le radici n-esime di 1 si ha

$$w_k = w_1^k.$$

Cioè il gruppo delle radici n-esime di 1 è *ciclico*, ovvero è generato da un elemento. Le radici n-esime di 1 che generano tutte le radici n-esime mediante potenze (come w_1) sono dette *radici n-esime primitive* di 1. Si ha che w_k è una radice n-esima primitiva di 1 $\Leftrightarrow$ MCD$(k, n) = 1$. In particolare se n è *primo* tutte le radici n-esime di 1 tranne 1 sono primitive.

Spazi vettoriali numerici complessi

Per ogni $n \in \mathbb{N}$, in analogia con $\mathbb{R}^n$, possiamo definire

$$\mathbb{C}^n \stackrel{\text{def}}{=} \underbrace{\mathbb{C} \times \cdots \times \mathbb{C}}_{n \text{ volte}} = \{(z_1, \ldots, z_n) \mid z_1, \ldots, z_n \in \mathbb{C}\}.$$

Per definizione poniamo anche $\mathbb{C}^0 = \mathbb{R}^0 \stackrel{\text{def}}{=} \{0\}$. Si ha $\mathbb{R}^n \subset \mathbb{C}^n$. Gli elementi di $\mathbb{C}^n$ sono chiamati vettori (complessi), e li scriveremo come vettori riga o vettori colonna a seconda della convenienza del momento.

Su $\mathbb{C}^n$ definiamo l'addizione componente per componente e la moltiplicazione per scalari complessi, analogamente al caso reale:

$$(w_1, \ldots, w_n) + (z_1, \ldots, z_n) \stackrel{\text{def}}{=} (w_1 + z_1, \ldots, w_n + z_n)$$

$$\alpha(z_1, \ldots, z_n) \stackrel{\text{def}}{=} (\alpha z_1, \ldots, \alpha z_n)$$

$\forall\, (w_1, \ldots, w_n), (z_1, \ldots, z_n) \in \mathbb{C}^n$, $\forall\, \alpha \in \mathbb{C}$.

Def. $\mathbb{C}^n$ con queste operazioni di addizione e moltiplicazione scalare è detto *spazio vettoriale numerico complesso di dimensione* n.

L'addizione è associativa e commutativa dato che queste proprietà valgono su $\mathbb{C}$ e inoltre il *vettore nullo*

$$0_{\mathbb{C}^n} \stackrel{\text{def}}{=} (\underbrace{0, \ldots, 0}_{n \text{ volte}})$$

è elemento neutro per l'addizione (spesso lo indicheremo con 0).

$\forall\, (z_1, \ldots, z_n) \in \mathbb{C}^n$ definiamo il *vettore opposto* come

$$-(z_1, \ldots, z_n) \stackrel{\text{def}}{=} (-z_1, \ldots, -z_n)$$

e si ha

$$(z_1, \ldots, z_n) - (z_1, \ldots, z_n) = 0_{\mathbb{C}^n}.$$

Polinomi

Un *polinomio in una indeterminata* è un'*espressione formale* del tipo

$$f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

dove $n \in \mathbb{N}$ e $a_0, a_1, \ldots, a_n$ sono numeri reali (*polinomio reale*) oppure complessi (*polinomio complesso*). La lettera x è detta *indeterminata* o *variabile*. I numeri $a_0, \ldots, a_n$ sono detti *coefficienti* e a_0 è detto *termine noto*. Se x non compare allora f è detto *polinomio costante* e in particolare si ha il *polinomio nullo* 0. Il numero n è detto *grado* di f se $a_n \neq 0$, indicato $n = \deg(f)$. Per definizione poniamo $\deg(0) = -\infty$. Quindi $\deg(a) = 0$ se a è un polinomio costante non nullo (i polinomi costanti sono numeri).

Due polinomi

$$f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$
$$g = b_0 + b_1x + b_2x^2 + \cdots + b_mx^m$$

sono *uguali* se $n = m$ e $a_0 = b_0, a_1 = b_1, \ldots, a_n = b_n$ (cioè se hanno uguale il grado e i coefficienti corrispondenti).

L'insieme dei polinomi reali o complessi nell'indeterminata x si indica rispettivamente con $\mathbb{R}[x]$ oppure $\mathbb{C}[x]$ e si ha

$$\mathbb{R} \subset \mathbb{R}[x],\ \mathbb{C} \subset \mathbb{C}[x] \text{ e } \mathbb{R}[x] \subset \mathbb{C}[x].$$

I polinomi si possono sommare, moltiplicare per scalari e moltiplicare tra loro (diamo per scontate queste operazioni)

$$\deg(fg) = \deg(f) + \deg(g)$$
$$\deg(f + g) \leqslant \max(\deg(f), \deg(g)).$$

Poniamo $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$ a seconda del campo che vogliamo considerare.

Funzioni polinomiali. Dato $f \in \mathbb{K}[x]$ possiamo associare la corrispondente *funzione polinomiale*

$$f : \mathbb{K} \to \mathbb{K}$$

che ad ogni $u \in \mathbb{K}$ associa $f(u)$, ottenuto mettendo u al posto di x.

Divisione. Si può sempre fare la *divisione* tra polinomi: dati $f, g \in \mathbb{K}[x]$, con $g \neq 0$, esistono unici $q, r \in \mathbb{K}[x]$ t.c.

$$f = gq + r \quad \text{e} \quad \deg r < \deg g.$$

I polinomi q e r sono detti risp. *quoziente* e *resto* della divisione di f (*dividendo*) per g (*divisore*). Se $r = 0$ diciamo che f è *divisibile* per g e si ha $f = gq$. L'algoritmo della divisione viene dato per scontato. Si raccomanda di ripassare questo algoritmo che è stato trattato alle superiori, inclusa la Regola di Ruffini.

Zeri di polinomi

Zeri. Un numero $\alpha \in \mathbb{K}$ è detto *zero* o *radice* di $f \in \mathbb{K}[x]$ se $f(\alpha) = 0$.

Di un polinomio reale possiamo considerare sia gli zeri reali che quelli complessi.

Oss. Se $\alpha \in \mathbb{C}$ è zero di $f \in \mathbb{R}[x]$ allora anche $\bar{\alpha}$ è zero di f. Infatti se $f = a_0 + a_1 x + \cdots + a_n x^n$ con $a_0, \ldots, a_n \in \mathbb{R}$ si ha $\bar{a}_i = a_i \ \forall\, i = 0, \ldots, n$ da cui

$$0 = f(\alpha) = \overline{f(\alpha)} = f(\bar{\alpha}) \quad \Rightarrow \quad \bar{\alpha} \text{ zero di } f.$$

Oss. Non è detto che un polinomio reale abbia zeri reali, ad es. $x^2 + 1$.

Teor. *$\alpha \in \mathbb{K}$ è zero di $f \in \mathbb{K}[x]$ $\Leftrightarrow$ f è divisibile per $x - \alpha$.*

Dim. $\boxed{\Rightarrow}$ Divisione di f per $x - \alpha \rightsquigarrow q, r \in \mathbb{K}[x]$ t.c. $f = (x - \alpha)q + r$ e $\deg r < \deg(x - a) = 1 \Rightarrow r =$ costante $\Rightarrow 0 = f(\alpha) = r$.

$\boxed{\Leftarrow}$ $f = (x - \alpha)q$ per un certo $q \in \mathbb{K}[x] \Rightarrow f(\alpha) = 0$. □

Gli zeri si determinano nel modo usuale risolvendo un'equazione a coefficienti reali o complessi. Consideriamo $a, b, c \in \mathbb{C}$, $a \neq 0$.

Grado 1. $\quad ax + b = 0 \quad \rightsquigarrow \quad x = -\dfrac{b}{a}$

Grado 2. $\quad ax^2 + bx + c = 0 \quad \rightsquigarrow \quad x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Oss. Se l'equazione è nella forma $ax^2 + 2bx + c = 0$ possiamo usare la formula ridotta $x = \dfrac{-b \pm \sqrt{b^2 - ac}}{a}$

Esempio. $ix + 4 - i = 0 \Rightarrow x = 1 + 4i$.
$x^2 + 1 = 0 \Rightarrow x = \pm i$.
$x^2 + 2ix - 1 = 0 \Rightarrow x = -i \pm \sqrt{i^2 + 1} = -i$.

Molteplicità. Se $\alpha \in \mathbb{K}$ è zero di $f \in \mathbb{K}[x]$, con $\deg(f) \geqslant 1$, abbiamo visto che f è divisibile per $x - \alpha$ ovvero $\exists\, q_1 \in \mathbb{K}[x]$ t.c. $f = (x - \alpha)q_1$. Può succedere che α sia zero anche di q_1, cioè $q_1(\alpha) = 0$ e quindi q_1 è divisibile per $x - \alpha$. Si ha $f = (x - \alpha)^2 q_2$. Si può continuare in questo modo al massimo un certo numero m di volte, ottenendo alla fine $f = (x - \alpha)^m q$ per un certo $q \in \mathbb{K}[x]$ t.c. $q(\alpha) \neq 0$. Si noti che $\deg(f) = m + \deg(q)$ quindi $m \leqslant \deg(f)$.

Def. Supponiamo che $\alpha \in \mathbb{K}$ sia zero di $f \in \mathbb{K}[x]$ con $\deg(f) \geqslant 1$. Si chiama *molteplicità* di α il massimo intero $m \in \mathbb{N}$ t.c. f sia divisibile per $(x - \alpha)^m$.

Oss. Per la molteplicità m di uno zero di f si ha $1 \leqslant m \leqslant \deg(f)$.

Polinomi interi. Per un polinomio a coefficienti *interi*

$$f = a_0 + a_1 x + \cdots + a_n x^n,$$

$a_0, \ldots, a_n \in \mathbb{Z}$, $a_n \neq 0$ gli zeri *razionali* se esistono vanno cercati tra le frazioni del tipo $\frac{p}{q}$ dove il numeratore p è divisore (positivo o negativo) del termine noto a_0, e il denominatore q è divisore positivo di a_n. Questo in certi casi consente di trovare zeri di polinomi di grado $\geqslant 3$.

Esempio. $x^3 - x^2 - 7x - 2 = 0$ cerchiamo zeri interi tra i divisori di 2 cioè ± 1, ± 2 e con qualche tentativo si trova lo zero $x = -2$. Successivamente si divide il polinomio dato per $x + 2$ e se ne ottiene uno di secondo grado di cui sappiamo calcolare gli zeri. Nel caso specifico si ha

$$\frac{x^3 - x^2 - 7x - 2}{x + 2} = x^2 - 3x - 1.$$

Esempio. $2x^3 - x^2 + x + 1 = 0$ cerchiamo zeri interi tra i divisori di 1, cioè ± 1, e poi tra le frazioni $\pm\frac{1}{2}$ e con qualche tentativo si trova lo zero $x = -\frac{1}{2}$. Si esegue quindi la divisione

$$\frac{2x^3 - x^2 + x + 1}{x + \frac{1}{2}} = 2(x^2 - x + 1).$$

Questo metodo non funziona se il polinomio non ha coefficienti interi o non ha radici razionali. Esistono formule per risolvere equazioni di terzo e quarto grado, ma non trattiamo questo argomento.

Prop. *Un polinomio reale o complesso di grado $n \geqslant 1$ ha al massimo n zeri distinti.*

Dim. Per induzione su n.

Base dell'induzione $n = 1$. Il polinomio $ax + b$, con $a, b \in \mathbb{K}$ e $a \neq 0$ ha soltanto lo zero $x = -\dfrac{b}{a}$ e quindi l'enunciato è vero.

Ipotesi induttiva Supponiamo vero l'enunciato per tutti i polinomi di grado $n - 1 \geqslant 1$ e dimostriamolo per i polinomi di grado n.

Sia $f \in \mathbb{K}[x]$ con $\deg(f) = n$. Se f non ha zeri allora non c'è nulla da dimostrare. Se $\alpha \in \mathbb{K}$ è zero di f allora $\exists\, q \in \mathbb{K}[x]$ t.c. $f = (x - \alpha)q \Rightarrow \deg(q) = n - 1$. Per l'ipotesi induttiva q ha al massimo $n - 1$ zeri, che sono anche zeri di f. Pertanto f ha al massimo n zeri (incluso α). □

Oss. Se $f \in \mathbb{K}[x]$ è un polinomio di grado $\deg(f) \leqslant n$ con più di n zeri allora $f = 0$ (polinomio nullo).

Principio d'Identità dei Polinomi. *Due polinomi reali o complessi sono uguali se e solo se sono uguali le corrispondenti funzioni polinomiali.*

In altre parole $f = g \Leftrightarrow f(u) = g(u)\ \forall\, u \in \mathbb{R}$ (o $\mathbb{C}$).

Dim. Siano $f, g \in \mathbb{K}[x]$. Se $f = g$ allora è ovvio che $f(u) = g(u)\ \forall\, u \in \mathbb{K}$.

Viceversa, se $f(u) = g(u)\ \forall\, u \in \mathbb{K}$ allora il polinomio $h = f - g \in \mathbb{K}[x]$ soddisfa $h(u) = f(u) - g(u) = 0\ \forall\, u \in \mathbb{K}$ e quindi h ha infiniti zeri $\Rightarrow$ $h = 0 \Rightarrow f = g$. □

N.B. Non è un enunciato banale perché l'uguaglianza di polinomi è espressa in termini dei coefficienti e del grado, concetto diverso dall'uguaglianza di funzioni.

Enunciamo senza dimostrare il seguente importante teorema.

Teorema Fondamentale dell'Algebra. *Ogni polinomio complesso di grado* $\geqslant 1$ *ha almeno uno zero complesso.*

Dato $f \in \mathbb{C}[x]$ con $\deg(f) = n \geqslant 1$, il Teorema fondamentale dell'Algebra garantisce l'esistenza di almeno uno zero $\alpha_1 \in \mathbb{C}$ e quindi $f = (x-\alpha_1)q$ con $\deg(q) = n - 1$. Se $n - 1 \geqslant 1$ allora q ha uno zero α_2 che è anche zero di f. Procedendo in questo modo si arriva ad un quoziente di grado 0, cioè costante. Alla fine f risulta scomposto in fattori di primo grado. Raccogliendo i fattori simili si ottiene la fattorizzazione

$$f = a(x - \alpha_1)^{m_1} \cdots (x - \alpha_k)^{m_k}$$

dove $\alpha_1, \ldots, \alpha_k$ sono gli *zeri distinti* di f, $a = a_n \neq 0$ è il coefficinete di x^n, e $m_1, \ldots, m_k \in \mathbb{N}$ sono le molteplicità degli zeri corrispondenti.

Oss. Questa fattorizzazione è unica a meno di riordinare i fattori.

Oss. $m_1 + \cdots + m_k = \deg(f) = n$.

Oss. Una fattorizzazione simile si ottiene per polinomi reali che abbiano tutti gli zeri reali.

Teor. *Sia* $f \in \mathbb{R}[x]$ *un polinomio reale di grado dispari. Allora* f *ha almeno uno zero reale.*

Dim. Supponiamo per assurdo che f non abbia zeri reali. Allora gli zeri di f sono tutti numeri complessi non reali $\alpha_1 \ldots, \alpha_k$. Dato che f è reale per ogni zero α_j c'è anche il coniugato $\bar{\alpha}_j \neq \alpha_j \Rightarrow k$ pari $\Rightarrow f$ è prodotto di un numero pari di fattori di primo grado $\Rightarrow \deg(f)$ pari, contraddizione. □

Matrici

Def. Dati due numeri positivi $m, n \in \mathbb{N}$, una *matrice* di tipo $m \times n$ è una tabella di numeri reali o complessi con m righe e n colonne. Se $m = n$ la matrice è detta *quadrata*. I numeri che formano la matrice sono detti *entrate* o *elementi*.

Una matrice quadrata $n \times n$ è detta anche matrice quadrata di ordine n.

Esempio.

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

è una matrice 2×3.

L'insieme delle matrici $m \times n$ a entrate reali si indica con $M_{m,n}(\mathbb{R})$ o anche $\mathbb{R}^{m \times n}$. Analogamente l'insieme delle matrici $m \times n$ a entrate complesse si indica con $M_{m,n}(\mathbb{C})$ o anche $\mathbb{C}^{m \times n}$. L'insieme delle matrici quadrate di ordine n si indica con $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ e similmente per quelle complesse $M_n(\mathbb{C}) = M_{n,n}(\mathbb{C})$. Si ha $M_{m,n}(\mathbb{R}) \subset M_{m,n}(\mathbb{C})$.

Esempio.

$$B = \begin{pmatrix} 2 & i \\ -1 + 3i & -1 + 7i \end{pmatrix} \in M_2(\mathbb{C}).$$

Le entrate di una matrice $m \times n$ si numerano a partire da quella più in alto a sinistra, mediante una coppia di indici (i, j) con $i \in \{1, \ldots, m\}$ (indice di riga) e $j \in \{1, \ldots, n\}$ (indice di colonna). Quindi i numera le righe dall'alto verso il basso, e j numera le colonne da sinistra verso destra. Indichiamo l'entrata di posto (i, j) di una matrice A con A_{ij}.

Esempio. Per la matrice A del primo esempio si ha $A_{11} = 3$, $A_{13} = -5$, $A_{21} = 0$, $A_{23} = 4$.

Una matrice A di tipo $m \times n$ la cui generica entrata sia un numero a_{ij} si indica con $A = (a_{ij})$, ossia $A_{ij} = a_{ij}$.

Addizione. Nel seguito indicheremo con $\mathbb{K}$ uno dei campi $\mathbb{R}$ o $\mathbb{C}$.
Date $A, B \in M_{m,n}(\mathbb{K})$ definiamo la *somma* $A + B \in M_{m,n}(\mathbb{K})$ ponendo

$$(A + B)_{ij} \stackrel{\text{def}}{=} A_{ij} + B_{ij}$$

ossia si sommano le entrate corrispondenti di A e B.

Moltiplicazione scalare. Dati $\alpha \in \mathbb{K}$ e $A \in M_{m,n}(\mathbb{K})$ definiamo la *moltiplicazione scalare* $\alpha A \in M_{m,n}(\mathbb{K})$ ponendo

$$(\alpha A)_{ij} \stackrel{\text{def}}{=} \alpha A_{ij}$$

ossia si moltiplicano le entrate di A per il numero α.

Prodotto righe per colonne

Abbiamo visto la definizione di matrice e le operazioni di somma e moltiplicazione scalare. Vedremo oggi una ulteriore operazione, prima in un caso speciale e poi in generale. Il campo sarà $\mathbb{R}$ o $\mathbb{C}$, che come sempre indicheremo con $\mathbb{K}$ (le cose che vedremo valgono su qualunque campo). L'insieme delle matrici $m \times n$, ovvero con m righe e n colonne, a entrate nel campo $\mathbb{K}$, è stato indicato con $M_{m,n}(\mathbb{K})$.

Def. Dati un vettore riga e un vettore colonna con n componenti

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \in M_{1,n}(\mathbb{K}) \quad \text{e} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in M_{n,1}(\mathbb{K})$$

chiamiamo *prodotto di A per B* lo scalare

$$AB \stackrel{\text{def}}{=} \sum_{k=1}^{n} a_k b_k = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \in \mathbb{K}.$$

Esempio.

$$\begin{pmatrix} 3 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = -5 \qquad \begin{pmatrix} 1 & 0 & 1+i \end{pmatrix} \begin{pmatrix} -2i \\ 3 \\ 1 \end{pmatrix} = 1 - i.$$

Data $A \in M_{m,n}(\mathbb{K})$ indichiamo con $A^{(i)} \in M_{1,n}(\mathbb{K})$ la i-esima riga di A, con $A_{(j)} \in M_{m,1}(\mathbb{K})$ la j-esima colonna di A e con $A_{ij} \in \mathbb{K}$ l'entrata di posto (i, j) di A, per $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Esempio.

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 5 & -3 \\ 0 & 7 & 0 \end{pmatrix} \rightsquigarrow A^{(1)} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}, \; A_{(2)} = \begin{pmatrix} 0 \\ 5 \\ 7 \end{pmatrix}, \; A_{32} = 7.$$

Generalizziamo l'operazione di moltiplicazione vista sopra.

Def. Consideriamo due matrici $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,\ell}(\mathbb{K})$. Il *prodotto righe per colonne di A per B*, è la matrice $AB \in M_{m,\ell}(\mathbb{K})$ avente come entrata di posto (i, j)

$$(AB)_{ij} \stackrel{\text{def}}{=} A^{(i)} B_{(j)} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

$\forall\, i = 1, \ldots, m$, $\forall\, j = 1, \ldots, \ell$.

Oss. Il prodotto righe per colonne è definito solo se il numero di colonne della prima matrice è uguale al numero di righe della seconda.

Oss. Se A è di tipo $m \times n$ e B è di tipo $n \times \ell$ allora AB è di tipo $m \times \ell$.

Oss. Due matrici quadrate $n \times n$ sono moltiplicabili e il risultato è una matrice quadrata $n \times n$.

Esempio.

$$\begin{pmatrix} 1 & 0 & -2 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 4 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -2 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -5 \\ 5 \end{pmatrix}$$

$$(2 \;\; 1) \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} = (7 \;\; 4)$$

Oss. Una matrice $m \times n$ è moltiplicabile per un vettore colonna a n componenti. Il risultato è un vettore colonna a m componenti.

Oss. Un vettore riga a m componenti è moltiplicabile per una matrice $m \times n$. Il risultato è un vettore riga a n componenti.

Oss. Il prodotto righe per colonne in generale non è commutativo, nemmeno per matrici quadrate $n \times n$, con $n \geqslant 2$.

Esempio.

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

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

$$AB \neq BA.$$

Proprietà delle matrici

Abbiamo visto le seguenti operazioni sulle matrici: somma, moltiplicazione scalare, prodotto righe per colonne.

La somma di matrici è associativa, commutativa e ha l'elemento neutro, la *matrice nulla* $0 \in M_{m,n}(\mathbb{K})$ le cui entrate sono tutte nulle. Ogni matrice $A \in M_{m,n}(\mathbb{K})$ ha la *matrice opposta* $-A$ le cui entrate sono gli opposti delle entrate di A, e soddisfa $A - A = 0$. Queste proprietà sono di verifica molto semplice e simile al caso già visto per i vettori di $\mathbb{R}^n$.

Il prodotto righe per colonne abbiamo visto che non è commutativo (non solo, se AB è definito non è detto che lo sia anche BA).

Prop. *Il prodotto righe per colonne, quando è definito, è associativo e distributivo rispetto alla somma, ovvero per tutte le matrici A, B, C per cui le operazioni indicate abbiano senso si ha*

$$A(BC) = (AB)C$$
$$A(B + C) = AB + AC$$
$$(A + B)C = AC + BC.$$

Dim. Dimostriamo la proprietà associativa. Consideriamo tre matrici

$$A \in M_{m,n}(\mathbb{K}),\ B \in M_{n,\ell}(\mathbb{K}),\ C \in M_{\ell,r}(\mathbb{K}).$$

Possiamo fare i prodotti AB, BC, $A(BC)$, $(AB)C$. Vogliamo dimostrare

$$A(BC) = (AB)C.$$

Due matrici sono uguali $\Leftrightarrow$ sono dello stesso tipo e hanno uguali le entrate corrispondenti. $A(BC)$ è di tipo $m \times r$, come $(AB)C$. Mostriamo che queste due matrici hanno uguali le entrate corrispondenti.

$$\big(A(BC)\big)_{ij} = \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k=1}^{n} A_{ik} \left(\sum_{h=1}^{\ell} B_{kh}C_{hj} \right) = \sum_{k=1}^{n} \sum_{h=1}^{\ell} A_{ik}B_{kh}C_{hj}$$

$$\big((AB)C\big)_{ij} = \sum_{h=1}^{\ell} (AB)_{ih}C_{hj} = \sum_{h=1}^{\ell} \left(\sum_{k=1}^{n} A_{ik}B_{kh} \right) C_{hj} = \sum_{k=1}^{n} \sum_{h=1}^{\ell} A_{ik}B_{kh}C_{hj}$$

$$\Rightarrow\ \big(A(BC)\big)_{ij} = \big((AB)C\big)_{ij} \quad \forall\, i = 1, \dots m,\ \forall\, j = 1, \dots, r.$$

La proprietà distributiva si dimostra con un calcolo analogo, lasciato come esercizio volontario (non è necessario conoscerlo). □

Matrice trasposta.

Def. Data $A \in M_{m,n}(\mathbb{K})$ chiamiamo *trasposta* di A la matrice ${}^tA \in M_{n,m}(\mathbb{K})$ ottenuta da A scambiando le righe con le colonne, ovvero t.c.

$$({}^tA)_{ij} := A_{ji} \qquad \forall\, i = 1, \dots, n,\ \forall\, j = 1, \dots, m.$$

Oss. $A \in M_n(\mathbb{K}) \Rightarrow {}^tA \in M_n(\mathbb{K})$.
${}^t({}^tA) = A$.

Prop. ${}^t(A + B) = {}^tA + {}^tB$, $\forall\, A, B \in M_{m,n}(\mathbb{K})$.

La dimostrazione segue subito dalle definizioni di somma e di trasposta.

Prop. ${}^t(AB) = {}^tB\,{}^tA$, $\forall\, A \in M_{m,n}(\mathbb{K})$, $\forall\, B \in M_{n,\ell}(\mathbb{K})$.

Dim. ${}^t(AB)$ e ${}^tB\,{}^tA$ sono di tipo $\ell \times m$.

$$\left({}^t(AB)\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki}$$

$$\left({}^tB\,{}^tA\right)_{ij} = \sum_{k=1}^{n} ({}^tB)_{ik}({}^tA)_{kj} = \sum_{k=1}^{n} B_{ki}A_{jk}$$

$$\Rightarrow\ \left({}^t(AB)\right)_{ij} = \left({}^tB\,{}^tA\right)_{ij} \qquad \forall\, i = 1, \ldots, \ell,\ \forall\, j = 1, \ldots, m.$$

□

Matrici diagonali

Def. Data una matrice quadrata $A \in M_n(\mathbb{K})$ chiamiamo *diagonale principale* il vettore

$$(A_{11},\ A_{22}, \ldots, A_{nn}) \in \mathbb{K}^n.$$

Def. Una matrice quadrata $D \in M_n(\mathbb{K})$ è *diagonale* se $D_{ij} = 0\ \forall\, i \neq j$.

Oss. D è una matrice diagonale se le entrate fuori dalla diagonale principale sono nulle.

Def. Dati $n \in \mathbb{N}$ e $a_1, \ldots, a_n \in \mathbb{K}$, indichiamo con

$$\mathrm{diag}(a_1, \ldots, a_n) \in M_n(\mathbb{K})$$

la matrice diagonale avente $(a_1, \ldots, a_n)$ come diagonale principale, ossia

$$\mathrm{diag}(a_1, \ldots, a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

Matrice identica. Per ogni intero $n \geqslant 1$ consideriamo la *matrice identica* (o *matrice identità*) di ordine n

$$I_n := \mathrm{diag}(1, 1, \ldots, 1) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in M_n(\mathbb{K})$$

avente 1 sulla diagonale principale e 0 altrove.

Def. Per ogni $i, j \in \mathbb{N}$ poniamo

$$\delta_{ij} := \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \qquad (\textit{Delta di Kronecker}).$$

Esempio. $\delta_{11} = 1$, $\delta_{12} = 0$, $\delta_{33} = 1$, $\delta_{41} = 0$.

Oss. δ_{ij} sono le entrate della matrice identica, ovvero $(I_n)_{ij} = \delta_{ij}$.

Prop. *Sia $A \in M_{m,n}(\mathbb{K})$. Allora*

$$A\,I_n = I_m\,A = A.$$

In altre parole la matrice identica è elemento neutro per il prodotto righe per colonne.

Dim. Dimostriamo $A\,I_n = A$.

$$(A\,I_n)_{ij} = \sum_{k=1}^{n} A_{ik}\delta_{kj} = A_{ij}\delta_{jj} = A_{ij}$$

dato che per $k \neq j$, $\delta_{kj} = 0$, mentre $\delta_{jj} = 1$.

La dimostrazione di $I_m\,A = A$ è simile. □

Matrici invertibili

Def. Sia $A \in M_n(\mathbb{K})$. Diciamo che A è *invertibile* se esiste una matrice $B \in M_n(\mathbb{K})$ t.c.

$$AB = BA = I_n.$$

In questo caso scriviamo $A^{-1} := B$ e chiamiamo A^{-1} l'*inversa* di A.

Se A è invertibile si ha quindi $AA^{-1} = A^{-1}A = I_n$.

Oss. L'inversa se esiste è unica. Infatti se $B, B' \in M_n(\mathbb{K})$ sono inverse di A, si ha

$$B = B\,I_n = B(AB') = (BA)B' = I_n\,B' = B'.$$

La matrice nulla 0 chiaramente non è invertibile dato che $0B = 0$.

Oss. $I_n^{-1} = I_n$.

Esempio.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

non è invertibile, infatti se

$$B = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

fosse l'inversa di A si avrebbe

$$AB = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

che è impossibile.

Pertanto esistono matrici quadrate non nulle non invertibili.

Esempio.

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

è invertibile e si può determinare l'inversa

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

risolvendo un sistema di equazioni

$$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{cases} x + 2z = 1 \\ y + 2t = 0 \\ x = 0 \\ y = 1 \end{cases} \qquad \begin{cases} z = \frac{1}{2} \\ t = -\frac{1}{2} \\ x = 0 \\ y = 1 \end{cases}$$

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

In seguito capiremo esattamente quali matrici quadrate sono invertibili e vedremo metodi più efficienti per determinare l'inversa.

Sistemi lineari

Continuiamo ad indicare con $\mathbb{K}$ uno dei campi $\mathbb{R}$ o $\mathbb{C}$.

Def. Un'espressione del tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

è detta *equazione lineare* nelle incognite $x_1, \ldots, x_n$, dove $a_1, \ldots, a_n \in \mathbb{K}$ sono i *coefficienti*, $b \in \mathbb{K}$ è il *termine noto* e $n \in \mathbb{N}$.

Un *sistema lineare* di m equazioni in n incognite $x_1, \ldots, x_n$ è un insieme di m equazioni lineari

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

Oss. I coefficienti a_{ij} hanno due indici: i per l'equazione e j per l'incognita. Un'equazione lineare è un caso particolare di sistema lineare con $m = 1$.

Forma matriciale. *Matrice dei coefficienti del sistema S*

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{K}).$$

Vettore dei termini noti

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m.$$

Vettore delle incognite

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

N.B. Le incognite sono simboli, non numeri.

Svolgendo il prodotto righe per colonne si ha:

$$AX = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

e quindi il sistema S può essere scritto nell'importante *forma matriciale*

$$S: AX = B.$$

Si osservi che $A \in M_{m,n}(\mathbb{K})$ fornisce i coefficienti del sistema, il numero m di equazioni (numero di righe) e il numero n di incognite (numero di colonne). Il vettore B dei termini noti deve avere m componenti.

Soluzioni. Si chiama *soluzione* del sistema S un vettore

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n$$

t.c.

$$AU = B,$$

ovvero le equazioni del sistema diventano identità sostituendo U al posto di X. Una soluzione ha tante componenti quante sono le incognite.

Un sistema lineare è *compatibile* se ammette almeno una soluzione, *incompatibile* o *impossibile* in caso contrario.

Esempio. $2x = 0$ è compatibile avendo la soluzione $x = 0$.

$\begin{cases} 2x = 3 \\ 2x = 4 \end{cases}$ è incompatibile (non ammette soluzioni).

Spazio delle soluzioni. Si chiama *spazio delle soluzioni* del sistema $S : AX = B$ l'insieme Σ_S di tutte le soluzioni di S, ovvero

$$\Sigma_S \stackrel{\text{def}}{=} \{U \in \mathbb{K}^n \mid AU = B\} \subset \mathbb{K}^n.$$

Si ha pertanto $\Sigma_S \neq \emptyset \Leftrightarrow S$ è compatibile.
Se il sistema S è incompatibile si ha $\Sigma_S = \emptyset$.

Sistemi omogenei. Un sistema lineare è *omogeneo* se i termini noti sono tutti nulli, ovvero è del tipo

$$AX = 0_{\mathbb{K}^m}.$$

Un sistema omogeneo è sempre compatibile avendo la *soluzione nulla*

$$X = 0_{\mathbb{K}^n}.$$

N.B. Potrebbero esserci altre soluzioni oltre a quella nulla.

Oss. $AX = B$ omogeneo $\Leftrightarrow B = 0_{\mathbb{K}^m} \Leftrightarrow$ ammette la soluzione nulla.

Pivot. Data una matrice A si chiamano *pivot* le prime entrate non nulle delle righe (ammesso che ve ne siano). Una riga ha un pivot $\Leftrightarrow$ è non nulla, e la prima entrata non nulla su tale riga è il suo pivot.
Pertanto A_{ip} è il pivot della riga i-esima $\Leftrightarrow A_{ip} \neq 0$ e $A_{ij} = 0 \ \forall j < p$.

Esempio.

$$\begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 5 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

il pivot della prima riga è 3, quello della seconda è 5 mentre la terza non ha pivot essendo nulla.

Matrici a gradini. Una matrice $A \in M_{m,n}(\mathbb{K})$ è *a gradini* o *a scala* se è nulla oppure se i suoi pivot sono $A_{1p_1}, A_{2p_2}, \ldots, A_{rp_r}$ con

$$p_1 < p_2 < \cdots < p_r,$$

mentre tutte le righe dopo la r-esima (se presenti) sono nulle. In altre parole A è a gradini se il pivot di ciascuna riga successiva alla prima è più a destra del pivot della riga precedente.

Oss. Il numero r di pivot di una matrice *a gradini* $m \times n$ non può eccedere il numero di righe e di colonne, ovvero $r \leqslant \min(m, n)$.

N.B. In inglese "matrice a gradini" si dice *"matrix in row echelon form"*.

Esempi. Le seguenti matrici sono a gradini.

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

Esempi. Le seguenti matrici non sono a gradini.

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

Sistemi a gradini. Un sistema lineare $AX = B$ è *a gradini* o *a scala* se la sua matrice A è a gradini.

Esempio. Il seguente sistema è a gradini

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 3x_2 - 4x_3 = 0 \\ 5x_3 = 1 \end{cases} \qquad A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -4 \\ 0 & 0 & 5 \end{pmatrix}.$$

Esempio. Il seguente sistema non è a gradini

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 3x_2 - 4x_3 = 0 \\ x_2 + 5x_3 = 1 \end{cases} \qquad A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -4 \\ 0 & 1 & 5 \end{pmatrix}.$$

Sistemi equivalenti. Due sistemi $S\colon AX = B$ e $S'\colon A'X = B'$ nelle stesse incognite $x_1, \ldots, x_n$ sono *equivalenti* se sono entrambi compatibili e hanno le stesse soluzioni, oppure se sono entrambi incompatibili: $S \sim S'$. In altre parole $S \sim S' \Leftrightarrow \Sigma_S = \Sigma_{S'}$.

Risoluzione. Risolvere un sistema lineare significa capire se è compatibile e in caso affermativo trovarne tutte le soluzioni. Un sistema si prende in considerazione per essere risolto. Come vedremo, un sistema lineare compatibile reale o complesso ammette una oppure infinite soluzioni. Un sistema incompatibile non ne ammette nessuna.

In primo luogo capiremo come si risolvono i sistemi a gradini. In un secondo momento vedremo come trasformare un sistema lineare qualunque in un sistema equivalente a gradini, che sapremo risolvere.

Risoluzione di un sistema a gradini

Consideriamo un sistema lineare a gradini di m equazioni in n incognite

$$S : AX = B$$

$$A = (a_{ij}) \in M_{m,n}(\mathbb{K}), \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Per ipotesi A è una matrice a gradini.

Sistema incompatibile

Se A ha una riga nulla, diciamo la i-esima, ma $b_i \neq 0$ allora S è incompatibile, dato che la i-esima equazione sarebbe $0 = b_i$.

Esempio.

$$\begin{cases} x_1 - 2x_2 = 1 \\ 0 = 3 \end{cases}$$

Sistema compatibile

Supponiamo ora che la situazione precedente non si presenti.

Soluzione unica. Tutte le incognite hanno un coefficiente pivot, quindi tutte le righe di A dopo la n-esima sono nulle e anche i corrispondenti termini noti sono nulli. Tali equazioni sono $0 = 0$ e possono essere eliminate.

L'ultima equazione sarà $a_{nn}x_n = b_n$ con $a_{nn} \neq 0$. Si trova subito

$$x_n = \frac{b_n}{a_{nn}}.$$

Sostituendo nella penultima equazione si trova x_{n-1} e così via procedendo a ritroso si trovano tutte le incognite e il sistema ha un'unica soluzione.

Esempio.

$$\begin{cases} x_1 - 2x_2 = 1 \\ x_2 = 3 \end{cases} \qquad \begin{cases} x_1 = 7 \\ x_2 = 3 \end{cases}$$

Infinite soluzioni. Se ci sono incognite $x_{i_1}, \ldots, x_{i_d}$ senza coefficiente pivot, a queste si assegnano *parametri liberi*

$$\begin{cases} x_{i_1} = t_1 \\ \cdots\cdots \\ x_{i_d} = t_d \end{cases}$$

Le altre incognite $x_{j_1}, \ldots, x_{j_r}$, aventi coefficiente pivot, si determinano a ritroso come nel caso precedente, partendo dall'ultima equazione.

Oss. $d = n - r$.

In definitiva tutte le incognite $x_1, \ldots, x_n$ risulteranno espresse come polinomi costanti o di primo grado aventi i parametri liberi come variabili. La soluzione così espressa è detta *soluzione generale*. Ogni scelta dei parametri liberi $t_1 \ldots, t_d \in \mathbb{K}$ determina una soluzione. Pertanto il sistema ammette infinite soluzioni.

Esempio.

$$\begin{cases} x_1 - 2x_2 + x_3 - 2x_4 = 1 \\ x_2 - 2x_3 = 3 \end{cases} \qquad \begin{cases} x_1 = 3t_1 + 2t_2 + 7 \\ x_2 = 2t_1 + 3 \\ x_3 = t_1 \\ x_4 = t_2 \end{cases}$$

Metodo di eliminazione di Gauss

Questo metodo permette di trasformare una matrice in una matrice a gradini, in modo tale che i sistemi lineari corrispondenti siano equivalenti. Come vedremo permette di fare anche varie altre cose tra cui calcolare l'inversa di una matrice.

Operazioni elementari sulle righe di una matrice. Data $A \in M_{m,n}(\mathbb{K})$ consideriamo i seguenti tre tipi di operazioni (con $i \neq j$).

Tipo I. Scambiare due righe di A.

$$\begin{pmatrix} \vdots \\ A^{(i)} \\ \vdots \\ A^{(j)} \\ \vdots \end{pmatrix} \overset{I}{\rightsquigarrow} \begin{pmatrix} \vdots \\ A^{(j)} \\ \vdots \\ A^{(i)} \\ \vdots \end{pmatrix}$$

Tipo II. Moltiplicare una riga di A per uno scalare $\alpha \neq 0$.

$$\begin{pmatrix} \vdots \\ A^{(i)} \\ \vdots \end{pmatrix} \overset{II}{\rightsquigarrow} \begin{pmatrix} \vdots \\ \alpha A^{(i)} \\ \vdots \end{pmatrix}$$

Tipo III. Sommare ad una riga di A un multiplo di un'altra riga di A.

$$\begin{pmatrix} \vdots \\ A^{(i)} \\ \vdots \\ A^{(j)} \\ \vdots \end{pmatrix} \overset{III}{\rightsquigarrow} \begin{pmatrix} \vdots \\ A^{(i)} \\ \vdots \\ A^{(j)} + \alpha A^{(i)} \\ \vdots \end{pmatrix}$$

Per ciascuna operazione le altre righe non vengono modificate.

Oss. Le operazioni inverse sono operazioni elementari dello stesso tipo. Questo è evidente per le operazioni di tipo I e II. Per le operazioni di tipo III si riottiene la matrice originale sommando $-\alpha A^{(i)}$ alla riga j.

Algoritmo di Gauss. Supponiamo che $A \in M_{m,n}(\mathbb{K})$ non sia a gradini.

Passo 1. Mediante un'operazione di tipo I (scambio di righe) mettiamo come prima riga una che abbia il pivot più a sinistra possibile (potrebbe già essere la prima riga). Supponiamo quindi che tale pivot sia $a_{1p} \neq 0$. Le colonne che precedono la p-esima, se presenti (cioè se $p > 1$), sono nulle.

Passo 2. Mediante operazioni di tipo III annulliamo tutti i pivot al di sotto del primo: se $a_{ip} \neq 0$ per $i > 1$, sostituiamo $A^{(i)}$ con

$$A^{(i)} - \frac{a_{ip}}{a_{1p}} A^{(1)}.$$

Passo 3. Tralasciamo (senza cancellare) la riga 1 e torniamo al Passo 1, iterando il procedimento sulle righe successive.

Al termine dell'algoritmo la matrice sarà a gradini.

Oss. Non abbiamo usato operazioni di tipo II ma queste possono comunque essere utili per eventuali semplificazioni.

Esempio.

$$\begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & -2 & 1 \\ 2 & 1 & 0 & 3 \end{pmatrix} \overset{I}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & 1 & 0 \\ 2 & 1 & 0 & 3 \end{pmatrix} \overset{III}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 4 & 1 \end{pmatrix} \overset{I}{\rightsquigarrow}$$

$$\begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 2 & 1 & 0 \end{pmatrix} \overset{I}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & -7 & -2 \end{pmatrix} \overset{II}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 7 & 2 \end{pmatrix}$$

Sistemi lineari arbitrari: metodo di Gauss

Consideriamo ora un sistema lineare qualunque in forma matriciale

$$S : AX = B$$

$$A = (a_{ij}) \in M_{m,n}(\mathbb{K}), \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Matrice completa. La *matrice completa* del sistema

$$\tilde{A} = (A \mid B) \in M_{m,n+1}(\mathbb{K})$$

si ottiene aggiungendo il vettore colonna dei termini noti alla matrice dei coefficienti. Le operazioni elementari sulle righe della matrice completa trasformano il sistema in un sistema *equivalente*. Infatti le operazioni inverse sono operazioni elementari dello stesso tipo, e a livello di sistema le operazioni elementari corrispondono alle seguenti.

Tipo I. Scambiare due equazioni di S.

Tipo II. Moltiplicare un'equazione di S per uno scalare $\alpha \neq 0$.

Tipo III. Sommare membro a membro un'equazione di S con un multiplo scalare di un'altra equazione di S.

Mediante operazioni elementari si trasforma il sistema S in un sistema S' e ciascuno è conseguenza dell'altro. Dunque hanno le stesse soluzioni.

Risoluzione di un sistema lineare. Si applica l'algoritmo di Gauss alla matrice completa del sistema per trasformarla in una matrice a gradini che rappresenta un sistema lineare a gradini equivalente a quello di partenza. Il sistema a gradini si risolve col metodo già studiato.

Esempio.

$$\begin{cases} x + 2y - 4z = 1 \\ 2x - y - z = -1 \end{cases} \left(\begin{array}{ccc|c} 1 & 2 & -4 & 1 \\ 2 & -1 & -1 & -1 \end{array}\right) \overset{III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 2 & -4 & 1 \\ 0 & -5 & 7 & -3 \end{array}\right)$$

Sistema compatibile, z non ha coefficiente pivot $\rightsquigarrow$ un parametro libero.

$$\begin{cases} x + 2y - 4z = 1 \\ -5y + 7z = -3 \end{cases} \quad \begin{cases} x = \frac{6}{5}t - \frac{1}{5} \\ y = \frac{7}{5}t + \frac{3}{5} \\ z = t \end{cases}$$

Le operazioni elementari trasformano un sistema omogeneo in un altro sistema omogeneo, quindi i termini noti se all'inizio sono *tutti* nulli, rimangono nulli. Per questa ragione, nel caso dei sistemi omogenei, possiamo operare sulla matrice dei coefficienti anziché sulla matrice completa.

Esempio. Risolviamo un sistema omogeneo.

$$\begin{cases} x_1 + x_2 - 3x_3 - x_4 = 0 \\ 2x_1 + 2x_3 - 3x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \end{cases} \begin{pmatrix} 1 & 1 & -3 & -1 \\ 2 & 0 & 2 & -3 \\ 0 & 1 & 1 & 1 \end{pmatrix} \overset{I}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 1 \\ 2 & 0 & 2 & -3 \end{pmatrix} \overset{III}{\rightsquigarrow}$$

$$\begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & -2 & 8 & -1 \end{pmatrix} \overset{III}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 10 & 1 \end{pmatrix} \quad \begin{cases} x_1 + x_2 - 3x_3 - x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \\ 10x_3 + x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 = \frac{8}{5}t \\ x_2 = -\frac{9}{10}t \\ x_3 = -\frac{1}{10}t \\ x_4 = t \end{cases}$$

Sistemi con infinite soluzioni

Prop. *Ogni sistema lineare compatibile con più incognite che equazioni ammette infinite soluzioni.*

Dim. L'algoritmo di Gauss non fa aumentare il numero di equazioni, mentre le incognite rimangono le stesse. Il sistema a gradini che si ottiene con Gauss sarà quindi compatibile perché equivalente a quello di partenza e avrà più incognite che equazioni. Pertanto c'è almeno un'incognita senza coefficiente pivot. Nella soluzione generale ci sarà almeno un parametro libero che assumendo infiniti valori determina infinite soluzioni. □

Dato che i sistemi omogenei sono sempre compatibili avendo almeno la soluzione nulla, si ottiene il seguente corollario.

Cor. *Ogni sistema lineare omogeneo con più incognite che equazioni ammette infinite soluzioni.*

Sistemi dipendenti da parametri

Spesso occorre risolvere sistemi che dipendono da uno o più *parametri*. I parametri del sistema non vanno confusi né con le incognite, né con i parametri liberi, coi quali non c'entrano niente, e devono essere indicati con lettere diverse. Cambiando i parametri del sistema, cambia il sistema e quindi le soluzioni. Per contro, cambiando soltanto i parametri liberi (se presenti) si ottengono diverse soluzioni dello stesso sistema.

Esempio. Nel seguente sistema lineare k è un parametro del sistema. Vogliamo trovare tutte le soluzioni del sistema al variare di $k \in \mathbb{R}$.

$$\begin{cases} kx + y - 2z = 0 \\ x + 2y - kz = k \\ y + z = k \end{cases} \quad \left(\begin{array}{ccc|c} k & 1 & -2 & 0 \\ 1 & 2 & -k & k \\ 0 & 1 & 1 & k \end{array}\right) \overset{I}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 2 & -k & k \\ 0 & 1 & 1 & k \\ k & 1 & -2 & 0 \end{array}\right) \overset{III}{\rightsquigarrow}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -k & k \\ 0 & 1 & 1 & k \\ 0 & -2k+1 & k^2-2 & -k^2 \end{array}\right) \overset{III}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 2 & -k & k \\ 0 & 1 & 1 & k \\ 0 & 0 & k^2+2k-3 & k^2-k \end{array}\right)$$

$k^2 + 2k - 3 = (k+3)(k-1)$ è pivot $\Leftrightarrow k \neq -3, 1$.

Si studiano quindi tre casi separatamente: $k = -3$, $k = 1$, $k \neq -3, 1$.

$\underline{k = -3}$ Sistema incompatibile.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & -3 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 12 \end{array}\right)$$

$\underline{k = 1}$ Sistema compatibile con infinite soluzioni, un parametro libero.

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \begin{cases} x + 2y - z = 1 \\ y + z = 1 \end{cases} \quad \begin{cases} x = 3t - 1 \\ y = 1 - t \\ z = t \end{cases}$$

$\underline{k \neq -3, 1}$ Sistema compatibile con un'unica soluzione.

Nell'ultima riga semplifichiamo $k - 1 \neq 0$, dato che $k^2 - k = k(k-1)$.

$$\left(\begin{array}{ccc|c} 1 & 2 & -k & k \\ 0 & 1 & 1 & k \\ 0 & 0 & k+3 & k \end{array}\right) \quad \begin{cases} x + 2y - kz = k \\ y + z = k \\ (k+3)z = k \end{cases} \quad \begin{cases} x = -\dfrac{k}{k+3} \\ y = \dfrac{k(k+2)}{k+3} \\ z = \dfrac{k}{k+3} \end{cases}$$

Determinante

Matrici di ordine 2. Data una matrice quadrata di ordine 2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{K})$$

vogliamo capire in quali casi il sistema lineare omogeneo

$$AX = 0_{\mathbb{K}^2}$$

ammette soltanto la soluzione banale. Per semplicità supponiamo $a \neq 0$. Con Gauss abbiamo

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{III}{\rightsquigarrow} \begin{pmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{pmatrix} \overset{II}{\rightsquigarrow} \begin{pmatrix} a & b \\ 0 & ad - bc \end{pmatrix}$$

Il sistema ammette un'unica soluzione (quella banale) $\Leftrightarrow$ $ad - bc \neq 0$.

Caso generale. Definiremo una funzione $\det : M_n(\mathbb{K}) \to \mathbb{K}$, chiamata *determinante*, che ad ogni matrice quadrata $A \in M_n(\mathbb{K})$ associa uno scalare $\det A \in \mathbb{K}$. Vogliamo che tale funzione abbia le seguenti proprietà:

0) $\det I_n = 1$

1) $\det A' = -\det A$, se A' si ottiene da A scambiando due righe

2) $\det A' = \alpha \det A$, se A' si ottiene da A moltiplicando una riga per $\alpha \in \mathbb{K}$

3) $\det A' = \det A$, se A' si ottiene da A con oper. element. tipo III.

La definizione di det è *ricorsiva*: prima definiamo det per matrici 1×1 (scalari). In secondo luogo, supponendo di saper calcolare det su matrici di ordine $n - 1$, lo definiamo sulle matrici di ordine $n \geqslant 2$.

Indichiamo con $A_{\widehat{ij}} \in M_{n-1}(\mathbb{K})$ la matrice ottenuta da A eliminando la i-esima riga e la j-esima colonna.

N.B. Non confondere $A_{\widehat{ij}} \in M_{n-1}(\mathbb{K})$ con $A_{ij} \in \mathbb{K}$, l'entrata (i, j) di A.

$\underline{n = 1}$. $\det(a) \overset{\text{def}}{=} a$, $\forall\, a \in M_1(\mathbb{K}) = \mathbb{K}$.

$\underline{n \geqslant 2}$. Supponiamo di aver definito det per le matrici di ordine $n - 1 \geqslant 1$. Data una matrice $A \in M_n(\mathbb{K})$ definiamo

$$\det A = |A| \overset{\text{def}}{=} \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det A_{\widehat{1j}}.$$

$\underline{n = 2}$. Riotteniamo la formula ricavata all'inizio.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

$\underline{n = 3}$.

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

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) =$$

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Prop. $\det A = \det {}^tA$.

Oss. Dalla proprietà (*1*) segue che $\det A = 0$ se A ha due righe (o due colonne) uguali. Infatti scambiandole sia ha $\det A = -\det A = 0$.

Matrici triangolari. $A = (a_{ij}) \in M_n(\mathbb{K})$ è *triangolare superiore* se le entrate sotto la diagonale principale sono nulle, ossia $a_{ij} = 0$, $\forall\, i > j$.

Similmente A è *triangolare inferiore* se le entrate sopra la diagonale principale sono nulle, ossia $a_{ij} = 0$, $\forall\, i < j$.

Oss. A a gradini $\not\Leftarrow\Rightarrow$ A triangolare superiore.

Oss. A triangolare superiore $\Leftrightarrow$ ${}^tA$ triangolare inferiore.

Esempio. La matrice seguente è triangolare superiore ma non a gradini.

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

Se $A \in M_n(\mathbb{K})$ è triangolare inferiore, dalla definizione si ha

$$\det A = a_{11} \cdots a_{nn}$$

prodotto delle entrate sulla diagonale principale, perché $a_{1j} = 0$, $\forall\, j > 1$. Questo vale per tutte le matrici triangolari, inferiori o superiori (in particolare per matrici diagonali), dato che det non cambia per trasposizione.

Calcolo di det mediante Gauss. In genere è conveniente usare le proprietà (0)–(3) per calcolare det, riducendo la matrice a gradini (quindi triangolare).

Se A' a gradini è ottenuta da A mediante operazioni elementari di tipo I e III (attenzione! non II) allora

$$\det A = (-1)^k \det A'$$

dove k è il numero di operazioni di tipo I che sono state utilizzate (ciascuna operazione di tipo I fa cambiare il segno a det). Con operazioni di tipo II si possono fare semplificazioni per esempio raccogliendo un coefficiente da una riga o colonna.

Cofattori o complementi algebrici. Data $A = (a_{ij}) \in M_n(\mathbb{K})$, il *cofattore* o *complemento algebrico* dell'entrata (i, j) è

$$\operatorname{cof}_{ij}(A) \stackrel{\text{def}}{=} (-1)^{i+j} \det A_{\widehat{ij}} \in \mathbb{K}.$$

Definiamo la *matrice cofattore* $\operatorname{cof}(A) \stackrel{\text{def}}{=} (\operatorname{cof}_{ij}(A)) \in M_n(\mathbb{K})$.

Formule di Laplace. Per ogni $i = 1, \dots, n$ si ha lo *sviluppo di* $\det A$ *secondo la* i*-esima riga*

$$\det A = \sum_{j=1}^{n} A_{ij} \operatorname{cof}_{ij}(A).$$

Per ogni $j = 1, \dots, n$ si ha lo *sviluppo di* $\det A$ *secondo la* j*-esima colonna*

$$\det A = \sum_{i=1}^{n} A_{ij} \operatorname{cof}_{ij}(A).$$

Oss. La formula ricorsiva usata per definire det è lo sviluppo di Laplace secondo la prima riga.

Conviene sviluppare $\det A$ secondo la riga o colonna col maggior numero di entrate nulle.

Gruppo lineare generale

Def. $\mathrm{GL}_n(\mathbb{K}) \stackrel{\text{def}}{=} \{M \in M_n(\mathbb{K}) \mid M \text{ invertibile}\}$ si chiama *gruppo lineare generale di ordine* n *su* $\mathbb{K}$. È l'insieme delle matrici invertibili $n \times n$.

Oss.

1) $I_n \in \mathrm{GL}_n(\mathbb{K})$.

2) $A \in \mathrm{GL}_n(\mathbb{K}) \Rightarrow A^{-1} \in \mathrm{GL}_n(K)$.

3) $A, B \in \mathrm{GL}_n(\mathbb{K}) \Rightarrow AB \in \mathrm{GL}_n(K)$ e si ha

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Infatti: $(AB)(B^{-1}A^{-1}) = I_n$.

Matrici simili

Def. Due matrici $A, B \in M_n(\mathbb{K})$ sono *simili* se $\exists\, P \in \mathrm{GL}_n(\mathbb{K})$ t.c.

$$B = P^{-1}AP.$$

Oss. I_n è simile solo a sé stessa: $P^{-1}I_nP = P^{-1}P = I_n$, $\forall\, P \in \mathrm{GL}_n(\mathbb{K})$. Analogamente la matrice nulla $0 \in M_n(\mathbb{K})$ è simile solo a sé stessa.

Teorema di Binet

Teor. $\forall\, A, B \in M_n(\mathbb{K}) \Rightarrow \det(AB) = \det(A)\det(B)$.

Cor. $\forall\, A \in \mathrm{GL}_n(\mathbb{K}) \Rightarrow \det(A^{-1}) = \dfrac{1}{\det A}$.

Dim. $1 = \det I_n = \det(AA^{-1}) = \det(A)\det(A^{-1})$, da cui la tesi. □

Cor. *Matrici simili hanno lo stesso determinante.*

Dim. $\det(P^{-1}AP) = \det(P)^{-1}\det(A)\det(P) = \det(A)$. □

Matrice inversa

Il metodo più efficiente per calcolare la matrice inversa di $A \in M_n(\mathbb{K})$ è il seguente: si applica il metodo di Gauss alla matrice $(A \mid I_n)$, portando A a gradini e se vi sono n pivot si può applicare di nuovo Gauss a ritroso per ottenere I_n nelle prime n colonne. Le colonne restanti formano A^{-1}.

$$(A \mid I_n) \overset{\text{Gauss}}{\rightsquigarrow} (I_n \mid A^{-1}).$$

Un altro metodo

Teor. $\forall\, A \in M_n(\mathbb{K})$ *si ha* $A\,{}^t\mathrm{cof}(A) = (\det A) I_n$.

Dim. Dimostriamo che $(A\,{}^t\mathrm{cof}(A))_{ij} = ((\det A) I_n)_{ij}$, $\forall\, i, j = 1, \dots, n$.

$\boxed{i = j}$
$$\begin{aligned}(A\,{}^t\mathrm{cof}(A))_{ii} &= \sum_{k=1}^{n} A_{ik} ({}^t\mathrm{cof}(A))_{ki} \\ &= \sum_{k=1}^{n} A_{ik}\, \mathrm{cof}_{ik}(A) \qquad \text{(Laplace su riga } i\text{)} \\ &= \det A = ((\det A) I_n)_{ii}\end{aligned}$$

$\boxed{i \neq j}$ Consideriamo la matrice $B \in M_n(\mathbb{K})$ ottenuta da A duplicando la riga i-esima al posto della j-esima

$$B^{(k)} = \begin{cases} A^{(k)}, & \text{se } k \neq j \\ A^{(i)}, & \text{se } k = j \end{cases}$$

B ha due righe uguali $B^{(i)} = B^{(j)} = A^{(i)} \Rightarrow \det B = 0$.
$A_{\widehat{jk}} = B_{\widehat{jk}} \Rightarrow \mathrm{cof}_{jk}(A) = \mathrm{cof}_{jk}(B)$ perché la riga j viene cancellata.

$$\begin{aligned}(A\,{}^t\mathrm{cof}(A))_{ij} &= \sum_{k=1}^{n} A_{ik} ({}^t\mathrm{cof}(A))_{kj} \\ &= \sum_{k=1}^{n} A_{ik}\, \mathrm{cof}_{jk}(A) \\ &= \sum_{k=1}^{n} B_{jk}\, \mathrm{cof}_{jk}(B) \qquad \text{(Laplace riga } j\text{)} \\ &= \det B = 0 = ((\det A) I_n)_{ij}.\end{aligned}$$

□

Cor. $\det A \neq 0 \Rightarrow A \in \mathrm{GL}_n(K)$ *e* $A^{-1} = \dfrac{1}{\det A}\,{}^t\mathrm{cof}\, A$.

Dim. $A \left(\frac{1}{\det A}\,{}^t\mathrm{cof}\, A\right) = I_n$. □

Cor. *Data* $A \in M_n(\mathbb{K})$ *si ha* $A \in \mathrm{GL}_n(\mathbb{K}) \Leftrightarrow \det A \neq 0$.

Teorema di Cramer

Un sistema lineare

$$AX = B$$

con $A \in \mathrm{GL}_n(\mathbb{K})$, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$ e $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, ha l'unica soluzione

$$X = A^{-1}B$$

che si ottiene moltiplicando a sinistra entrambi i membri per A^{-1}.
Il teorema seguente fornisce una formula per calcolare la soluzione.

Teor. *Siano $A \in \mathrm{GL}_n(\mathbb{K})$ e $B \in \mathbb{K}^n$. La soluzione del sistema lineare*

$$AX = B$$

è data da

$$x_i = \frac{\det \widehat{A}_i}{\det A} \quad \forall\, i = 1, \ldots, n$$

dove $\widehat{A}_i$ è la matrice ottenuta da A sostituendo B al posto della i-esima colonna.

Dim. $X = A^{-1}B$ da cui si ha:

$$\begin{aligned} x_i &= \sum_{k=1}^{n} (A^{-1})_{ik} b_k = \frac{1}{\det A} \sum_{k=1}^{n} ({}^t\mathrm{cof}(A))_{ik} b_k \\ &= \frac{1}{\det A} \sum_{k=1}^{n} b_k \,\mathrm{cof}(A)_{ki} \\ &= \frac{1}{\det A} \sum_{k=1}^{n} b_k \,\mathrm{cof}(\widehat{A}_i)_{ki} \qquad \text{(Laplace su colonna } i\text{)} \\ &= \frac{\det \widehat{A}_i}{\det A}. \end{aligned}$$

□

Spazi vettoriali

Indichiamo con $\mathbb{K}$ uno dei campi $\mathbb{R}$ o $\mathbb{C}$ (oppure un campo qualsiasi).

Def (Assiomi di spazio vettoriale)**.** Un $\mathbb{K}$-*spazio vettoriale* è un insieme *non vuoto* V dove sono date due funzioni dette *operazioni*

$$+ : V \times V \to V \qquad \cdot : \mathbb{K} \times V \to V$$
$$(v, w) \mapsto v + w \qquad (\alpha, v) \mapsto \alpha v$$

dette rispettivamente *addizione* e *moltiplicazione (per uno) scalare*, che soddisfano i seguenti assiomi $\forall\, u, v, w \in V$, $\forall\, \alpha, \beta \in \mathbb{K}$:

1)	$(u + v) + w = u + (v + w)$	*Proprietà associativa per l'addizione*
2)	$\exists\, 0_V \in V$ t.c. $v + 0_V = v$	*Vettore nullo*
3)	$\exists\, -v \in V$ t.c. $v - v = 0_V$	*Vettore opposto*
4)	$u + v = v + u$	*Proprietà commutativa per l'addizione*
5)	$\alpha(u + v) = \alpha u + \alpha v$	*Proprietà distributiva vettoriale*
6)	$(\alpha + \beta)v = \alpha v + \beta v$	*Proprietà distributiva scalare*
7)	$(\alpha\beta)v = \alpha(\beta v)$	*Proprietà associativa scalare*
8)	$1v = v$	*Elemento neutro moltiplicativo*

Gli elementi di V sono detti *vettori* e gli elementi di $\mathbb{K}$ sono detti *scalari*.

Un $\mathbb{K}$-spazio vettoriale è detto anche *spazio vettoriale su* $\mathbb{K}$. Se consideriamo $\mathbb{K} = \mathbb{R}$ (risp. $\mathbb{K} = \mathbb{C}$) diciamo anche spazio vettoriale reale (risp. complesso). Scriviamo anche 0 anziché 0_V senza confonderlo con $0 \in \mathbb{K}$.

Esempio. I seguenti sono $\mathbb{K}$-spazi vettoriali con le operazioni di addizione e moltiplicazione scalare definite in precedenza. La validità degli assiomi segue subito dalle proprietà dei campi e dei polinomi.

$\mathbb{K}^n$ (ossia $\mathbb{R}^n$ o $\mathbb{C}^n$) $\mathbb{K}$-spazio vettoriale numerico di dimensione n.
$M_{m,n}(\mathbb{K})$ spazio delle matrici $m \times n$ a entrate in $\mathbb{K}$.
$M_n(\mathbb{K})$ spazio delle matrici quadrate di ordine n a entrate in $\mathbb{K}$.
$\mathbb{K}[x]$ spazio dei polinomi a coefficienti in $\mathbb{K}$.

Oss. Per gli assiomi (*1*) e (*4*) possiamo scrivere

$$v_1 + \cdots + v_n$$

senza parentesi e in un ordine qualsiasi $\forall\, n \in \mathbb{N}$, $\forall\, v_1, \ldots, v_n \in V$.

Prop. *Per qualunque spazio vettoriale V valgono le seguenti proprietà,* $\forall\, u, v, w \in V$, $\forall\, \alpha \in \mathbb{K}$.

a) *Il vettore nullo* 0_V *è unico.*

b) $u + w = v + w \Rightarrow u = v$ *(legge di cancellazione).*

c) $0v = 0_V$.

d) $\alpha 0_V = 0_V$.

e) $\alpha v = 0_V \Leftrightarrow \alpha = 0$ *oppure* $v = 0_V$.

f) *L'opposto* $-v$ *è unico.*

g) $-v = (-1)v$.

Dim.

a) Supponiamo $0'_V \in V$ altro vettore nullo: $0_V = 0_V + 0'_V = 0'_V$.

b) Basta sommare a entrambi i membri $-w$.

c) $0v = (0 + 0)v = 0v + 0v \Rightarrow 0_V = 0v$.

d) $\alpha 0_V = \alpha(00_V) = (\alpha 0)0_V = 00_V = 0_V$.

e) $\alpha v = 0_V$ e $\alpha \neq 0 \Rightarrow 0_V = \alpha^{-1}0_V = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1v = v$.

f) $v' \in V$ altro opposto di $v \Rightarrow v + v' = 0_V = v - v \Rightarrow v' = -v$ per (b).

g) $0_V = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$. □

Combinazioni lineari. Consideriamo un $\mathbb{K}$-spazio vettoriale V, vettori $v_1, \ldots, v_n \in V$ e scalari $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, con $n \in \mathbb{N}$. Il vettore

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$$

è detto *combinazione lineare* di $v_1, \ldots, v_n$ con *coefficienti* $\alpha_1, \ldots, \alpha_n$. Scriveremo anche

$$u = \sum_{i=1}^{n} \alpha_i v_i.$$

Una combinazione lineare è *nulla* se ha come risultato il vettore nullo:

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V.$$

Una combinazione lineare è *banale* se tutti i coefficienti sono nulli. Una combinazione lineare banale è nulla ma il viceversa non è sempre vero.

Esempio. Combinazione lineare nulla non banale in $\mathbb{R}^2$

$$2\begin{pmatrix}1\\-1\end{pmatrix} + \begin{pmatrix}3\\2\end{pmatrix} - 5\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

Dipendenza lineare. I vettori $v_1, \ldots, v_n \in V$ sono *linearmente dipendenti* se esiste una loro combinazione lineare nulla non banale, ossia se esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ non tutti nulli t.c.

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V.$$

Indipendenza lineare. I vettori $v_1, \dots, v_n \in V$ sono *linearmente indipendenti* se ammettono un'unica combinazione lineare nulla, ossia quella banale. In altri termini $v_1, \dots, v_n$ sono linearmente indipendenti se e solo se $\forall\, \alpha_1, \dots, \alpha_n \in \mathbb{K}$ si ha:

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V \;\Rightarrow\; \alpha_1 = \cdots = \alpha_n = 0.$$

Questa è l'implicazione che va dimostrata per far vedere che i vettori dati sono linearmente indipendenti. Si noti che l'implicazione inversa è banale (sempre verificata), quindi non serve.

Pertanto le due nozioni di dipendenza e indipendenza lineare sono complementari (o l'una o l'altra).

I vettori dell'esempio precedente sono linearmente dipendenti.

Esempio. Consideriamo i seguenti vettori di $\mathbb{R}^2$

$$\begin{pmatrix}1\\3\end{pmatrix},\ \begin{pmatrix}3\\-2\end{pmatrix}$$

e proviamo a capire se sono linearmente indipendenti oppure no. Dobbiamo se esistono due scalari non entrambi nulli $x, y \in \mathbb{R}$ t.c.

$$x\begin{pmatrix}1\\3\end{pmatrix} + y\begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \qquad \begin{cases} x + 3y = 0 \\ 3x - 2y = 0 \end{cases}$$

$$\begin{pmatrix}1 & 3\\3 & -2\end{pmatrix} \rightsquigarrow \begin{pmatrix}1 & 3\\0 & -11\end{pmatrix} \qquad \begin{cases} x + 3y = 0 \\ -11y = 0 \end{cases} \qquad \begin{cases} x = 0 \\ y = 0 \end{cases}$$

L'unica soluzione è quella banale, quindi i vettori sono linearmente indipendenti.

Per stabilire la dipendenza o indipendenza lineare di vettori in $\mathbb{K}^n$ possiamo impostare un sistema omogeneo, risolverlo e vedere se ci sono soluzioni non nulle. Più avanti vedremo un metodo più veloce.

Oss. Se uno dei vettori $v_1, \dots, v_n$ è nullo allora questi vettori sono linearmente dipendenti. Infatti se $v_1 = 0_V$ si ha:

$$1 \cdot 0_V + 0v_2 + \cdots + 0v_n = 0_V$$

è combinazione lineare nulla non banale essendo 1 il primo coefficiente.

Vettori proporzionali. $v, w \in V$ sono *proporzionali* se $\exists\, \alpha \in \mathbb{K}$ t.c.

$$v = \alpha w \quad \text{oppure} \quad w = \alpha v.$$

Prop. *$v, w \in V$ sono linearmente dipendenti $\Leftrightarrow$ v, w sono proporzionali.*

Dim. $\boxed{\Rightarrow}$ $\exists\, \alpha, \beta \in \mathbb{K}$ non entrambi nulli t.c. $\alpha v + \beta w = 0_V$.
$\beta \neq 0 \Rightarrow w = -\frac{\alpha}{\beta} v$. $\alpha \neq 0 \Rightarrow v = -\frac{\beta}{\alpha} w$. Quindi sono proporzionali.

$\boxed{\Leftarrow}$ $\exists\, \alpha \in \mathbb{K}$ t.c. $v = \alpha w \Rightarrow v - \alpha w = 0_V$ oppure $w = \alpha v \Rightarrow \alpha v - w = 0_V$.
In entrambi i casi otteniamo una combinazione lineare nulla non banale. Quindi sono linearmente dipendenti. □

Esempio.

$$v = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, \ w = \begin{pmatrix} 6 \\ 21 \end{pmatrix} \in \mathbb{R}^2$$

sono proporzionali: $w = 3v$. Quindi sono linearmente dipendenti.

Esempio.

$$v = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, \ w = \begin{pmatrix} -1 \\ 4 \end{pmatrix} \in \mathbb{R}^2$$

non sono proporzionali. Quindi sono linearmente indipendenti.

N.B. Questo metodo per stabilire la lineare dipendenza/indipendenza è valido se i vettori sono due, non se sono più di due, come illustrato nel primo esempio sopra e in quello seguente.

Esempio. I seguenti vettori di $\mathbb{R}^3$ sono linearmente dipendenti?

$$\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \ \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix}$$

$$x \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + y \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + z \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{cases} 2x + y + 4z = 0 \\ x + 2y - z = 0 \\ 2x - y + 8z = 0 \end{cases}$$

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 2 & -1 & 8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 6 \\ 0 & -5 & 10 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{cases} x + 2y - z = 0 \\ y - 2z = 0 \end{cases} \quad \begin{cases} x = -3t \\ y = 2t \\ z = t \end{cases} \quad -3 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Il sistema ammette soluzioni non nulle quindi i vettori dati sono linearmente dipendenti.

Sottospazi vettoriali

Def. Consideriamo un $\mathbb{K}$-spazio vettoriale V. Un sottoinsieme $U \subseteq V$ è detto *sottospazio vettoriale* se $\forall\, u, v \in U$, $\forall\, \alpha \in \mathbb{K}$ si ha:

1) $U \neq \emptyset$ (U non vuoto).
2) $u + v \in U$ (U chiuso rispetto all'addizione).
3) $\alpha u \in U$ (U chiuso rispetto a prodotto per scalari).

Teor. *Sia $U \subset V$ sottospazio vettoriale. Allora $0_V \in U$. Inoltre U è un $\mathbb{K}$-spazio vettoriale rispetto alle operazioni ereditate da V.*

Dim. $U \neq \emptyset \Rightarrow \exists\, u \in U \overset{(3)}{\Rightarrow} 0u = 0_V \in U$.
$\forall\, u \in U \Rightarrow -u = (-1)u \in U$.
U è chiuso rispetto all'addizione e al prodotto per scalari, contiene il vettore nullo e l'opposto di ogni suo vettore. Valgono gli assiomi di spazio vettoriale perché discendono da V. Quindi U è un $\mathbb{K}$-spazio vettoriale. □

Oss. Il sottoinsieme $\{0_V\} \subset V$ che contiene solo il vettore nullo è un sottospazio vettoriale di V detto *sottospazio nullo* e indicato con 0 (non confondere il simbolo 0 nelle sue varie accezioni).

Oss. Sia $U \subset V$ sottospazio vettoriale. Dati $n \in \mathbb{N}$, $u_1, \ldots, u_n \in U$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{K} \Rightarrow \alpha_1 u_1 + \cdots + \alpha_n u_n \in U$ (infatti U è $\mathbb{K}$-spazio vettoriale).

Prop. *Sia $U \subseteq V$ un sottoinsieme non vuoto di un $\mathbb{K}$-spazio vettoriale V. Allora U è sottospazio vettoriale di $V \Leftrightarrow \forall\, u, v \in U$, $\forall\, \alpha, \beta \in \mathbb{K}$ si ha:*

2′) $\alpha u + \beta v \in U$.

Dim. La (*1*) è soddisfatta per ipotesi.
$\boxed{\Rightarrow}$ Supponiamo $U \subseteq V$ sottospazio vettoriale $\Rightarrow U$ contiene tutte le combinazioni lineari di suoi vettori (vedi Oss. precedente).

$\boxed{\Leftarrow}$ Supponiamo che U soddisfi la (*2′*). Si ottiene (*2*) ponendo $\alpha = \beta = 1$ nella (*2′*). Si ottiene (*3*) ponendo $\beta = 0$ nella (*2′*). □

Spazio delle soluzioni di un sistema omogeneo. Dato un sistema lineare omogeneo

$$S : AX = 0_{\mathbb{K}^m}$$

con $A \in M_{m,n}(\mathbb{K})$, lo spazio delle soluzioni è il sottoinsieme

$$\Sigma_S = \{u \in \mathbb{K}^n \mid Au = 0_{\mathbb{K}^m}\} \subset \mathbb{K}^n.$$

Teor. *Lo spazio delle soluzioni di un sistema lineare omogeneo in n incognite è un sottospazio vettoriale di $\mathbb{K}^n$.*

Dim. $0_{\mathbb{K}^n} \in \Sigma_S \neq \emptyset$. $\forall\, u, v \in \Sigma_S$, $\forall\, \alpha, \beta \in \mathbb{K}$ si ha:
$A(\alpha u + \beta v) = A\alpha u + A\beta v = \alpha Au + \beta Av = 0_{\mathbb{K}^m} \Rightarrow \alpha u + \beta v \in \Sigma_S$. □

N.B. Se il sistema non è omogeneo allora lo spazio delle soluzioni non è un sottospazio vettoriale perché non contiene il vettore nullo (e se il sistema è incompatibile lo spazio delle soluzioni è vuoto).

Span. Consideriamo uno spazio vettoriale V e vettori $v_1, \dots, v_n \in V$. Definiamo il sottoinsieme

$$\text{span}(v_1, \dots, v_n) \stackrel{\text{def}}{=} \{\alpha_1 v_1 + \cdots + \alpha_n v_n \mid \alpha_1, \dots, \alpha_n \in \mathbb{K}\} \subset V.$$

Gli elementi di $\text{span}(v_1, \dots, v_n)$ sono tutte e sole le combinazioni lineari di $v_1, \dots, v_n$, al variare dei coefficienti $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ in tutti i modi.

Per definizione poniamo anche $\text{span}(\emptyset) = \{0\}$, lo *spazio vettoriale nullo*.

Oss.

1) $v \in \text{span}(v_1, \dots, v_n) \Leftrightarrow \exists\, \alpha_1, \dots, \alpha_n \in \mathbb{K}$ t.c. $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

2) $\text{span}(0_V) = \{0_V\} = \text{span}(\emptyset)$ spazio vettoriale nullo.

3) $\text{span}(v) = \{\alpha v \mid \alpha \in \mathbb{K}\}$ insieme di tutti i multipli scalari di v.

Prop. $\text{span}(v_1, \dots, v_n)$ *è un sottospazio vettoriale di* V.

Dim. Verifichiamo le proprietà della definizione.

1) $0_V = 0v_1 + \cdots + 0v_n \in \text{span}(v_1, \dots, v_n) \neq \emptyset$.

2) $\forall\, u, v \in \text{span}(v_1, \dots, v_n)$ si ha per certi α_i e $\beta_i \in \mathbb{K}$

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n$$
$$v = \beta_1 v_1 + \cdots + \beta_n v_n$$
$$u + v = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_n + \beta_n)v_n \in \text{span}(v_1, \dots, v_n).$$

3) $\forall\, \beta \in \mathbb{K}$, $\forall\, u \in \text{span}(v_1, \dots, v_n)$ scritto come sopra si ha

$$\beta u = (\beta\alpha_1)v_1 + \cdots + (\beta\alpha_n)v_n \in \text{span}(v_1, \dots, v_n). \qquad \square$$

Def. $\text{span}(v_1, \dots, v_n)$ è detto *sottospazio vettoriale generato da* $v_1, \dots, v_n$ e questi vettori sono detti *generatori*. Si indica anche con $\langle v_1, \dots, v_n \rangle$.

Oss. $v_1, \dots, v_n \in \text{span}(v_1, \dots, v_n)$, $\forall\, i = 1, \dots, n$.

Oss. $\text{span}(v_1, \dots, v_n)$ è il più piccolo sottospazio vettoriale di V che contiene $v_1, \dots, v_n$, ossia $\forall\, U \subseteq V$ sottospazio vettoriale t.c. $v_1 \dots, v_n \in U \Rightarrow \text{span}(v_1, \dots, v_n) \subseteq U$.

Def. V è *finitamente generato* se $\exists\, n \in \mathbb{N}$ e $\exists\, v_1 \dots, v_n \in V$ t.c.

$$V = \text{span}(v_1, \dots, v_n).$$

$v_1, \dots, v_n$ sono detti *generatori* di V.

Basi di spazi vettoriali

Def. Una *base* per V è un insieme di generatori linearmente indipendenti.

Oss. $(v_1, \dots, v_n)$ base per $V \Leftrightarrow V = \text{span}(v_1, \dots, v_n)$ e $v_1, \dots, v_n$ sono linearmente indipendenti.

Base canonica di $\mathbb{K}^n$. Consideriamo i seguenti vettori di $\mathbb{K}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

colonne di $I_n \Rightarrow e_1, \dots, e_n$ linearmente indipendenti (perché $\text{rg}\, I_n = n$).

$$\forall\, v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n \;\Rightarrow\; v = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

$\Rightarrow \mathbb{K}^n = \text{span}(e_1, \dots, e_n)$. Pertanto $e_1, \dots, e_n$ formano una base per $\mathbb{K}^n$.

Def. $\mathcal{E}_n := (e_1, \dots, e_n)$ è detta *base canonica* per $\mathbb{K}^n$.

Oss. $\mathbb{K}^n$ è finitamente generato.

Esempio.

$\underline{n = 1}$. $e_1 = 1 \rightsquigarrow \mathcal{E}_1 = (e_1)$ base canonica per $\mathbb{K}$.

$\underline{n = 2}$. $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightsquigarrow \mathcal{E}_2 = (e_1, e_2)$ base canonica per $\mathbb{K}^2$.

$\underline{n = 3}$. $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rightsquigarrow \mathcal{E}_3 = (e_1, e_2, e_3)$.

Prop. *$v_1, \dots, v_n \in V$ linearmente dipendenti $\Leftrightarrow$ uno di essi è combinazione lineare degli altri.*

Dim. $\boxed{\Rightarrow}$ $\exists\, \alpha_1, \dots \alpha_n \in \mathbb{K}$ non tutti nulli t.c.

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V.$$

A meno di riordinare i vettori possiamo assumere $\alpha_n \neq 0 \Rightarrow$

$$v_n = -\alpha_n^{-1}(\alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1})$$

$\Rightarrow v_n$ combinazione lineare di $v_1, \dots, v_{n-1}$.

$\boxed{\Leftarrow}$ A meno di riordinarli possiamo assumere v_n combinazione lineare di $v_1, \dots, v_{n-1} \Rightarrow \exists\, \alpha_1, \dots, \alpha_{n-1} \in \mathbb{K}$ t.c. $v_n = \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} \Rightarrow$

$$\alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} - v_n = 0_V$$

combinazione lineare nulla non banale $\Rightarrow v_1, \dots, v_n$ lin. indipendenti. $\square$

Lem. $v \in \text{span}(v_1, \ldots, v_n) \Leftrightarrow \text{span}(v, v_1, \ldots, v_n) = \text{span}(v_1, \ldots, v_n)$.

Dim. $\boxed{\Leftarrow}$ Evidente perché $v \in \text{span}(v, v_1, \ldots, v_n)$.

$\boxed{\Rightarrow}$ Ogni combinazione lineare di $v_1, \ldots, v_n$ è anche combinazione lineare di $v, v_1, \ldots, v_n$ ponendo 0 come coefficiente di $v \Rightarrow$

$$\text{span}(v_1, \ldots, v_n) \subset \text{span}(v, v_1, \ldots, v_n).$$

Dimostriamo l'altra inclusione.
$v \in \text{span}(v_1, \ldots, v_n) \Rightarrow \exists\, \alpha_1 \ldots, \alpha_n \in \mathbb{K}$ t.c.

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

$\forall\, w \in \text{span}(v, v_1, \ldots, v_n) \Rightarrow \exists\, \beta, \beta_1, \ldots, \beta_n \in \mathbb{K}$ t.c.

$$\begin{aligned} w &= \beta v + \beta_1 v_1 + \cdots + \beta_n v_n \\ &= \beta(\alpha_1 v_1 + \cdots + \alpha_n v_n) + \beta_1 v_1 + \cdots + \beta_n v_n \\ &= (\beta\alpha_1 + \beta_1)v_1 + \cdots + (\beta\alpha_n + \beta_n)v_n \in \text{span}(v_1, \ldots, v_n) \\ &\Rightarrow \text{span}(v, v_1, \ldots, v_n) \subset \text{span}(v_1, \ldots, v_n). \end{aligned}$$

Dalle due inclusioni si ha l'uguaglianza. □

Oss. Un generatore può essere eliminato $\Leftrightarrow$ è combinazione lin. degli altri.

Teor. *Ogni spazio vettoriale finitamente generato ammette una base.*

Dim. Supponiamo $V = \text{span}(v_1, \ldots, v_n)$. Facciamo induzione su $n \geqslant 0$.

Base dell'induzione: $n = 0$. Non ci sono generatori e quindi $V = \{0\}$ è lo spazio vettoriale nullo avente l'insieme vuoto come base.

Ipotesi induttiva. Supponiamo che ogni spazio vettoriale con $n - 1 \geqslant 0$ generatori ammetta una base e sia V generato da n vettori $v_1, \ldots, v_n$.

Se $v_1, \ldots, v_n$ sono linearmente indipendenti allora formano base per V.

Se invece sono linearmente dipendenti, per la Prop. un certo v_k è combinazione lineare degli altri v_i, e per il Lem. v_k può essere eliminato. Quindi V ammette $n - 1$ generatori. Per l'ipotesi induttiva V ammette base. □

Oss. Ogni insieme finito di generatori per V contiene una base per V.

Coordinate

Teor. *Sia V un $\mathbb{K}$-spazio vettoriale. I vettori $b_1, \dots, b_n \in V$ formano una base per $V \Leftrightarrow \forall\, v \in V$ esistono unici $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ t.c.*

$$v = \alpha_1 b_1 + \cdots + \alpha_n b_n.$$

Dim. $\boxed{\Rightarrow}$ $b_1, \dots, b_n \in V$ generatori $\Rightarrow \forall\, v \in V\ \exists\ \alpha_1, \dots, \alpha_n \in \mathbb{K}$ t.c.

$$v = \alpha_1 b_1 + \cdots + \alpha_n b_n.$$

Mostriamo che i coefficienti sono unici. Supponiamo che

$$v = \alpha_1 b_1 + \cdots + \alpha_n b_n$$
$$v = \beta_1 b_1 + \cdots + \beta_n b_n$$

da cui sottraendo membro a membro

$$(\alpha_1 - \beta_1) b_1 + \cdots + (\alpha_n - \beta_n) b_n = 0_V.$$

Dato che $b_1, \dots, b_n$ sono lin. indip. i coefficienti sono tutti nulli

$$\begin{cases} \alpha_1 - \beta_1 = 0 \\ \cdots\cdots\cdots \\ \alpha_n - \beta_n = 0 \end{cases} \Rightarrow \begin{cases} \alpha_1 = \beta_1 \\ \cdots\cdots \\ \alpha_n = \beta_n. \end{cases}$$

Quindi i coefficienti corrispondenti sono uguali, da cui si ha l'unicità.

$\boxed{\Leftarrow}$ Per ipotesi $b_1, \dots, b_n$ sono generatori di V. Inoltre 0_V si scrive in un unico modo come combinazione lin. di $b_1, \dots, b_n$: la combinazione lin. banale. Quindi $b_1, \dots, b_n$ sono linearmente indipendenti. □

Def. Sia $\mathcal{B} = (b_1, \dots, b_n)$ una base per V e sia $v = \alpha_1 b_1 + \cdots + \alpha_n b_n \in V$. I coefficienti $\alpha_1, \dots, \alpha_n$ sono detti *coordinate* o *componenti* di v rispetto alla base $\mathcal{B}$. Le scriviamo come n-upla

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n.$$

Scriviamo anche

$$v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}^{\mathcal{B}} = \alpha_1 b_1 + \cdots + \alpha_n b_n.$$

N.B. Le componenti di un vettore dipendono dalla base, oltre che dal vettore stesso.

Data una base $\mathcal{B} = (b_1, \dots, b_n)$ per un $\mathbb{K}$-spazio vettoriale V, definiamo l'*applicazione coordinata*

$$\Phi_{\mathcal{B}} : V \to \mathbb{K}^n$$
$$\Phi_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad v = \alpha_1 b_1 + \cdots + \alpha_n b_n$$

$\Phi_{\mathcal{B}}$ è invertibile (quindi biiettiva) e l'inversa è l'applicazione

$$\Phi_{\mathcal{B}}^{-1} : \mathbb{K}^n \to V$$

$$\Phi_{\mathcal{B}}^{-1}\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = \alpha_1 b_1 + \cdots + \alpha_n b_n$$

Oss. Le componenti del vettore nullo sono tutte nulle:

$$0_V = \begin{pmatrix}0\\ \vdots \\ 0\end{pmatrix}^{\mathcal{B}}.$$

Oss. Fissata una base $\mathcal{B} = (b_1, \ldots, b_n)$ e dati $u, v \in V$ e $\lambda \in \mathbb{K}$ abbiamo

$$\begin{aligned} u &= \alpha_1 b_1 + \cdots + \alpha_n b_n \\ v &= \beta_1 b_1 + \cdots + \beta_n b_n \\ u + v &= (\alpha_1 + \beta_1) b_1 + \cdots + (\alpha_n + \beta_n) b_n \\ \lambda u &= (\lambda \alpha_1) b_1 + \cdots + (\lambda \alpha_n) b_n \end{aligned}$$

Segue che le combinazioni lineari di vettori di V corrispondono a combinazioni lineari dei vettori delle coordinate.

Dimensione

Teor. *Sia $\mathcal{B} = (b_1, \dots, b_n)$ base per V e siano $u_1, \dots, u_k \in V$ vettori linearmente indipendenti. Allora $k \leqslant n$.*

Dim. Per assurdo supponiamo $k > n$. Coordinate rispetto a $\mathcal{B} \rightsquigarrow$

$$u_1 = \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{n1} \end{pmatrix}^{\mathcal{B}}, \dots, u_k = \begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{nk} \end{pmatrix}^{\mathcal{B}}$$

$$x_1 u_1 + \cdots + x_k u_k = 0_V$$

deve avere come unica soluzione quella nulla. In coordinate equivale a

$$x_1 \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{n1} \end{pmatrix} + \cdots + x_k \begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{nk} \end{pmatrix} = 0_{\mathbb{K}^n}$$

$$\begin{cases} \alpha_{11} x_1 + \cdots + \alpha_{1k} x_k = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{n1} x_1 + \cdots + \alpha_{nk} x_k = 0 \end{cases}$$

sistema omogeneo di n equazioni in k incognite.
$k > n \Rightarrow \exists$ soluzione non nulla $\Rightarrow u_1, \dots, u_k$ linearmente dipendenti, contraddizione. Pertanto si deve avere $k \leqslant n$. □

Teor. *Due basi qualsiasi di uno spazio vettoriale finitamente generato hanno lo stesso numero di vettori.*

Dim. Siano $\mathcal{B} = (b_1, \dots, b_n)$, $\mathcal{C} = (c_1, \dots, c_k)$ basi di V. Per il teorema, $k \leqslant n$ dato che $c_1, \dots, c_k$ sono linearmente indipendenti. Scambiando il ruolo delle due basi si ha anche $n \leqslant k$. Quindi $n = k$. □

Def. Il numero di vettori di una base per V si chiama *dimensione* di V e si indica con $\dim_{\mathbb{K}} V$. Se V non è finitamente generato si pone $\dim_{\mathbb{K}} V = \infty$.

N.B. Scriviamo semplicemente $\dim V$ quando $\mathbb{K}$ è chiaro dal contesto.

Oss. $\mathbb{K}^n$ ha la base canonica $\mathcal{E}_n = (e_1, \dots, e_n) \Rightarrow \dim_{\mathbb{K}} \mathbb{K}^n = n$.

Oss. V finitamente generato $\Leftrightarrow \dim V < \infty$.

Oss. $\dim V = 0 \Leftrightarrow V = \{0\}$ (spazio vettoriale nullo).

Oss. $\dim V = 1 \Leftrightarrow \exists\, v \in V$ non nullo t.c. $V = \operatorname{span}(v) = \{\alpha v \mid \alpha \in \mathbb{K}\}$ e V è anche detto *retta vettoriale*.

N.B. La dimensione è un numero naturale (oppure ∞) ed è molto importante. Occorre sempre tenerne conto e imparare a calcolarla.

Il seguente teorema lo enunciamo senza dimostrazione.

Teor. *Se $\dim V < \infty$ e $U \subset V$ sottospazio vettoriale $\Rightarrow \dim U \leqslant \dim V$. Vale uguaglianza $\Leftrightarrow U = V$.*

Oss. $\dim V =$ massimo numero di vettori linearmente indipendenti di V.

Cor. *Sia* $\dim V = n$ *e siano* $v_1, \dots, v_n \in V$*. Le seguenti sono equivalenti:*

a) $(v_1, \dots, v_n)$ *linearmente indipendenti;*

b) $(v_1, \dots, v_n)$ *generano* V*;*

c) $(v_1, \dots, v_n)$ *base per* V*.*

Dim. $\boxed{a \Rightarrow b}$ $\dim(\mathrm{span}(v_1, \dots, v_n)) = \dim V \Rightarrow \mathrm{span}(v_1, \dots, v_n) = V$.

$\boxed{b \Rightarrow c}$ $(v_1, \dots, v_n)$ contiene base con n vettori $\Rightarrow (v_1, \dots, v_n)$ base.

$\boxed{c \Rightarrow a}$ Per definizione di base. □

Oss. Per stabilire se certi vettori formano una base per V basta che il loro numero sia uguale a $\dim V$ e che siano generatori oppure che siano linearmente indipendenti.

Oss. Se $\dim V = 1$ allora un vettore $v \in V$ è base per $V \Leftrightarrow v \neq 0_V$.

Esempio. Due vettori di $\mathbb{R}^2$ formano una base $\Leftrightarrow$ non sono proporzionali.

Rango di una matrice

Spazio delle colonne e spazio delle righe. $A \in M_{m,n}(\mathbb{K}) \rightsquigarrow$

$\mathrm{CS}(A) \stackrel{\text{def}}{=} \mathrm{span}(A_{(1)}, \dots, A_{(n)}) \subset \mathbb{K}^m$ *Spazio delle colonne di A. (Column Space)*

$\mathrm{RS}(A) \stackrel{\text{def}}{=} \mathrm{span}(A^{(1)}, \dots, A^{(m)}) \subset \mathbb{K}^n$ *Spazio delle righe di A. (Row Space)*

Enunciamo la seguente proposizione senza dimostrarla.

Prop. $\dim(\mathrm{CS}(A)) = \dim(\mathrm{RS}(A))$.

Def. Sia $A \in M_{m,n}(\mathbb{K})$. Si chiama *rango* di A la dimensione dello spazio delle colonne di A, ossia

$$\mathrm{rg}\, A \stackrel{\text{def}}{=} \dim(\mathrm{CS}(A)).$$

Oss. In virtù della Proposizione si ha $\mathrm{rg}(A) = \dim(\mathrm{RS}(A)) \leqslant \min(m, n)$. Il rango si può anche definire come il *massimo numero* di righe (o di colonne) linearmente indipendenti.

Le operazioni elementari sulle righe (o sulle colonne) di una matrice preservano il rango.

Il rango di una matrice a gradini è il numero di pivot. Pertanto per calcolare $\mathrm{rg}(A)$ si trasforma A in una matrice a gradini A' mediante il metodo di Gauss e si contano i pivot

$$\mathrm{rg}(A) = \text{numero dei pivot di } A'.$$

Teorema di Rouché-Capelli. *Consideriamo un sistema lineare*

$$S : AX = B$$

con $A \in M_{m,n}(\mathbb{K})$ *e* $B \in \mathbb{K}^m$*. Allora* S *è compatibile se e solo se*

$$\mathrm{rg}(A) = \mathrm{rg}(A\,|B).$$

Se questo accade allora la soluzione generale si rappresenta con $n - \mathrm{rg}(A)$ *parametri liberi, ed è unica se e solo se* $\mathrm{rg}(A) = n$*.*

Dim. Con l'algoritmo di Gauss trasformiamo $\tilde{A} = (A\,|B)$ in una matrice a gradini $\tilde{A}' = (A'\,|B')$ che rappresenta un sistema lineare $S' : A'X = B'$ equivalente ad S. Allora S compatibile $\Leftrightarrow$ non risultano equazioni del tipo $0 = b \neq 0 \Leftrightarrow$ i pivot di $\tilde{A}'$ sono tutti sulle prime n colonne $\Leftrightarrow$ $\mathrm{rg}(\tilde{A}') = \mathrm{rg}(A')$. D'altra parte $\mathrm{rg}(\tilde{A}) = \mathrm{rg}(\tilde{A}')$ e $\mathrm{rg}(A) = \mathrm{rg}(A')$.

Se S è compatibile la soluzione generale ha $d = n - \mathrm{rg}(A)$ parametri liberi perché $\mathrm{rg}(A) =$ numero di pivot di A', ed è unica se $d = 0$. □

Oss. Il numero $d = n - \mathrm{rg}(A)$ di parametri liberi è una proprietà dello spazio delle soluzioni.

Calcolo del rango col determinante

Def. Un *minore* di ordine r di una matrice $A \in M_{m,n}(\mathbb{K})$ è il determinante di una sottomatrice quadrata di A di ordine r.

Teor (di Kronecker)**.** $\mathrm{rg}\, A = r \Leftrightarrow$ *A ha un minore di ordine r non nullo e ogni minore di ordine $r+1$ ottenuto aggiungendo ad esso una riga e una colonna è nullo.*

Esempio.

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

quindi $\mathrm{rg}\, A = 2$.

Calcoli con le basi

Data una base $\mathcal{B} = (v_1, \dots, v_n)$ per V e dati $u_1 \dots, u_k \in V$ possiamo determinare una base di

$$U = \operatorname{span}(u_1, \dots, u_k) \subset V$$

considerando la matrice $A \in M_{n,k}(\mathbb{K})$ avente per colonne le coordinate dei vettori $u_1, \dots, u_k$ rispetto a $\mathcal{B}$. Mediante l'algoritmo di Gauss trasformiamo A in A' a gradini. Si ha:

$$\dim U = \operatorname{rg}(A) = \text{numero dei pivot di } A'.$$

Una base di U è data da quei generatori $u_1, \dots, u_k$ corrispondenti alle colonne dei pivot.

Esempio. Determinare base e dimensione di $U = \operatorname{span}(u_1, u_2, u_3, u_4) \subset \mathbb{R}^4$

$$u_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}, \ u_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 2 \end{pmatrix}, \ u_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 3 \end{pmatrix}, \ u_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 2 & 0 & 2 & 1 \\ -1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 3 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

3 pivot nelle colonne $1, 2, 4 \Rightarrow \dim U = 3$ e (u_1, u_2, u_4) base per U.

Equazioni di sottospazi vettoriali

Equazione vettoriale. Dato un sottospazio vettoriale $U \subset V$, scegliamo una base $\mathcal{C} = (u_1, \dots, u_k)$ per U. Ogni vettore $v \in U$ è combinazione lineare unica di $u_1, \dots, u_k$ e possiamo scrivere

$$U: \ v = t_1 u_1 + \cdots + t_k u_k.$$

Questa si chiama *equazione vettoriale* di U. I coefficienti $t_1, \dots, t_k \in \mathbb{K}$ prendono il nome di *parametri* e variano in tutti i modi possibili.

Equazioni parametriche. Scegliamo una base $\mathcal{B} = (b_1, \dots, b_n)$ per V.

$$v = \sum_{i=1}^{n} x_i b_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^{\mathcal{B}}, \quad u_j = \sum_{i=1}^{n} a_{ij} b_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}^{\mathcal{B}}, \quad j = 1, \dots, k.$$

Poniamo anche

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \in \mathbb{K}^n.$$

L'equazione vettoriale di U si scrive come

$$U: \ X = t_1 A_1 + \cdots + t_k A_k$$

da cui si ottengono le *equazioni parametriche* di U

$$U\colon \begin{cases} x_1 = a_{11}t_1 + \cdots + a_{1k}t_k \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = a_{n1}t_1 + \cdots + a_{nk}t_k \end{cases}$$

Equazioni cartesiane. Nelle equazioni parametriche possiamo eliminare i parametri $t_1, \ldots, t_k$. Otteniamo le *equazioni cartesiane* di U

$$U\colon \begin{cases} c_{11}x_1 + \cdots + c_{1n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ c_{m1}x_1 + \cdots + c_{mn}x_n = 0 \end{cases}$$

sistema omogeneo il cui spazio delle soluzioni rappresenta U: $v \in U \Leftrightarrow$ il vettore delle coordinate di v è soluzione del sistema.

Il passaggio da equazioni cartesiane a equazioni parametriche si fa risolvendo il sistema e determinando la soluzione generale, che esprime le equazioni parametriche.

Esempio. $U = \operatorname{span}(u_1, u_2, u_3) \subset \mathbb{R}^4$, da un esempio della lezione scorsa:

$$u_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix},\ u_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 2 \end{pmatrix},\ u_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$U\colon\ X = t_1u_1 + t_2u_2 + t_3u_3 \qquad \text{equazione vettoriale}$$

$$U\colon \begin{cases} x_1 = t_1 - t_2 \\ x_2 = 2t_1 + t_3 \\ x_3 = -t_1 + 2t_2 + t_3 \\ x_4 = t_1 + 2t_2 \end{cases} \qquad \text{equazioni parametriche}$$

Per trovare le equazioni cartesiane risolviamo il sistema col metodo di Gauss, trattando i parametri t_i come incognite e le x_i come termini noti. Ricaveremo le equazioni cartesiane come condizioni di compatibilità.

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & x_1 \\ 2 & 0 & 1 & x_2 \\ -1 & 2 & 1 & x_3 \\ 1 & 2 & 0 & x_4 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & x_1 \\ 0 & 2 & 1 & -2x_1 + x_2 \\ 0 & 1 & 1 & x_1 + x_3 \\ 0 & 3 & 0 & -x_1 + x_4 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & x_1 \\ 0 & 1 & 1 & x_1 + x_3 \\ 0 & 2 & 1 & -2x_1 + x_2 \\ 0 & 3 & 0 & -x_1 + x_4 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & x_1 \\ 0 & 1 & 1 & x_1 + x_3 \\ 0 & 0 & -1 & -4x_1 + x_2 - 2x_3 \\ 0 & 0 & -3 & -4x_1 - 3x_3 + x_4 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & x_1 \\ 0 & 1 & 1 & x_1 + x_3 \\ 0 & 0 & -1 & -4x_1 + x_2 - 2x_3 \\ 0 & 0 & 0 & 8x_1 - 3x_2 + 3x_3 + x_4 \end{array}\right)$$

da cui si ricava l'equazione cartesiana

$$U\colon\ 8x_1 - 3x_2 + 3x_3 + x_4 = 0.$$

Esempio. $W \subset \mathbb{R}^4$ sottospazio vettoriale di equazioni cartesiane

$$W\colon \begin{cases} x_1 - 2x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

Determiniamo equazioni parametriche, una base e la dimensione di W.

$$\begin{pmatrix} 1 & 0 & -2 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3 & -1 \end{pmatrix} \quad \begin{cases} x_1 \quad - 2x_3 + x_4 = 0 \\ x_2 + 3x_3 - x_4 = 0 \end{cases}$$

$$W: \begin{cases} x_1 = \quad 2t_1 - t_2 \\ x_2 = -3t_1 + t_2 \\ x_3 = \quad t_1 \\ x_4 = \quad\quad t_2 \end{cases} \qquad w_1 = \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \; w_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$\mathcal{B}_W = (w_1, w_2)$ base per W, ricavabile dai coefficienti dei parametri o anche assegnando ai parametri (t_1, t_2) i valori $(1, 0)$ e $(0, 1)$ risp.
La dimensione coincide dunque col numero dei parametri liberi: $\dim W = 2$.

Completamento della base

Sia V un $\mathbb{K}$-spazio vettoriale, $\dim V = n$, $\mathcal{B} = (b_1 \ldots, b_n)$ base per V.

Oss. Il vettore coordinate di b_i rispetto alla base $\mathcal{B}$ è e_i, i-esimo vettore della base canonica di $\mathbb{K}^n$:

$$b_i = 0b_1 + \cdots + b_i + \cdots + 0b_n = \begin{pmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}^{\mathcal{B}} \leftarrow i\text{-esima posizione}$$

Teorema di completamento della base. *Dati $u_1 \ldots, u_k \in V$ linearmente indipendenti $\Rightarrow \exists\, u_{k+1}, \ldots, u_n \in \mathcal{B}$ t.c. $(u_1, \ldots, u_n)$ base per V.*

Dim. $A \in M_{n,k+n}(\mathbb{K})$ matrice avente per colonne le coordinate dei vettori $u_1, \ldots, u_k, b_1, \ldots, b_n \Rightarrow \operatorname{rg} A = n$. $A \overset{\text{Gauss}}{\rightsquigarrow} A'$ a gradini con n pivot sulle prime k colonne (linearmente indipendenti) e su altre $n - k$ colonne corrispondenti a vettori della base $\mathcal{B}$. Le colonne con i pivot sono base. □

Esempio.

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \; u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3 \qquad \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 \end{pmatrix} \rightsquigarrow$$

(u_1, u_2, e_2) base $\qquad \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$

Applicazioni lineari

Def. Un'applicazione $f : V \to W$ tra $\mathbb{K}$-spazi vettoriali è *lineare* se valgono le seguenti, $\forall\, u, v \in V$, $\forall\, \alpha \in \mathbb{K}$

1) $f(u+v) = f(u) + f(v)$.

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

Oss. $f : V \to W$ lineare $\Rightarrow f(0_V) = 0_W$. $f(0_V) = f(0\, 0_V) = 0 f(0_V) = 0_W$.

Esempio. $\mathrm{id}_V : V \to V$, $\mathrm{id}_V(v) = v$, $\forall\, v \in V$ *applicazione identica* lineare.

Esempio. $0 : V \to W$, $0(v) = 0_W$, $\forall\, v \in V$ *applicazione nulla* lineare.

Esempio. $\alpha \in \mathbb{K} - \{0\} \rightsquigarrow h_\alpha : V \to V$, $h_\alpha(v) = \alpha v$ *omotetia* lineare.

Applicazione lineare associata ad una matrice.

$$A \in M_{m,n}(\mathbb{K}) \rightsquigarrow L_A : \mathbb{K}^n \to \mathbb{K}^m$$
$$L_A(X) = AX.$$

$L_A(X+Y) = A(X+Y) = AX + AY = L_A(X) + L_A(Y)$, $\forall\, X, Y \in \mathbb{K}^n$
$L_A(\alpha X) = A(\alpha X) = \alpha AX = \alpha L_A(X)$, $\forall\, X \in \mathbb{K}^n$, $\forall\, \alpha \in \mathbb{K} \Rightarrow L_A$ lineare.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \rightsquigarrow L_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

Def. Un polinomio omogeneo di primo grado in $x_1, \ldots, x_n$ è un'espressione del tipo $a_1x_1 + \cdots + a_nx_n$, con $a_1, \ldots, a_n \in \mathbb{K}$ non tutti nulli.

Oss. $L_A : \mathbb{K}^n \to \mathbb{K}^m$ ha per componenti polinomi omogenei di primo grado oppure nulli. Le applicazioni di questo tipo sono lineari.

Esempio.

$$A = \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix} \in M_2(\mathbb{R}) \rightsquigarrow L_A : \mathbb{R}^2 \to \mathbb{R}^2$$
$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ -3x + 5y \end{pmatrix}$$

Prop. *$f : V \to W$ lineare* $\Leftrightarrow$

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v), \quad \forall\, u, v \in V,\ \forall\, \alpha, \beta \in \mathbb{K}.$$

Dim. $\boxed{\Rightarrow}$ $f(\alpha u + \beta v) = f(\alpha u) + f(\beta v) = \alpha f(u) + \beta f(v)$.

$\boxed{\Leftarrow}$ Si ottiene *1)* per $\alpha = \beta = 1$ e *2)* per $\beta = 0$. □

Oss. $f : V \to W$ lineare $\Leftrightarrow \forall\, n \in N$, $\forall\, \alpha_1, \ldots, \alpha_n \in \mathbb{K}$, $\forall\, v_1, \ldots, v_n \in V$,

$$f(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 f(v_1) + \cdots + \alpha_n f(v_n), \quad \text{ossia}$$
$$f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f(v_i).$$

Esempio. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y$ non è lineare, infatti
$4 = f(2, 0) = f(2(1, 0)) \neq 2f(1, 0) = 2$.

Esempio. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x+y+1$ non è lineare, infatti $f(0, 0) = 1$.

Nucleo e immagine

Nucleo. $f : V \to W$ lineare $\rightsquigarrow$

$$\ker f \stackrel{\text{def}}{=} \{v \in V \mid f(v) = 0_W\} \subset V$$

Def. $\ker f$ si chiama *nucleo* di f (dall'inglese *kernel*).

Oss. Dato $v \in V$, si ha che $v \in \ker f \Leftrightarrow f(v) = 0_W$.

Oss. $A \in M_{m,n}(\mathbb{K}) \rightsquigarrow \ker L_A = \Sigma_S$ con $S : AX = 0_{\mathbb{K}^m}$.

Teor. *$f : V \to W$ lineare $\Rightarrow$ $\ker f$ sottospazio vettoriale di V.*

Dim. $f(0_V) = 0_W \Rightarrow 0_V \in \ker f \neq \emptyset$.

$\forall\, \alpha, \beta \in \mathbb{K}$, $\forall\, u, v \in \ker f \Rightarrow f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) = 0_W \Rightarrow \alpha u + \beta v \in \ker f$. □

Def. $f : V \to W$ è *iniettiva* se $\forall\, u, v \in V$, $f(u) = f(v) \Rightarrow u = v$.

Teor. *$f : V \to W$ lineare. Allora f iniettiva $\Leftrightarrow \ker f = \{0_V\}$.*

Dim. $\boxed{\Rightarrow}$ $\forall\, v \in \ker f \Rightarrow f(v) = f(0_V) = 0_W \Rightarrow v = 0_V \Rightarrow \ker f = \{0_V\}$.

$\boxed{\Leftarrow}$ $\forall\, u, v \in V$, $f(u) = f(v) \Rightarrow f(u) - f(v) = 0_W \Rightarrow f(u - v) = 0_W \Rightarrow u - v \in \ker f = \{0_V\} \Rightarrow u - v = 0_V \Rightarrow u = v$. □

Cor. *$f : V \to W$ lineare. Allora f iniettiva $\Leftrightarrow \dim(\ker f) = 0$.*

Nucleo di L_A. Data $A \in M_{m,n}(\mathbb{K}) \rightsquigarrow L_A : \mathbb{K}^n \to \mathbb{K}^m$, $L_A(X) = AX$.
$X \in \ker L_A \Leftrightarrow L_A(X) = 0_{\mathbb{K}^m} \Leftrightarrow AX = 0_{\mathbb{K}^m} \Leftrightarrow X$ è soluzione del sistema

$$S_A : AX = 0_{\mathbb{K}^m}$$

Prop. *$\ker L_A$ è lo spazio delle soluzioni di $S_A : AX = 0_{\mathbb{K}^m}$.*

Immagine. $f : V \to W$ lineare $\rightsquigarrow$

$$\operatorname{im} f \stackrel{\text{def}}{=} \{f(v) \mid v \in V\} \subset W$$

Def. $\operatorname{im} f$ si chiama *immagine* di f.

Oss. Dato $w \in W$, si ha che $w \in \operatorname{im} f \Leftrightarrow \exists\, v \in V$ t.c. $f(v) = w$.

Teor. *$f : V \to W$ lineare $\Rightarrow$ $\operatorname{im} f$ sottospazio vettoriale di W.*

Dim. $f(0_V) = 0_W \Rightarrow 0_W \in \operatorname{im} f \neq \emptyset$.

$\forall\, \alpha, \beta \in \mathbb{K}$, $\forall\, w_1, w_2 \in \operatorname{im} f \Rightarrow \exists\, v_1, v_2 \in V$ t.c. $f(v_1) = w_1$ e $f(v_2) = w_2$ $\Rightarrow \alpha w_1 + \beta w_2 = \alpha f(v_1) + \beta f(v_2) = f(\alpha v_1 + \beta v_2) \in \operatorname{im} f$. □

Def. $f : V \to W$ è *suriettiva* se $\operatorname{im} f = W$.

Oss. f suriettiva $\Leftrightarrow \forall\, w \in W$, $\exists\, v \in V$ t.c. $f(v) = w$.

Def. $\operatorname{rg} f \stackrel{\text{def}}{=} \dim(\operatorname{im} f)$ si chiama *rango* di f.

Oss. $f : V \to W$ lineare, $\dim W < \infty$. Allora f suriettiva $\Leftrightarrow \operatorname{rg} f = \dim W$.

Generatori per l'immagine

Prop. $V = \operatorname{span}(v_1, \dots, v_n) \Rightarrow \operatorname{im} f = \operatorname{span}(f(v_1), \dots, f(v_n))$.

Dim. $f(v_1), \dots, f(v_n) \in \operatorname{im} f \Rightarrow \operatorname{span}(f(v_1), \dots, f(v_n)) \subset \operatorname{im} f$.

$\forall\, w \in \operatorname{im} f$, $\exists\ v = \sum_{i=1}^{n} \alpha_i v_i \in V$ t.c.

$$w = f(v) = f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f(v_i) \in \operatorname{span}(f(v_1), \dots, f(v_n))$$

$\Rightarrow \operatorname{im} f \subset \operatorname{span}(f(v_1), \dots, f(v_n))$ e quindi si ha la tesi. □

Cor. *Per ogni matrice* $A \in M_{m,n}(\mathbb{K})$ *si ha* $\operatorname{im} L_A = \operatorname{span}\big(A_{(1)}, \dots, A_{(n)}\big)$.

Dim. $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$, $L_A(X) = AX$. $(e_1, \dots, e_n)$ base canonica di $\mathbb{K}^n \Rightarrow$

$$\operatorname{im} L_A = \operatorname{span}(L_A(e_1), \dots, L_A(e_n)).$$

$$L_A(e_1) = A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A_{(1)}$$

$$\vdots$$

$$L_A(e_n) = A \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = A_{(n)}$$ □

Cor. $\operatorname{rg} L_A = \operatorname{rg} A$.

Dim. $\operatorname{rg} L_A = \dim(\operatorname{im} L_A) = \dim\big(\operatorname{span}\big(A_{(1)}, \dots, A_{(n)}\big)\big) = \operatorname{rg} A$. □

Cor. $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$ *suriettiva* $\Leftrightarrow \operatorname{rg} A = m$.

Teorema della dimensione

Teor. *$f : V \to W$ lineare e* $\dim V < \infty \Rightarrow \dim V = \dim(\ker f) + \operatorname{rg} f$.

Dim. $\dim V = n$, $\dim(\ker f) = k \leqslant n$.
$(u_1, \ldots, u_k)$ base per $\ker f \rightsquigarrow (u_1, \ldots, u_k, v_1, \ldots, v_{n-k})$ base per V.
completamento della base
$f(u_i) = 0_W$, $w_i := f(v_i) \Rightarrow w_1, \ldots, w_{n-k}$ generatori di $\operatorname{im} f$.
Mostriamo che $w_1, \ldots, w_{n-k}$ sono linearmente indipendenti.

$$\sum_{i=1}^{n-k} \alpha_i w_i = 0_W \Rightarrow 0_W = \sum_{i=1}^{n-k} \alpha_i f(v_i) = f\left(\sum_{i=1}^{n-k} \alpha_i v_i\right) \Rightarrow$$

$$\sum_{i=1}^{n-k} \alpha_i v_i \in \ker f \Rightarrow \sum_{i=1}^{n-k} \alpha_i v_i = \sum_{i=1}^{k} \beta_i u_i \Rightarrow$$

$$\sum_{i=1}^{n-k} \alpha_i v_i - \sum_{i=1}^{k} \beta_i u_i = 0_V \Rightarrow \alpha_i = 0,\ \beta_i = 0,\ \forall\, i.$$

$(w_1, \ldots, w_{n-k})$ base per $\operatorname{im} f \Rightarrow \operatorname{rg} f = \dim(\operatorname{im} f) = n - k \Rightarrow$

$$\dim(\ker f) + \operatorname{rg} f = k + n - k = \dim V. \qquad \square$$

Isomorfismi di spazi vettoriali

Consideriamo $\mathbb{K}$-spazi vettoriali V e W.

Def. $f : V \to W$ *biiettiva* (o *biunivoca*) $\overset{\text{def}}{\iff}$ f iniettiva e suriettiva.

Oss. $f : V \to W$ biiettiva $\Leftrightarrow \forall\, w \in W$, $\exists!\, v \in V$ t.c. $f(v) = w$
$\Leftrightarrow \exists\, f^{-1} : W \to V$ *applicazione inversa:* $f^{-1} \circ f = \mathrm{id}_V$ e $f \circ f^{-1} = \mathrm{id}_W$.

Oss. $f^{-1}(f(v)) = (f^{-1} \circ f)(v) = \mathrm{id}_V(v) = v$ e similmente $f(f^{-1}(w)) = w$.

Def. $f : V \to W$ *isomorfismo* $\overset{\text{def}}{\iff}$ f lineare e biiettiva.

Oss. $\mathrm{id}_V : V \to V$ isomorfismo.

Prop. *$f : V \to W$ isomorfismo $\Rightarrow$ $f^{-1} : W \to V$ lineare.*

Dim. $\forall\, \alpha, \beta \in \mathbb{K}$, $\forall\, w_1, w_2 \in W \rightsquigarrow v_1 := f^{-1}(w_1)$, $v_2 := f^{-1}(w_2) \Rightarrow$ $w_1 = f(v_1)$, $w_2 = f(v_2)$.

$$\begin{aligned} f^{-1}(\alpha w_1 + \beta w_2) &= f^{-1}(\alpha f(v_1) + \beta f(v_2)) \\ &= f^{-1}(f(\alpha v_1 + \beta v_2)) \qquad (f \text{ lineare}) \\ &= \alpha v_1 + \beta v_2 \\ &= \alpha f^{-1}(w_1) + \beta f^{-1}(w_2). \end{aligned}$$ □

Def. V e W sono *isomorfi* $\overset{\text{def}}{\iff}$ $\exists\, f : V \to W$ isomorfismo.
In questo caso scriviamo $V \cong W$ (si legge V isomorfo a W).

Consideriamo $\mathbb{K}$-spazi vettoriali U, V, W.

Prop. *$f : U \to V$ e $g : V \to W$ lineari $\Rightarrow$ $g \circ f : U \to W$ lineare.*

Dim. $\forall\, \alpha, \beta \in \mathbb{K}$, $\forall\, u_1, u_2 \in U \Rightarrow (g \circ f)(\alpha u_1 + \beta u_2) = g(f(\alpha u_1 + \beta u_2)) =$
$g(\alpha f(u_1) + \beta f(u_2)) = \alpha g(f(u_1)) + \beta g(f(u_2)) =$
$\alpha(g \circ f)(u_1) + \beta(g \circ f)(u_2)$. □

Oss. In altre parole, composizioni di applicazioni lineari sono lineari.

Cor. *$f : U \to V$ e $g : V \to W$ isomorfismi $\Rightarrow$ $g \circ f : U \to W$ isomorfismo.*

Dim. $g \circ f$ è lineare e invertibile con inversa $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

Oss. In altre parole, composizioni di isomorfismi sono isomorfismi.

Teor. *Siano $f : V \to W$ lineare e $\mathcal{B} = (b_1, \ldots, b_n)$ base per V. Allora*

1) *f suriettiva $\Leftrightarrow$ $f(b_1), \ldots, f(b_n)$ generano W.*
2) *f iniettiva $\Leftrightarrow$ $f(b_1), \ldots, f(b_n)$ linearmente indipendenti.*
3) *f isomorfismo $\Leftrightarrow$ $(f(b_1), \ldots, f(b_n))$ base per W.*

Dim. 1) Segue subito dal fatto che $\operatorname{im} f = \operatorname{span}(f(b_1), \ldots, f(b_n))$.
2) f iniettiva $\Leftrightarrow \dim(\ker f) = 0 \Leftrightarrow \dim(\operatorname{im} f) = \operatorname{rg} f \overset{\text{Teor dim}}{=} \dim V = n$.
Segue la tesi dal fatto che $\dim(\operatorname{im} f) = n \Leftrightarrow f(b_1), \ldots, f(b_n)$ lin. indip.
3) Segue subito dalle precedenti. □

Cor (Invarianza della dimensione)**.** $V \cong W \Rightarrow \dim V = \dim W$.

Cor. $\mathbb{K}^m \cong \mathbb{K}^n \Rightarrow m = n$.

Teor. *Sia* $f : V \to W$ *lineare e* $\dim V = \dim W < \infty$*. Sono equivalenti:*

1) f *suriettiva;*
2) f *iniettiva;*
3) f *isomorfismo;*

Dim. f suriettiva $\Leftrightarrow \operatorname{rg} f = \dim W \Leftrightarrow \dim(\ker f) = \dim V - \operatorname{rg} f = 0 \Leftrightarrow$ f iniettiva. Pertanto $(1) \Leftrightarrow (2) \Rightarrow (3) \Rightarrow (1)$. $\overset{\text{Teor dim}}{}$ □

Esempio.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \rightsquigarrow L_A : \mathbb{R}^2 \to \mathbb{R}^2$$

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x + y \end{pmatrix}$$

$\operatorname{rg} L_A = \operatorname{rg} A = 2 \Rightarrow L_A$ suriettiva e quindi isomorfismo.

Applicazioni lineari associate a matrici

La seguente proposizione è la ragione principale alla base della definizione del prodotto righe per colonne: il prodotto di matrici corrisponde alla composizione di applicazioni.

Prop. *Date* $A \in M_{m,n}(\mathbb{K})$, $B \in M_{n,\ell}(\mathbb{K}) \Rightarrow L_{AB} = L_A \circ L_B$

Dim. $AB \in M_{m,\ell}(\mathbb{K}) \Rightarrow L_{AB} : \mathbb{K}^\ell \to \mathbb{K}^m$.
Abbiamo anche $L_B : \mathbb{K}^\ell \to \mathbb{K}^n$ e $L_A : \mathbb{K}^n \to \mathbb{K}^m \Rightarrow L_A \circ L_B : \mathbb{K}^\ell \to \mathbb{K}^m$.
$L_{AB}(X) = ABX = L_A(L_B(X)) = (L_A \circ L_B)(X)$, $\forall\, X \in \mathbb{K}^\ell$. □

Oss. $L_{I_n} = \mathrm{id}_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$.

Oss. $A \in \mathrm{GL}_n(\mathbb{K}) \Rightarrow L_{A^{-1}} = (L_A)^{-1}$.
Infatti $L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_{I_n} = \mathrm{id}_{\mathbb{K}^n}$ e similmente $L_{A^{-1}} \circ L_A = \mathrm{id}_{\mathbb{K}^n}$.

Teor. $A \in M_n(\mathbb{K})$ *invertibile* $\Leftrightarrow \operatorname{rg} A = n$.

Dim. $\boxed{\Rightarrow}$ $A \in \mathrm{GL}_n(\mathbb{K}) \Rightarrow L_A : \mathbb{K}^n \xrightarrow{\cong} \mathbb{K}^n$ iso $\Rightarrow \operatorname{rg} A = \operatorname{rg} L_A = n$.

$\boxed{\Leftarrow}$ $\mathcal{E}_n = (e_1, \dots, e_n)$ base canonica di $\mathbb{K}^n$. $\operatorname{rg} A = n \Rightarrow \operatorname{rg}(A \mid e_j) = n \Rightarrow$ $AX = e_j$ sistema compatibile con un'unica soluzione $B_j \in \mathbb{K}^n$ (Teorema di Rouché-Capelli) $\Rightarrow B = (B_1 \ \cdots \ B_n) \in M_n(\mathbb{K})$ inversa di A. □

Teorema di determinazione

Teor. *Dati V e W $\mathbb{K}$-spazi vettoriali, $\mathcal{B} = (b_1, \dots, b_n)$ base per V e $w_1, \dots, w_n \in W$ vettori arbitrari $\Rightarrow \exists!$ applicazione lineare*

$$f : V \to W \quad t.c.$$

$$\begin{cases} f(b_1) = w_1 \\ \cdots\cdots\cdots \\ f(b_n) = w_n \end{cases}$$

Dim. *Unicità* Sia $f : V \to W$ lineare t.c. $f(b_i) = w_i$, $\forall\, i = 1, \dots, n$.

$$\forall\, v \in V \rightsquigarrow v = \sum_{i=1}^{n} \alpha_i b_i \Rightarrow f(v) = f\left(\sum_{i=1}^{n} \alpha_i b_i\right) = \sum_{i=1}^{n} \alpha_i f(b_i) = \sum_{i=1}^{n} \alpha_i w_i.$$

Le coordinate $(\alpha_1, \dots, \alpha_n)$ di v rispetto alla base $\mathcal{B}$ sono uniche $\Rightarrow f(v)$ ha l'unico possibile valore dato dalla formula

$$f(v) = \sum_{i=1}^{n} \alpha_i w_i$$

e quindi f, se esiste, è unica.

Esistenza L'ultima formula suggerisce la definizione di f:

$$f : V \to W$$

$$f(v) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \alpha_i w_i, \quad \text{con } v = \sum_{i=1}^{n} \alpha_i b_i.$$

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^{\mathcal{B}}, \dots, b_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}^{\mathcal{B}} \Rightarrow \begin{cases} f(b_1) = w_1 \\ \cdots\cdots\cdots \\ f(b_n) = w_n. \end{cases}$$

Resta da dimostrare che f è lineare.

$$v = \sum_{i=1}^{n} \alpha_i b_i \text{ e } u = \sum_{i=1}^{n} \beta_i b_i \Rightarrow v + u = \sum_{i=1}^{n} (\alpha_i + \beta_i) b_i \Rightarrow$$

$$f(v + u) = \sum_{i=1}^{n} (\alpha_i + \beta_i) w_i = \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \beta_i w_i = f(v) + f(u).$$

$$\beta \in \mathbb{K} \Rightarrow \beta v = \sum_{i=1}^{n} \beta \alpha_i b_i \Rightarrow f(\beta v) = \sum_{i=1}^{n} \beta \alpha_i w_i = \beta \sum_{i=1}^{n} \alpha_i w_i = \beta f(v). \quad \square$$

Cor. *Siano V e W di dimensione finita. Allora $V \cong W \Leftrightarrow \dim V = \dim W$.*

Dim. $\Rightarrow$ Corollario della Lezione 20 (Invarianza della dimensione).

$\boxed{\Leftarrow}$ $\dim V = \dim W = n$, $\mathcal{B} = (b_1, \dots, b_n)$ base per V e $\mathcal{C} = (c_1, \dots, c_n)$ base per $W \rightsquigarrow f : V \to W$ lineare t.c. $f(b_i) = c_i$, $\forall\, i \Rightarrow f$ isomorfismo perché f manda una base in una base. □

Cor. $V \cong \mathbb{K}^n \Leftrightarrow \dim V = n$.

Oss. $\mathcal{B} = (b_1, \dots, b_n)$ base per V, $\mathcal{E}_n = (e_1, \dots, e_n)$ base canonica per $\mathbb{K}^n$

$$\Phi_{\mathcal{B}} : V \to \mathbb{K}^n \quad \text{isomorfismo t.c.}$$

$$\Phi_{\mathcal{B}}(b_i) = e_i, \quad \forall\, i = 1, \dots, n.$$

$$v = \sum_{i=1}^{n} \alpha_i b_i \;\Rightarrow\; \Phi_{\mathcal{B}}(v) = \sum_{i=1}^{n} \alpha_i e_i = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Quindi $\Phi_{\mathcal{B}}$ associa a $v \in V$ le sue coordinate rispetto alla base $\mathcal{B}$.

Oss. Viceversa, dato un isomorfismo $\Phi : V \to \mathbb{K}^n$, con $\dim V = n$, resta determinata una base per V data dai vettori $b_i = \Phi^{-1}(e_i)$, $i = 1, \dots, n$, e abbiamo $\Phi_{\mathcal{B}} = \Phi$. Scegliere una base per V equivale sostanzialmente a scegliere un isomorfismo $\Phi : V \to \mathbb{K}^n$.

Matrici di applicazioni lineari

V e W $\mathbb{K}$-spazi vettoriali con $\dim V = n$ e $\dim W = m$.
$f : V \to W$ applicazione lineare. Scegliamo basi
$\mathcal{B} = (b_1, \dots, b_n)$ per V

$$\mathcal{C} = (c_1, \dots, c_m) \text{ per } W \rightsquigarrow f(b_j) = \sum_{i=1}^{m} a_{ij} c_i \rightsquigarrow A = (a_{ij}) \in M_{m,n}(\mathbb{K})$$

A è costruita per colonne: $A_{(j)}$ è la colonna delle coordinate di $f(b_j)$.

$$v \in V \rightsquigarrow v = \sum_{j=1}^{n} x_j b_j = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^{\mathcal{B}}$$

$$\begin{aligned} f(v) &= \sum_{j=1}^{n} x_j f(b_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} c_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) c_i \\ &= \sum_{i=1}^{m} y_i c_i = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}^{\mathcal{C}}, \quad \text{con } y_i = \sum_{j=1}^{n} a_{ij} x_j. \end{aligned}$$

Poniamo ora

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \; Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

f si rappresenta in coordinate con l'equazione

$$f: Y = AX$$

quindi A permette di calcolare f in coordinate. Scriviamo anche

$$f(X) = AX.$$

Def. A si chiama *matrice di f rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$*. Si indica con $A = M_{\mathcal{B}}^{\mathcal{C}}(f) \in M_{m,n}(\mathbb{K})$.

Viceversa, fissate le basi e data $A \in M_{m,n}(\mathbb{K})$, ponendo $f(X) = AX$ resta definita un'applicazione lineare $f: V \to W$. Abbiamo quindi il teorema.

Teor. *Scelte le basi $\mathcal{B}$ e $\mathcal{C}$ risp. per V e W, c'è una biiezione tra applicazioni lineari $f: V \to W$ e matrici $m \times n$, che associa a f la matrice $M_{\mathcal{B}}^{\mathcal{C}}(f)$.*

N.B. Questo ha senso solo se sono fissate basi per dominio e codominio. Senza basi non si associa niente.

N.B. Se $f: V \to V$ è un endomorfismo (lineare con stesso dominio e codominio), spesso useremo la stessa base $\mathcal{B}$ sia nel dominio che nel codominio $\rightsquigarrow M_{\mathcal{B}}^{\mathcal{B}}(f) \in M_n(\mathbb{K})$.

Oss. $M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V) = I_n$ perché $\mathrm{id}_V(b_j) = b_j$, $j = 1, \ldots, n$.

Applicazioni lineari $\mathbb{K}^n \to \mathbb{K}^m$. Usiamo le basi canoniche su $\mathbb{K}^n$ e $\mathbb{K}^m$
$\forall\, f: \mathbb{K}^n \to \mathbb{K}^m$ lineare $\rightsquigarrow A = M_{\mathcal{E}_n}^{\mathcal{E}_m}(f) \in M_{m,n}(\mathbb{K}) \Rightarrow$

$$f(X) = AX = L_A(X).$$

Quindi $f = L_A$. Abbiamo dimostrato il teorema seguente.

Teor. *Le applicazioni lineari $\mathbb{K}^n \to \mathbb{K}^m$ sono del tipo L_A con $A \in M_{m,n}(\mathbb{K})$.*

Oss. $\forall\, A \in M_{m,n}(\mathbb{K}) \Rightarrow M_{\mathcal{E}_n}^{\mathcal{E}_m}(L_A) = A$ perché $L_A(e_j) = Ae_j = A_{(j)}$.

Matrice dell'applicazione composta

V, W, U spazi vettoriali su $\mathbb{K}$, $\dim V = n$, $\dim W = m$, $\dim U = \ell$
$\mathcal{B} = (b_1, \dots, b_n)$ base per V con coordinate $X = (x_1, \dots, x_n)$
$\mathcal{C} = (c_1, \dots, c_m)$ base per W con coordinate $Y = (y_1, \dots, y_m)$
$\mathcal{D} = (d_1, \dots, d_\ell)$ base per U con coordinate $Z = (z_1, \dots, z_\ell)$

Applicazioni lineari

$$f : V \to W \qquad g : W \to U \qquad g \circ f : V \to U$$
$$A = M_{\mathcal{B}}^{\mathcal{C}}(f) \qquad B = M_{\mathcal{C}}^{\mathcal{D}}(g) \qquad C = M_{\mathcal{B}}^{\mathcal{D}}(g \circ f)$$
$$m \times n \qquad \ell \times m \qquad \ell \times n$$

$$V \xrightarrow{f} W \xrightarrow{g} U \qquad (g \circ f)$$

$$Y \overset{g}{\longmapsto} BY$$
$$X \overset{f}{\longmapsto} AX \overset{g}{\longmapsto} BAX$$
$$\| \qquad (C = BA) \quad CX \qquad (g \circ f)$$

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

Matrici di isomorfismi

$\dim V = \dim W = n$
$\mathcal{B} = (b_1, \dots, b_n)$ base per V
$\mathcal{C} = (c_1, \dots, c_n)$ base per W
$f : V \to W$ isomorfismo

$$f^{-1} \circ f = \mathrm{id}_V \Rightarrow M_{\mathcal{C}}^{\mathcal{B}}(f^{-1}) \cdot M_{\mathcal{B}}^{\mathcal{C}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V) = I_n$$
$$f \circ f^{-1} = \mathrm{id}_W \Rightarrow M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(f^{-1}) = M_{\mathcal{C}}^{\mathcal{C}}(\mathrm{id}_W) = I_n$$

Dunque $M_{\mathcal{B}}^{\mathcal{C}}(f) \in M_n(\mathbb{K})$ è invertibile e si ha

$$M_{\mathcal{C}}^{\mathcal{B}}(f^{-1}) = \left(M_{\mathcal{B}}^{\mathcal{C}}(f)\right)^{-1}$$

Cor. *$A \in M_n(\mathbb{K})$ invertibile $\Leftrightarrow \operatorname{rg} A = n$ (rango massimo).*

Matrice del cambio di base

$\mathcal{B} = (b_1, \dots, b_n)$ base per V con coordinate $X = (x_1, \dots, x_n)$
$\mathcal{B}' = (b'_1, \dots, b'_n)$ altra base per V con coordinate $X' = (x'_1, \dots, x'_n)$
$M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \in \mathrm{GL}_n(\mathbb{K})$ si chiama *matrice del cambio di base da $\mathcal{B}$ a $\mathcal{B}'$*.

$$X' = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \cdot X$$

Oss. La matrice del cambio di base da $\mathcal{B}$ a $\mathcal{B}'$ ha come j-esima colonna le coordinate di b_j rispetto alla base $\mathcal{B}'$.

N.B. Una matrice del cambio di base è sempre quadrata e *invertibile*.

Cambio di base per applicazioni lineari

$f : V \to W$ lineare, $\dim V = n$, $\dim W = m$
$\mathcal{B} = (b_1, \dots, b_n)$, $\mathcal{B}' = (b'_1, \dots, b'_n)$ basi per V
$\mathcal{C} = (c_1, \dots, c_m)$, $\mathcal{C}' = (c'_1, \dots, c'_m)$ basi per W

$$f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$$

$$V \xrightarrow{\ f\ } W$$

$$\begin{array}{ccc} \mathcal{B} & M_{\mathcal{B}}^{\mathcal{C}}(f) & \mathcal{C} \\ \mathcal{B}' & M_{\mathcal{B}'}^{\mathcal{C}'}(f) & \mathcal{C}' \end{array}$$

$M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)$ $M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W)$

$$M_{\mathcal{B}'}^{\mathcal{C}'}(f) = M_{\mathcal{C}}^{\mathcal{C}'}(\mathrm{id}_W) \cdot M_{\mathcal{B}}^{\mathcal{C}}(f) \cdot M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)$$

Cambio di base per gli endomorfismi. $f : V \to V$ endomorfismo. Consideriamo due basi $\mathcal{B}$ e $\mathcal{B}'$ per V. Poniamo $\mathcal{C} = \mathcal{B}$ e $\mathcal{C}' = \mathcal{B}'$.

$$M_{\mathcal{B}'}^{\mathcal{B}'}(f) = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \cdot M_{\mathcal{B}}^{\mathcal{B}}(f) \cdot M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)$$

Poniamo $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$, $A' = M_{\mathcal{B}'}^{\mathcal{B}'}(f) \in M_n(\mathbb{K})$, $P = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V) \in \mathrm{GL}_n(\mathbb{K})$.
Ricordiamo che

$$M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) = \left(M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)\right)^{-1}$$

$$A' = P^{-1}AP$$

Prop. *$A, A' \in M_n(\mathbb{K})$ rappresentano lo stesso endomorfismo rispetto a due basi se e solo se A e A' sono simili.*

Oss. $A, B \in M_n(\mathbb{K})$ simili $\Rightarrow \mathrm{rg}\, A = \mathrm{rg}\, B$.

Nucleo e immagine. Per trovare il nucleo di $f : V \to W$ in coordinate, si fissano basi $\mathcal{B}$ per V e $\mathcal{C}$ per $W \rightsquigarrow A = M_{\mathcal{B}}^{\mathcal{C}}(f)$ e si risolve il sistema

$$\ker f : \ AX = 0_{\mathbb{K}^m}.$$

Lo spazio delle soluzioni di questo sistema omogeneo rappresenta $\ker f$.
Lo spazio delle colonne di A, ossia $\mathrm{span}(A_{(1)}, \dots, A_{(n)})$, rappresenta $\mathrm{im}\, f$.

$$\mathrm{rg}\, f = \mathrm{rg}\, A.$$

Determinante come funzione multilineare

Una matrice $n \times n$ è essenzialmente una n-upla ordinata di vettori colonna di $\mathbb{K}^n$, ossia $A = (v_1, \dots, v_n)$ con $v_i = A_{(i)}$.

$$\det : M_n(\mathbb{K}) = \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n \text{ volte}} \to \mathbb{K}$$

è funzione delle colonne di una matrice. Si può dimostrare che det soddisfa le seguenti proprietà (dove ci sono ... non modifichiamo nulla):

$$\det(\dots, v + w, \dots) = \det(\dots, v, \dots) + \det(\dots, w, \dots)$$
$$\det(\dots, \alpha v, \dots) = \alpha \det(\dots, v, \dots).$$

In altre parole det è lineare rispetto a ciascuna colonna, tenendo fisse le altre. Per questo motivo si dice che det è *multilineare nelle colonne*, ossia è lineare rispetto a ciascuna colonna.

Dato che $\det A = \det {}^tA$ si ha che det è *multilineare nelle righe*.

Spazio delle applicazioni lineari

Definiamo lo *spazio delle applicazioni lineari* da V a W

$$\mathrm{Hom}(V, W) \stackrel{\text{def}}{=} \{f : V \to W \mid f \text{ lineare}\}$$

Date applicazioni lineari $f, g : V \to W$ definiamo la *somma*

$$f + g : V \to W$$
$$(f + g)(v) \stackrel{\text{def}}{=} f(v) + g(v)$$

e la *moltiplicazione scalare* per $\alpha \in \mathbb{K}$

$$\alpha f : V \to W$$
$$(\alpha f)(v) \stackrel{\text{def}}{=} \alpha f(v).$$

È facile dimostrare che $f + g$ e αf sono lineari e con queste operazioni $\mathrm{Hom}(V, W)$ è un $\mathbb{K}$-spazio vettoriale.

$\mathcal{B} = (b_1 \dots, b_n)$ base per V
$\mathcal{C} = (c_1 \dots, c_m)$ base per W
$\forall f, g \in \mathrm{Hom}(V, W)$, $\forall \alpha \in \mathbb{K} \Rightarrow$

$$M_{\mathcal{B}}^{\mathcal{C}}(f + g) = M_{\mathcal{B}}^{\mathcal{C}}(f) + M_{\mathcal{B}}^{\mathcal{C}}(g)$$
$$M_{\mathcal{B}}^{\mathcal{C}}(\alpha f) = \alpha M_{\mathcal{B}}^{\mathcal{C}}(f)$$

Prop. *$M_{\mathcal{B}}^{\mathcal{C}} : \mathrm{Hom}(V, W) \to M_{m,n}(\mathbb{K})$, $f \mapsto M_{\mathcal{B}}^{\mathcal{C}}(f)$, è un isomorfismo.*

$M_{m,n}(\mathbb{K}) \cong \mathbb{K}^{mn} \Rightarrow \dim M_{m,n}(\mathbb{K}) = mn \Rightarrow \dim \mathrm{Hom}(V, W) = \dim V \dim W$.

Definiamo lo *spazio degli endomorfismi* di V

$$\mathrm{End}(V) \stackrel{\text{def}}{=} \mathrm{Hom}(V, V) = \{f : V \to V \mid f \text{ lineare}\}.$$

$\mathrm{End}(V)$ è uno spazio vettoriale e $\dim \mathrm{End}(V) = (\dim V)^2$.

Determinante di un endomorfismo

Def. Sia $f \in \mathrm{End}(V)$ e sia $\mathcal{B} = (b_1, \dots, b_n)$ una base per V. Si chiama *determinante* di f il numero

$$\det f \stackrel{\text{def}}{=} \det M_{\mathcal{B}}^{\mathcal{B}}(f) \in \mathbb{K}.$$

Oss. Il determinante di un endomorfismo è definito come il determinante di una sua matrice rispetto ad una base arbitraria di V, usando la *stessa base* sia nel dominio che nel codominio.

Oss. $\mathcal{C} = (c_1, \dots, c_n)$ altra base per $V \Rightarrow \det M_{\mathcal{C}}^{\mathcal{C}}(f) = \det M_{\mathcal{B}}^{\mathcal{B}}(f)$ perché sono matrici simili. Quindi $\det f$ è ben definito (non dipende dalla base).

Oss. $f, g \in \mathrm{End}(V) \Rightarrow \det(f \circ g) = \det(f)\det(g)$ (Teorema di Binet).

Oss. f isomorfismo $\Leftrightarrow M_{\mathcal{B}}^{\mathcal{B}}(f) \in \mathrm{GL}_n(\mathbb{K}) \Leftrightarrow \det f \neq 0$. $\det(f^{-1}) = \dfrac{1}{\det f}$.

Prop. *Sia $f \in \mathrm{End}(V)$ e $0 \neq \dim V < \infty$. Allora* $\ker f \neq 0 \Leftrightarrow \det f = 0$.

Problema della diagonalizzazione

Dato $f \in \mathrm{End}(V)$ vorremmo trovare una base $\mathcal{D} = (v_1, \dots, v_n)$ per V t.c. $M_{\mathcal{D}}^{\mathcal{D}}(f)$ sia la più semplice possibile: una matrice diagonale.

Def. $f \in \mathrm{End}(V)$ è *diagonalizzabile* se esiste una base $\mathcal{D}$ per V t.c. $M_{\mathcal{D}}^{\mathcal{D}}(f)$ sia una matrice diagonale. $\mathcal{D}$ è detta *base diagonalizzante*.

Ricordando che la i-esima colonna di $M_{\mathcal{D}}^{\mathcal{D}}(f)$ è il vettore delle coordinate di $f(v_i)$ si ha subito la seguente proposizione.

Prop. $\mathcal{D} = (v_1, \dots, v_n)$ *base diagonalizzante* $f \Leftrightarrow \exists\, \lambda_1, \dots, \lambda_n \in \mathbb{K}$ *t.c.*

$$f(v_i) = \lambda_i v_i, \quad \forall\, i = 1, \dots, n.$$

Inoltre $M_{\mathcal{D}}^{\mathcal{D}}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Oss. Se $M_{\mathcal{D}}^{\mathcal{D}}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ allora f si scrive in coordinate come

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix}$$

Autovalori e autovettori

Def. Dato $f \in \mathrm{End}(V)$, uno scalare $\lambda \in \mathbb{K}$ è detto *autovalore* di f se esiste un vettore *non nullo* $v \in V$ t.c. $f(v) = \lambda v$. In questo caso v è detto *autovettore per f relativo all'autovalore* λ.

Oss. $v \in V$ autovettore per $f \Leftrightarrow v \neq 0_V$ e $\exists\, \lambda \in \mathbb{K}$ t.c. $f(v) = \lambda v$.

N.B. Vogliamo $v \neq 0_V$ perché altrimenti la definizione sarebbe banale: per ogni $\lambda \in \mathbb{K}$ si ha $f(0_V) = \lambda 0_V$.

Cor. *Una base per V diagonalizza $f \in \mathrm{End}(V) \Leftrightarrow$ è formata da autovettori.*

Polinomio caratteristico

Supponiamo $\dim V = n < \infty$.

Def. Il *polinomio caratteristico* di $f \in \mathrm{End}(V)$ è $p_f(x) \stackrel{\text{def}}{=} \det(f - x\,\mathrm{id}_V)$.

Def. Il *polinomio caratteristico* di $A \in M_n(\mathbb{K})$ è $p_A(x) \stackrel{\text{def}}{=} \det(A - xI_n)$.

Oss. p_A è un polinomio di grado n perché si esprime mediante prodotti e somme di entrate della matrice

$$A - xI_n = \begin{pmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{pmatrix}$$

Il prodotto delle entrate sulla diagonale principale è l'unico termine da cui scaturisce $(-1)^n x^n$ (gli altri termini sono monomi di grado $< n$). Quindi $p_A(x)$ ha grado n e coefficiente direttore $(-1)^n = \pm 1$.
Il termine noto è $p_A(0) = \det A$.

Esempio.

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad p_A(x) = \begin{vmatrix} 1 - x & 2 \\ -1 & 3 - x \end{vmatrix} = (1 - x)(3 - x) + 2$$

$$p_A(x) = x^2 - 4x + 5.$$

Prop. *Sia $f : V \to V$ endomorfismo di un $\mathbb{K}$-spazio vettoriale V. Allora $\lambda \in \mathbb{K}$ è autovalore di $f \Leftrightarrow \lambda \in \mathbb{K}$ è radice di p_f, ossia $p_f(\lambda) = 0$.*

Dim. $\lambda \in \mathbb{K}$ autovalore di $f \Leftrightarrow \exists\, v \in V$ non nullo t.c.
$f(v) = \lambda v \Leftrightarrow (f - \lambda\,\mathrm{id}_V)(v) = 0_V \Leftrightarrow \ker(f - \lambda\,\mathrm{id}_V) \neq \{0_V\} \Leftrightarrow$
$p_f(\lambda) = \det(f - \lambda\,\mathrm{id}_V) = 0$. □

Prop. *$\mathcal{B} = (b_1, \ldots, b_n)$ base per V e $A = M_{\mathcal{B}}^{\mathcal{B}}(f) \Rightarrow p_f(x) = p_A(x)$.*
Quindi p_f è un polinomio di grado $n = \dim V$.

Dim. $p_f(x) = \det(f - x\,\mathrm{id}_V) = \det M_{\mathcal{B}}^{\mathcal{B}}(f - x\,\mathrm{id}_V) = \det(M_{\mathcal{B}}^{\mathcal{B}}(f) - xM_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V))$
$= \det(A - xI_n) = p_A(x)$. □

N.B. Se $\mathbb{K} = \mathbb{R}$ gli autovalori di $f \in \mathrm{End}(V)$ sono le radici *reali* di p_f.
Se $\mathbb{K} = \mathbb{C}$ si considerano tutte le radici di p_f.

Def. Gli autovalori e gli autovettori di una matrice $A \in M_n(\mathbb{K})$ sono gli autovalori e autovettori di $L_A \in \mathrm{End}(\mathbb{K}^n)$.

Oss.
$\lambda \in \mathbb{K}$ autovalore di $A \Leftrightarrow p_A(\lambda) = 0$.
$v \in \mathbb{K}^n - \{0_{\mathbb{K}^n}\}$ autovettore di A relativo a $\lambda \Leftrightarrow (A - \lambda I_n)v = 0_{\mathbb{K}^n} \Leftrightarrow$ v soluzione non nulla del sistema omogeneo

$$(A - \lambda I_n)X = 0_{\mathbb{K}^n}.$$

Def. Se $\lambda \in \mathbb{K}$ è un autovalore di $f \in \mathrm{End}(V)$, l'*autospazio* di λ è

$$\mathrm{Aut}(\lambda) \stackrel{\text{def}}{=} \ker(f - \lambda\,\mathrm{id}_V) \subset V.$$

Oss. $\mathrm{Aut}(\lambda)$ sottospazio vettoriale di V in quanto ker e $\dim \mathrm{Aut}(\lambda) \geqslant 1$.

Oss. $\mathrm{Aut}(\lambda) = \{v \in V \mid f(v) = \lambda v\}$. $v \in \mathrm{Aut}(\lambda) \Leftrightarrow f(v) - \lambda v = 0_V$. $\mathrm{Aut}(\lambda)$ contiene tutti gli autovettori di f relativi a λ, più il vettore nullo.

Oss. Per diagonalizzare $f \in \mathrm{End}(V)$ si determina in primo luogo una matrice $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$, si calcola il polinomio caratteristico $p_f = p_A$ e si determinano gli zeri in $\mathbb{K}$ di p_f (autovalori). Per ciascun autovalore si determina una base dell'autospazio e si mettono insieme tali basi. Se in questo modo si ottiene una base $\mathcal{D}$ per V allora $\mathcal{D}$ è una base diagonalizzante.

Esempio. Diagonalizziamo $L_B : \mathbb{R}^2 \to \mathbb{R}^2$ con

$$B = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \quad p_B(x) = \begin{vmatrix} 1-x & 2 \\ 1 & -x \end{vmatrix} = x^2 - x - 2,\ \lambda_1 = -1,\ \lambda_2 = 2$$

$$-1)\ \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad x + y = 0 \quad \begin{cases} x = -t \\ y = t \end{cases} \qquad v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$2)\ \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad x - 2y = 0 \quad \begin{cases} x = 2t \\ y = t \end{cases} \qquad v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

v_1 base per $\mathrm{Aut}(-1)$; v_2 base per $\mathrm{Aut}(2)$.
$\mathcal{D} = (v_1, v_2)$ base diagonalizzante per L_B (e per B).

$$D = M_{\mathcal{D}}^{\mathcal{D}}(L_B) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} = P^{-1}BP, \quad P = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} = M_{\mathcal{D}}^{\mathcal{E}_2}(\mathrm{id}_{\mathbb{R}^2})$$

Molteplicità algebrica e geometrica. Supponiamo che V sia un $\mathbb{K}$-spazio vettoriale con $1 \leqslant \dim V = n < \infty$.

Def. La *molteplicità algebrica* $m_a(\lambda)$ di un autovalore λ è la molteplicità di λ come radice del polinomio caratteristico $p_f(x)$.

Oss. $p_f(x)$ divisibile per $(x - \lambda)^{m_a(\lambda)}$ e non divisibile per $(x - \lambda)^{m_a(\lambda)+1}$.

Oss. $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ autovalori distinti di $f \in \mathrm{End}(V)$, $m_i = m_a(\lambda_i) \Rightarrow$

$$p_f(x) = (-1)^n(x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k} g(x)$$

con $g(x)$ polinomio senza radici in $\mathbb{K}$ ($g = 1$ se p_f ha tutte le radici in $\mathbb{K}$).

Oss. $m_a(\lambda_1) + \cdots + m_a(\lambda_k) \leqslant \dim V$ e $= \Leftrightarrow p_f$ ha tutte le radici in $\mathbb{K}$.

Def. La *molteplicità geometrica* $m_g(\lambda)$ di un autovalore λ è la dimensione del corrispondente autospazio, ossia $m_g(\lambda) \stackrel{\text{def}}{=} \dim \mathrm{Aut}(\lambda)$.

Oss. $m_g(\lambda) = \dim V - \mathrm{rg}(f - \lambda\,\mathrm{id}_V) = n - \mathrm{rg}(A - \lambda I_n)$, $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$.

Valgono le disuguaglianze seguenti (la prima e l'ultima per definizione)

$$1 \leqslant m_g(\lambda) \leqslant m_a(\lambda) \leqslant \dim V.$$

Oss. $m_g(\lambda_1) + \cdots + m_g(\lambda_k) \leqslant \dim V$.

Teor. $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ *autovalori distinti di* $f \in \mathrm{End}(V)$ *con autovettori rispettivamente* $v_1, \ldots, v_k \in V \Rightarrow v_1, \ldots, v_k$ *linearmente indipendenti.*

Dim. Per ipotesi $f(v_i) = \lambda_i v_i$. Procediamo per induzione su $k \geqslant 1$.
Base dell'induzione: $k = 1$. Evidente perché $v_1 \neq 0_V$.

Ipotesi induttiva: vero per $k - 1 \geqslant 1$. Consideriamo una qualunque combinazione lineare nulla di $v_1 \ldots, v_k$ e mostriamo che è banale.

$$\sum_{i=1}^{k} \alpha_i v_i = 0_V \;\Rightarrow\; 0_V = f\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i f(v_i) = \sum_{i=1}^{k} \alpha_i \lambda_i v_i =$$

$$\sum_{i=1}^{k} \alpha_i \lambda_i v_i - \underbrace{\lambda_k \sum_{i=1}^{k} \alpha_i v_i}_{0_V} = \sum_{i=1}^{k-1} \alpha_i (\lambda_i - \lambda_k) v_i.$$

Ipotesi induttiva $\Rightarrow \alpha_i(\lambda_i - \lambda_k) = 0$, $\forall\, i = 1, \ldots, k-1$.
$\lambda_i \neq \lambda_k$, $\forall\, i \leqslant k-1 \Rightarrow \alpha_i = 0$, $\forall\, i \leqslant k-1 \Rightarrow \alpha_k v_k = 0_V \Rightarrow \alpha_k = 0$. □

Teorema di diagonalizzazione. *Siano* $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ *gli autovalori distinti di* f. *Le seguenti sono equivalenti:*

1) p_f *ha tutte le radici in* $\mathbb{K}$ *e* $m_g(\lambda_i) = m_a(\lambda_i)$ $\forall\, i$.

2) $m_g(\lambda_1) + \cdots + m_g(\lambda_k) = \dim V$.

3) f *diagonalizzabile.*

In questo caso, scelta $\mathcal{D}_i$ *base per* $\mathrm{Aut}(\lambda_i)$ $\forall\, i \Rightarrow \mathcal{D} = \mathcal{D}_1 \cup \cdots \cup \mathcal{D}_k$ *base diagonalizzante per* f.

Dim. *1) ⇒ 2)* $\sum_{i=1}^{k} m_g(\lambda_i) = \sum_{i=1}^{k} m_a(\lambda_i) = \dim V$ (vedi osservazione sopra).

2) ⇒ 1) $\dim V = \sum_{i=1}^{k} m_g(\lambda_i) \leqslant \sum_{i=1}^{k} m_a(\lambda_i) \leqslant \dim V \Rightarrow$ tutte uguaglianze.

2) ⇒ 3) $\mathcal{D} = (v_1, \ldots, v_n)$ insieme di autovettori.

$$\sum_{i=1}^{n} \alpha_i v_i = u_1 + \cdots + u_k = 0_V$$

con u_j somma dei termini $\in \mathrm{Aut}(\lambda_j) \Rightarrow u_j = 0 \Rightarrow \alpha_i = 0 \Rightarrow \mathcal{D}$ base.

$\boxed{3) \Rightarrow 2)}$ $\mathcal{D}$ base diagonalizzante (formata da autovettori), n_i numero di autovettori di $\mathcal{D}$ che appartengono a $\mathrm{Aut}(\lambda_i) \Rightarrow n_i \leqslant m_g(\lambda_i)$.

$\dim V = \sum_{i=1}^{k} n_i \leqslant \sum_{i=1}^{k} m_g(\lambda_i) \leqslant \dim V \Rightarrow$ tutte uguaglianze. □

Cor. *Se f ha $n = \dim V$ autovalori distinti $\lambda_1, \ldots, \lambda_n \in \mathbb{K} \Rightarrow f$ diagonalizzabile. In questo caso $m_g(\lambda_i) = m_a(\lambda_i) = 1$, $\forall\, i = 1, \ldots, n$.*

Oss. p_f ha una radice $\lambda \notin \mathbb{K} \Rightarrow f$ non diagonalizzabile.

Metodo per diagonalizzare un endomorfismo

Dato $f \in \mathrm{End}(V)$ con $\dim(V) = n$ per capire se f è diagonalizzabile e trovare una base diagonalizzante si procede come segue:

1) si sceglie (se già non è data) una base $\mathcal{B}$ per V;

2) si determina $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$;

3) si calcola $p_f(x) = p_A(x) = \det(A - xI_n)$;

4) si calcolano le radici distinte $\lambda_1, \ldots, \lambda_k$ di $p_f(x)$;

5) se $\lambda_i \notin \mathbb{K}$ o $m_g(\lambda_i) < m_a(\lambda_i)$ per un certo $\lambda_i \Rightarrow f$ non diagonalizzabile;

6) altrimenti per ogni λ_i si risolve il sistema omogeneo

$$(A - \lambda_i I_n)X = 0_{\mathbb{K}^n}$$

trovando una base $\mathcal{D}_i$ per lo spazio delle soluzioni $\mathrm{Aut}(\lambda_i)$;

7) $\mathcal{D} = \mathcal{D}_1 \cup \cdots \cup \mathcal{D}_k$ base diagonalizzante.

Se al punto (*4*) non abbiamo verificato la disuguaglianza, si contano i vettori ottenuti al punto (*6*): se il numero è uguale a $\dim V$ abbiamo base diagonalizzante. Altrimenti f non è diagonalizzabile.

In pratica si lavora sempre su A, ossia si *diagonalizza una matrice*.

Esempio. Diagonalizziamo $L_A : \mathbb{C}^2 \to \mathbb{C}^2$

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad p_A(x) = x^2 - 4x + 5, \ \lambda_1 = 2 + i, \ \lambda_2 = 2 - i$$

$\lambda_1)$ $\begin{pmatrix} -1-i & 2 \\ -1 & 1-i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0, \quad -x + (1-i)y = 0 \quad \begin{cases} x = (1-i)t \\ y = t \end{cases}$

$$v_1 = \begin{pmatrix} 1-i \\ 1 \end{pmatrix} \text{ base di } \mathrm{Aut}(2+i)$$

$\lambda_2)$ $\begin{pmatrix} -1+i & 2 \\ -1 & 1+i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0, \quad -x + (1+i)y = 0 \quad \begin{cases} x = (1+i)t \\ y = t \end{cases}$

$$v_2 = \begin{pmatrix} 1+i \\ 1 \end{pmatrix} \text{ base di } \mathrm{Aut}(2-i)$$

$\mathcal{D} = (v_1, v_2)$ base diagonalizzante $D = M_{\mathcal{D}}^{\mathcal{D}}(L_A) = \begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix}$

$P = \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix} = M_{\mathcal{D}}^{\mathcal{E}_2}(\mathrm{id}_{\mathbb{R}^2})$, $D = P^{-1}AP$.

Oss. $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ non è diagonalizzabile perché non ha autovalori reali.

Esempio. Diagonalizziamo $L_B : \mathbb{R}^3 \to \mathbb{R}^3$

$$B = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \quad p_B(x) = \begin{vmatrix} -x & 1 & -1 \\ -1 & 2-x & -1 \\ 1 & -1 & 2-x \end{vmatrix} = -x^3 + 4x^2 - 5x + 2$$

Si cercano le radici calcolando p_B sui divisori di 2 e si fattorizza.

$$p_B(x) = -(x-1)^2(x-2)$$
$$\lambda_1 = 1, \ m_a(1) = 2$$
$$\lambda_2 = 2, \ m_a(2) = 1$$

$$B - I_3 = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \ \mathrm{rg}(B - I_3) = 1 \Rightarrow m_g(1) = 2$$

$$B - 2I_3 = \begin{pmatrix} -2 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}, \ \mathrm{rg}(B - 2I_3) = 2 \Rightarrow m_g(2) = 1$$

L_B diagonalizzabile. Calcoliamo basi degli autospazi.

1) $x - y + z = 0 \qquad \begin{cases} x = t - u \\ y = t \\ z = u \end{cases} \qquad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

2) $\begin{cases} -x - z = 0 \\ x - y = 0 \end{cases} \qquad \begin{cases} x = t \\ y = t \\ z = -t \end{cases} \qquad v_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

(v_1, v_2) base per Aut(1); v_3 base per Aut(2)
$\mathcal{D} = (v_1, v_2, v_3)$ base diagonalizzante per L_B.

$$D = M_{\mathcal{D}}^{\mathcal{D}}(L_B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad P = M_{\mathcal{D}}^{\mathcal{E}_3}(\mathrm{id}_{\mathbb{R}^3}) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

$$D = P^{-1}BP \qquad B = PDP^{-1}$$

Esempio. Proviamo a capire se $L_J : \mathbb{R}^2 \to \mathbb{R}^2$ è diagonalizzabile

$$J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad p_J(x) = \begin{vmatrix} 1-x & 1 \\ 0 & 1-x \end{vmatrix} = (1-x)^2, \ \lambda = 1, \ m_a(1) = 2$$

$$J - I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ \mathrm{rg}(J - I_2) = 1 \Rightarrow m_g(1) = 2 - 1 = 1 < m_a(1)$$

$L_J : \mathbb{R}^2 \to \mathbb{R}^2$ non è diagonalizzabile.
$L_J : \mathbb{C}^2 \to \mathbb{C}^2$ non è diagonalizzabile (per lo stesso motivo).

Forme bilineari

Da questo momento V sarà uno *spazio vettoriale reale* ($\mathbb{K} = \mathbb{R}$).

Def. Una *forma bilineare* su V è una funzione

$$b : V \times V \to \mathbb{R}$$
$$(v, w) \mapsto b(v, w)$$

che sia lineare in ciascun argomento, ossia $\forall\, u, v, w \in V$, $\forall\, \alpha \in \mathbb{R}$ si ha

1) $b(u + v, w) = b(u, w) + b(v, w)$

2) $b(u, v + w) = b(u, v) + b(u, w)$

3) $b(\alpha v, w) = b(v, \alpha w) = \alpha b(v, w)$.

Oss. $b(0_V, v) = b(v, 0_V) = 0$. Infatti $b(0_V, v) = b(00_V, v) = 0b(0_V, v) = 0$.

Def. $b : V \times V \to \mathbb{R}$ è *simmetrica* se $b(v, w) = b(w, v)$, $\forall\, v, w \in V$.

Def. $b : V \times V \to \mathbb{R}$ è *definita positiva* se $b(v, v) > 0$, $\forall\, v \in V - \{0_V\}$.

N.B. $b(0_V, 0_V) = 0$.

Def. Un *prodotto scalare* su V è una forma bilineare, simmetrica, definita positiva $b : V \times V \to \mathbb{R}$. Scriviamo $b(v, w) = \langle v, w \rangle$ (leggi v *scalare* w). Uno *spazio vettoriale Euclideo* è uno spazio vettoriale reale V con un prodotto scalare $\langle\ ,\ \rangle$.

Forma bilineare associata ad una matrice quadrata.

$A \in M_n(\mathbb{R}) \rightsquigarrow b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

$$b_A(X, Y) \stackrel{\text{def}}{=} {}^tXAY,\ \forall\, X, Y \in \mathbb{R}^n \quad \text{(vettori colonna)}.$$

b_A bilineare per la proprietà distributiva del prodotto righe per colonne.

Def. b_A è detta *forma bilineare associata ad* $A \in M_n(\mathbb{R})$.

$$A = (a_{ij}) \rightsquigarrow b_A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij} x_i y_j$$

Oss. b_A simmetrica $\Leftrightarrow$ A simmetrica, ossia $A = {}^tA$. Infatti: $\forall\, X, Y \in \mathbb{R}^n$ $b_A(X, Y) = b_A(Y, X) \Leftrightarrow {}^tXAY = {}^tYAX = {}^t({}^tYAX) = {}^tX\,{}^tAY \Leftrightarrow A = {}^tA$.

N.B. Non confondere la forma bilineare e l'applicazione lineare associata

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \qquad L_A : \mathbb{R}^n \to \mathbb{R}^n.$$

Prodotto scalare canonico su $\mathbb{R}^n$. Forma bilineare associata a I_n.

$$\langle\ ,\ \rangle = b_{I_n} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
$$\langle X, Y \rangle = {}^tXY = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n$$

$\langle X, Y \rangle = \langle Y, X \rangle$, $\forall\, X, Y \in \mathbb{R}^n$ (simmetrica).
$\langle X, X \rangle = x_1^2 + \cdots + x_n^2 > 0$, $\forall\, X \in \mathbb{R}^n - \{0\}$ (definita positiva).

Def. Una matrice quadrata simmetrica $A \in M_n(\mathbb{R})$ è *definita positiva* se $b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ è una forma bilineare simmetrica definita positiva.

Oss. $A \in M_n(\mathbb{R})$ simmetrica $\Rightarrow b_A(X,X) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$.

Esempio. I_n definita positiva perché induce il prodotto scalare canonico.

Esempio.

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

b_A simmetrica perché $A = {}^tA$. Abbiamo

$${}^tXAY = (x_1\ x_2)\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$$

b_A definita positiva infatti

$$\begin{aligned} {}^tXAX &= 2x_1^2 + 2x_1x_2 + x_2^2 = \\ &= x_1^2 + (x_1+x_2)^2 > 0, \quad \forall\, X \in \mathbb{R}^2 - \{0\}. \end{aligned}$$

Quindi A definisce un prodotto scalare (non canonico) su $\mathbb{R}^2$.

Def. Un *minore principale* di $A \in M_n(\mathbb{R})$ è il det di una sottomatrice quadrata formata dalle prime p righe e colonne di A, con $1 \leqslant p \leqslant n$.

Un altro metodo per verificare se A è definita positiva è il seguente.

Prop. *Una matrice simmetrica $A \in M_n(\mathbb{R})$ è definita positiva $\Leftrightarrow$ tutti i minori principali di A sono > 0.*

Oss. $A \in M_2(\mathbb{R})$ simmetrica è definita positiva $\Leftrightarrow a_{11} > 0$ e $\det A > 0$.

Matrice di una forma bilineare. $b : V \times V \to \mathbb{R}$ forma bilineare $\mathcal{C} = (c_1, \ldots, c_n)$ base per $V \rightsquigarrow \rightsquigarrow M_{\mathcal{C}}(b) \stackrel{\text{def}}{=} (b(c_i, c_j)) \in M_n(\mathbb{R})$.

Def. $M_{\mathcal{C}}(b)$ è detta *matrice di b rispetto alla base $\mathcal{C}$*.

$$A = M_{\mathcal{C}}(b),\ a_{ij} = b(c_i, c_j),\ v = X^{\mathcal{C}},\ w = Y^{\mathcal{C}} \in V \rightsquigarrow$$

$$b(v,w) = b\left(\sum_{i=1}^{n} x_ic_i, \sum_{j=1}^{n} y_jc_j\right) = \sum_{i,j=1}^{n} x_iy_jb(c_i,c_j)$$

$$= \sum_{i,j=1}^{n} a_{ij}x_iy_j = {}^tXAY = b_A(X,Y).$$

$M_{\mathcal{C}}(b)$ consente di calcolare b in coordinate mediante b_A.

Matrice di b_A rispetto alla base canonica. $A \in M_n(\mathbb{R}) \rightsquigarrow$ $b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $b_A(X,Y) = {}^tXAY \Rightarrow M_{\mathcal{E}_n}(b_A) = A$.

Prop. *$b : V \times V \to \mathbb{R}$ prod. scal. $\Leftrightarrow$ $M_{\mathcal{C}}(b)$ simmetrica definita positiva.*

Matrici congruenti

Def. Due matrici $A, B \in M_n(\mathbb{K})$ sono *congruenti* se $\exists\, P \in \mathrm{GL}_n(\mathbb{K})$ t.c.

$$B = {}^tPAP.$$

Oss. Matrici congruenti hanno lo stesso rango.

N.B. Non confondere congruenza $B = {}^tPAP$ e similitudine $B = P^{-1}AP$.

Cambio di base per forme bilineari

Teor. *$b : V \times V \to \mathbb{R}$ bilineare, $\mathcal{C}$, $\mathcal{C}'$ basi per V $\Rightarrow$*

$$M_{\mathcal{C}'}(b) = {}^tM_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_V) \cdot M_{\mathcal{C}}(b) \cdot M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_V).$$

Dim. $A = M_{\mathcal{C}}(b)$, $A' = M_{\mathcal{C}'}(b)$, $P = M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_V)$. $\forall\, v, w \in V$ con coordinate

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^{\mathcal{C}} = X^{\mathcal{C}}, \qquad w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^{\mathcal{C}} = Y^{\mathcal{C}} \text{ rispetto a } \mathcal{C}$$

$$v = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}^{\mathcal{C}'} = (X')^{\mathcal{C}'}, \qquad w = \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix}^{\mathcal{C}'} = (Y')^{\mathcal{C}'} \text{ rispetto a } \mathcal{C}'$$

$$X = PX', \qquad Y = PY'$$

$b(v, w) = {}^tXAY = {}^t(PX')A(PY') = {}^tX'\,{}^tPAP\,Y' = {}^tX'A'\,Y' \Rightarrow A' = {}^tPAP.$ □

Cor. *Due matrici A e $A' \in M_n(\mathbb{K})$ rappresentano la stessa forma bilineare su V rispetto a due basi opportune $\Leftrightarrow$ A e A' sono congruenti.*

Spazi vettoriali Euclidei

V spazio vettoriale Euclideo con prodotto scalare $\langle\ ,\ \rangle$.

Def. Si chiama *norma* o *lunghezza* di un vettore $v \in V$ il numero

$$\|v\| \stackrel{\text{def}}{=} \sqrt{\langle v, v\rangle}.$$

Oss.

1) $\|v\|$ ben definita perché $\langle\ ,\ \rangle$ definito positivo: $\langle v, v\rangle \geqslant 0$.

2) $\|v\| \geqslant 0$, $\forall\, v \in V$ e $\|v\| = 0 \Leftrightarrow v = 0_V$.

3) $\|\alpha v\| = |\alpha|\|v\|$.

4) $\langle v, v\rangle = \|v\|^2$.

5) $v \neq 0_V \Rightarrow \left\| \dfrac{v}{\|v\|} \right\| = 1$. $\dfrac{v}{\|v\|} \in V$ è detto *normalizzazione* di v.

Disuguaglianza di Cauchy-Schwarz. *$\forall\, v, w \in V$ si ha*

$$|\langle v, w\rangle| \leqslant \|v\|\|w\|.$$

Vale l'uguaglianza $\Leftrightarrow$ v e w sono linearmente dipendenti.

Dim. Se $w = 0_V$ vale l'uguale. Supponiamo $w \neq 0_V \rightsquigarrow \alpha := -\dfrac{\langle v, w\rangle}{\|w\|^2}$

$$0 \leqslant \langle v + \alpha w, v + \alpha w\rangle = \langle v, v\rangle + \langle v, \alpha w\rangle + \langle \alpha w, v\rangle + \langle \alpha w, \alpha w\rangle$$

$$= \|v\|^2 + 2\alpha\langle v, w\rangle + \alpha^2\|w\|^2 = \|v\|^2 - 2\frac{\langle v, w\rangle}{\|w\|^2}\langle v, w\rangle + \frac{\langle v, w\rangle^2}{\|w\|^4}\|w\|^2$$

$$= \|v\|^2 - \frac{\langle v, w\rangle^2}{\|w\|^2} \Rightarrow \langle v, w\rangle^2 \leqslant \|v\|^2\|w\|^2 \Rightarrow |\langle v, w\rangle| \leqslant \|v\|\|w\|.$$

Non dimostriamo l'ultima affermazione. □

Disuguaglianza triangolare per la norma. $\|v + w\| \leqslant \|v\| + \|w\|$.

Dim. $\|v + w\|^2 = \langle v + w, v + w\rangle = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \leqslant$

$$\leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$ □

Metrica Euclidea. La *metrica* o *distanza Euclidea* su V è la funzione

$$d : V \times V \to \mathbb{R}$$

$$d(v, w) \stackrel{\text{def}}{=} \|v - w\|.$$

Proprietà della metrica. *Valgono le seguenti $\forall\, u, v, w \in V$*

1) $d(v, w) \geqslant 0$ *e* $d(v, w) = 0 \Leftrightarrow v = w$.

2) $d(v, w) = d(w, v)$.

3) $d(u, w) \leqslant d(u, v) + d(v, w)$ *(disuguaglianza triangolare per d).*

Dim. *(1)*−*(2)* sono immediate. Dimostriamo la *(3)*:

$d(u, w) = \|u - v + v - w\| \leqslant \|u - v\| + \|v - w\| = d(u, v) + d(v, w)$. □

Oss. $d(v, 0_V) = \|v\|$.

Angolo convesso tra due vettori non nulli. $\forall\, v, w \in V - \{0_V\} \Rightarrow$

$$-1 \leqslant \frac{\langle v, w\rangle}{\|v\|\|w\|} \leqslant 1 \;\Rightarrow\; \exists!\, \vartheta \in [0, \pi] \text{ t.c. } \cos\vartheta = \frac{\langle v, w\rangle}{\|v\|\|w\|}$$

Def. ϑ si chiama *angolo tra* v e w. Scriviamo $\vartheta = \widehat{vw}$.

$$\boxed{\widehat{vw} \stackrel{\text{def}}{=} \arccos\frac{\langle v, w\rangle}{\|v\|\|w\|}} \qquad \boxed{\langle v, w\rangle = \|v\|\|w\|\cos\widehat{vw}}$$

N.B. L'angolo non è definito se uno dei vettori è nullo.

Def. $v, w \in V$ sono *ortogonali* se $\langle v, w\rangle = 0$. Scriviamo $v \perp w$.

Oss. Supponiamo $v, w \in V - \{0_V\}$.

1) $\widehat{vw} = 0 \Leftrightarrow \exists\, \alpha > 0$ t.c. $v = \alpha w$ (*angolo nullo*).

2) $0 < \widehat{vw} < \dfrac{\pi}{2} \Leftrightarrow \langle v, w\rangle > 0$ (*angolo acuto*).

3) $\widehat{vw} = \dfrac{\pi}{2} \Leftrightarrow \langle v, w\rangle = 0$ (*angolo retto*).

4) $\dfrac{\pi}{2} < \widehat{vw} < \pi \Leftrightarrow \langle v, w\rangle < 0$ (*angolo ottuso*).

5) $\widehat{vw} = \pi \Leftrightarrow \exists\, \alpha > 0$ t.c. $v = -\alpha w$ (*angolo piatto*).

Spazi vettoriali Euclidei numerici

Consideriamo $\mathbb{R}^n$ col prodotto scalare canonico

$$\langle v, w\rangle = {}^t v\, w = \sum_{i=1}^{n} x_i y_i$$

$$\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2} \qquad v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$d(v, w) = \|v - w\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\widehat{vw} = \arccos\frac{\langle v, w\rangle}{\|v\|\|w\|}$$

Esempio. Consideriamo in $\mathbb{R}^2$ col prodotto scalare canonico i vettori $v = (1, 2)$, $w = (-2, 1) \Rightarrow \|v\| = \|w\| = \sqrt{5}$, $\langle v, w\rangle = 0 \Rightarrow \widehat{vw} = \dfrac{\pi}{2}$ $d(v, w) = \|v - w\| = \sqrt{10}$.

Esempio. Consideriamo in $\mathbb{R}^3$ col prodotto scalare canonico i vettori $v = (-1, 0, 1)$, $w = (0, 1, 1) \Rightarrow \|v\| = \|w\| = \sqrt{2}$, $\langle v, w \rangle = 1 \Rightarrow \widehat{vw} = \frac{\pi}{3}$ $d(v, w) = \|v - w\| = \sqrt{2}$.

Basi ortonormali

V spazio vettoriale Euclideo con prodotto scalare $\langle\ ,\ \rangle$.

Prop. *Siano $v_1, \dots, v_k \in V$ a due a due ortogonali e non nulli. Allora $v_1, \dots, v_k$ sono linearmente indipendenti.*

Dim. Per ipotesi $\langle v_i, v_j \rangle = 0$, $\forall\, i \neq j$.

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_V$$

$$\langle v_i, \alpha_1 v_1 + \cdots + \alpha_k v_k \rangle = \alpha_i \langle v_i, v_i \rangle = \alpha_i \|v_i\|^2 = 0.$$

$\|v_i\| \neq 0 \Rightarrow \alpha_i = 0$, $\forall\, i$, quindi sono linearmente indipendenti. □

Def. Una base $\mathcal{B} = (b_1, \dots, b_n)$ per V è *ortogonale* se $\langle b_i, b_j \rangle = 0$, $\forall\, i \neq j$. $\mathcal{B}$ è *ortonormale* se $\langle b_i, b_j \rangle = \delta_{ij}$, $\forall\, i, j = 1, \dots, n$.

Oss. $\mathcal{B}$ ortonormale $\Leftrightarrow M_{\mathcal{B}}(\langle\ ,\ \rangle) = I_n$.

Oss. Normalizzando una base ortogonale se ne ottiene una ortonormale.

Oss. La base canonica di $\mathbb{R}^n$ è ortonormale: $\langle e_i, e_j \rangle = \delta_{ij}$.

Ortogonalizzazione di Gram-Schmidt. *$v_1, \dots, v_n \in V$ linearmente indipendenti $\Rightarrow \exists\, u_1, \dots, u_n \in V$ non nulli e a due a due ortogonali t.c.*

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(u_1, \dots, u_k), \quad \forall\, k = 1, \dots, n.$$

Dim. Definiamo gli u_k in modo ricorsivo e procediamo per induzione su k.

$$u_1 := v_1$$

$$u_k := v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i \rangle}{\langle u_i, u_i \rangle} u_i \qquad \text{per } k \geqslant 2.$$

Ipotesi induttiva: per $i, j \leqslant k-1$, $i \neq j$ si ha $\langle u_i, u_j \rangle = 0$. Pertanto:

$$\langle u_k, u_j \rangle = \langle v_k, u_j \rangle - \frac{\langle v_k, u_j \rangle}{\langle u_j, u_j \rangle} \langle u_j, u_j \rangle = 0$$

$\operatorname{span}(u_1, \dots, u_{k-1}) = \operatorname{span}(v_1, \dots, v_{k-1}) \Rightarrow v_k \in \operatorname{span}(u_1, \dots, u_k)$ e
$u_k \in \operatorname{span}(v_1, \dots, v_k) \Rightarrow \operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(u_1, \dots, u_k)$, $u_k \neq 0_V$. □

Oss. I vettori u_k possono essere moltiplicati per scalari non nulli preservando le proprietà dell'enunciato. Questo a volte semplifica i calcoli.

Cor. *V spazio vett. Euclideo,* $\dim V < \infty \Rightarrow V$ *ammette base ortonormale.*

Esempio. $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$

$$u_1 = v_1$$

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{3}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Normalizzando si ha la base ortonormale per $\mathbb{R}^2$

$$w_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad w_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Esempio. $U \subset \mathbb{R}^4$ sottospazio vettoriale di equazione

$$U\colon\ x_1 - x_2 + x_3 + x_4 = 0.$$

Determiniamo una base di U.

$$\begin{cases} x_1 = t_1 - t_2 - t_3 \\ x_2 = t_1 \\ x_3 = t_2 \\ x_4 = t_3 \end{cases} \qquad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ v_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$u_1 = v_1$$

$$u_2 = v_2 - \frac{-1}{2}\,u_1 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix}$$

$$u_3 = v_3 - \frac{-1}{2}\,u_1 - \frac{1}{6}\,u_2 = \frac{1}{6}\begin{pmatrix} -2 \\ 2 \\ -2 \\ 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 \\ 1 \\ -1 \\ 3 \end{pmatrix}$$

Per determinare una base ortonormale per U basta normalizzare gli u_i.

$$w_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ w_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix},\ w_3 = \frac{1}{2\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ -1 \\ 3 \end{pmatrix}.$$

Calcoli mediante basi ortonormali

$\mathcal{B} = (b_1, \ldots, b_n)$ *base ortonormale* per V.

$$\langle v, w\rangle = {}^tXY = \langle X, Y\rangle_{\mathbb{R}^n} = \sum_{i=1}^{n} x_i y_i$$

$$\|v\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

$$d(v, w) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

$$\widehat{vw} = \arccos \frac{\langle v, w\rangle}{\|v\|\|w\|}$$

$$x_i = \langle v, b_i\rangle, \qquad v = \sum_{i=1}^{n} \langle v, b_i\rangle b_i.$$

$$v = \sum_{i=1}^{n} x_i b_i \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$w = \sum_{i=1}^{n} y_i b_i \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Pertanto le formule per prodotto scalare, la norma, distanza rispetto ad una base ortonormale coincidono con le analoghe formule viste su $\mathbb{R}^n$ col prodotto scalare canonico. Le coordinate rispetto ad una base ortonormale sono i prodotti scalari con i vettori di base.

Prop. $\forall\, A \in M_n(\mathbb{R})$ *simmetrica definita positiva* $\Rightarrow \exists\, P \in \mathrm{GL}_n(\mathbb{R})$ *t.c.* $A = {}^tPP$. *Quindi* $\det A > 0$ *e* $\operatorname{rg} A = n$.

Dim. $b_A \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ prodotto scalare $\Rightarrow \exists\, \mathcal{B} = (b_1, \dots, b_n)$ base ortonormale $\Rightarrow M_{\mathcal{B}}(b_A) = I_n$ congruente ad $A \rightsquigarrow P = M^{\mathcal{B}}_{\mathcal{E}_n}(\mathrm{id}_{\mathbb{R}^n}) \in \mathrm{GL}_n(\mathbb{R})$ $\rightsquigarrow A = {}^tPI_nP = {}^tPP \Rightarrow \det A = (\det P)^2 > 0$. □

Esempio. Consideriamo il prodotto scalare su $\mathbb{R}^2$ indotto dalla matrice

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

Abbiamo già visto quest'esempio e sappiamo che A è definita positiva. La base canonica di $\mathbb{R}^2$ non è ortonormale. Usiamo Gram-Schmidt

$$u_1 = e_1$$
$$u_2 = e_2 - \frac{\langle e_2, u_1\rangle}{\langle u_1, u_1\rangle} u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$$\|u_1\| = \sqrt{2} \qquad b_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\|u_2\| = \sqrt{{}^tu_2Au_2} = \sqrt{2} \qquad b_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$\mathcal{B} = (b_1, b_2)$ base ortonormale $\Rightarrow A = {}^tPP$ con $P = (b_1\ b_2)^{-1}$.

Completamento della base ortonormale. $\dim V = n$ *e* $u_1, \dots, u_k \in V$ *vettori ortonormali* $\Rightarrow \exists\, u_{k+1}, \dots, u_n \in V$ *t.c.* $(u_1, \dots, u_n)$ *base ortonormale per* V.

Dim. Completiamo $u_1, \dots, u_k$ ad una base per V aggiungendo certi vettori $v_{k+1}, \dots, v_n$. Applicando Gram-Schmidt e normalizzazione a

$$u_1, \dots, u_k, v_{k+1}, \dots, v_n$$

i primi k vettori non cambiano perché ortonormali tra loro $\rightsquigarrow u_1, \dots, u_n$ base ortonormale per V. □

Sottospazi vettoriali ortogonali. Due sottospazi vettoriali $U, W \subset V$ sono *ortogonali* se $\langle u, w\rangle = 0$, $\forall\, u \in U$ e $\forall\, w \in W$. Scriviamo $U \perp W$.

Oss. $U \perp W \Leftrightarrow$ ogni vettore di U è ortogonale ad ogni vettore di W.

Complemento ortogonale. $U \subset V$ sottospazio vettoriale

$$U^\perp \stackrel{\text{def}}{=} \{v \in V \mid \langle v, u\rangle = 0,\ \forall\, u \in U\}$$

Def. $U^\perp$ è detto *complemento ortogonale* o anche solo *ortogonale* di U.

Oss. $U \cap U^\perp = \{0_V\}$. Infatti $\forall\, v \in U \cap U^\perp \Rightarrow \langle v, v\rangle = 0 \Rightarrow v = 0_V$.

Teor. *$U^\perp \subset V$ è il più grande sottospazio vettoriale di V ortogonale a U.*

Dim. $0_V \in U^\perp$ dato che $\langle 0_V, u\rangle = 0$. $\forall\, v, w \in U^\perp$, $\forall\, \alpha, \beta \in \mathbb{R}$, $\forall\, u \in U \Rightarrow$

$$\langle \alpha v + \beta w, u\rangle = \alpha\langle v, u\rangle + \beta\langle w, u\rangle = 0 \;\Rightarrow\; \alpha v + \beta w \in U^\perp \qquad \square$$

Complemento ortogonale

V spazio vettoriale Euclideo con prodotto scalare $\langle\ ,\ \rangle$.

Teor. $\dim V < \infty \Rightarrow \dim U^\perp = \dim V - \dim U$.

Dim. $\dim V = n$. $(u_1, \dots, u_k)$ base ortonormale per $U \rightsquigarrow u_1, \dots, u_n$ completamento della base ortonormale $\Rightarrow u_{k+1}, \dots, u_n \in U^\perp$. $\forall\, w \in U^\perp$

$$w = \sum_{i=1}^{n} \langle w, u_i\rangle u_i = \sum_{i=k+1}^{n} \langle w, u_i\rangle u_i \in \operatorname{span}(u_{k+1}, \dots, u_n)$$

$\Rightarrow (u_{k+1}, \dots, u_n)$ base ortonormale per $U^\perp \Rightarrow \dim U^\perp = n - k$. □

Cor. *Dato $U \subset V$ sottospazio vettoriale, $\forall\, v \in V \Rightarrow \exists!\, v_U \in U$, $v_{U^\perp} \in U^\perp$ t.c. $v = v_U + v_{U^\perp}$.*

Dim. $(u_1, \dots, u_n)$ base ortonormale per V ottenuta come nella dimostrazione precedente: $U = \operatorname{span}(u_1, \dots, u_k)$, $U^\perp = \operatorname{span}(u_{k+1}, \dots, u_n)$

$$v_U = \sum_{i=1}^{k} \langle v, u_i\rangle u_i, \qquad v_{U^\perp} = \sum_{i=k+1}^{n} \langle v, u_i\rangle u_i.$$ □

Def. v_U è detta *proiezione ortogonale* di v su U.
$v_{U^\perp}$ è detta *componente normale* di v rispetto a U.

Oss. $\mathcal{B} = (b_1, \dots, b_n)$ base ortonormale per V, $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ non tutti nulli

$$U:\ \alpha_1 x_1 + \cdots + \alpha_n x_n = 0$$
$$\dim U = n - 1 \quad (\textit{iperpiano vettoriale})$$
$$v = \alpha_1 b_1 + \cdots + \alpha_n b_n \ \text{ base di } U^\perp$$

infatti l'equazione di U è il prodotto scalare tra il vettore di coordinate $(\alpha_1, \dots, \alpha_n)$ e il vettore di coordinate $(x_1, \dots, x_n)$.

Più in generale i vettori riga della matrice di un sistema lineare omogeneo generano l'ortogonale dello spazio delle soluzioni (e ne sono base se in numero uguale al rango della matrice).

Matrici ortogonali

Def. Una matrice $P \in M_n(\mathbb{R})$ è *ortogonale* se ${}^tPP = I_n$

Oss. $P \in M_n(\mathbb{R})$ ortogonale $\Leftrightarrow P^{-1} = {}^tP$

Oss. P ortogonale $\Rightarrow \det P = \pm 1$. Infatti $1 = \det({}^tPP) = (\det P)^2$.

Def. $\mathrm{O}(n) \stackrel{\text{def}}{=} \{P \in M_n(\mathbb{R}) \mid P \text{ ortogonale}\}$ *gruppo ortogonale*.

Def. $\mathrm{SO}(n) \stackrel{\text{def}}{=} \{P \in \mathrm{O}(n) \mid \det P = 1\}$ *gruppo ortogonale speciale*.

Oss. $I_n \in \mathrm{SO}(n)$ e $\mathrm{SO}(n) \subset \mathrm{O}(n) \subset \mathrm{GL}_n(\mathbb{R})$.

Oss. $\mathrm{O}(n)$ e $\mathrm{SO}(n)$ sono chiusi rispetto a prodotti e inverse:
$A, B \in \mathrm{O}(n) \Rightarrow AB \in \mathrm{O}(n)$ e $A^{-1} \in \mathrm{O}(n)$.
$A, B \in \mathrm{SO}(n) \Rightarrow AB \in \mathrm{SO}(n)$ e $A^{-1} \in \mathrm{SO}(n)$.

Prop. *Le seguenti sono equivalenti*
1) le colonne di P formano una base ortonormale per $\mathbb{R}^n$;

2) le righe di P formano una base ortonormale per $\mathbb{R}^n$;

3) $P \in O(n)$.

Dim. Dimostriamolo per le colonne.
$({}^tPP)_{ij} = ({}^tP)^{(i)} P_{(j)} = {}^tP_{(i)} P_{(j)} = \langle P_{(i)}, P_{(j)} \rangle = \delta_{ij} \ \forall i, j \Leftrightarrow P \in O(n)$. □

Oss. $P \in O(n)$ è la matrice del cambio di base dalla base canonica alla base ortonormale di $\mathbb{R}^n$ formata dale colonne di P.

Teor. *Siano $\mathcal{B} = (b_1, \dots, b_n)$ e $\mathcal{C} = (c_1, \dots, c_n)$ basi per V con $\mathcal{B}$ ortonormale. Allora $\mathcal{C}$ è ortonormale $\Leftrightarrow M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V) \in O(n)$.*

Endomorfismi autoaggiunti

Def. $f \in \mathrm{End}(V)$ è *autoaggiunto* o *simmetrico* se $\forall v, w \in V$ si ha

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Teor. *Sia $\mathcal{B} = (b_1, \dots, b_n)$ base ortonormale per V.*
Allora $f \in \mathrm{End}(V)$ autoaggiunto $\Leftrightarrow M_{\mathcal{B}}^{\mathcal{B}}(f)$ simmetrica.

Dim. $A := M_{\mathcal{B}}^{\mathcal{B}}(f)$. $\forall v = X^{\mathcal{B}}, w = Y^{\mathcal{B}} \in V$, $X, Y \in \mathbb{R}^n$

$$\langle f(v), w \rangle = \langle AX, Y \rangle_{\mathbb{R}^n} = {}^t(AX)Y = {}^tX\,{}^tAY$$
$$\langle v, f(w) \rangle = \langle X, AY \rangle_{\mathbb{R}^n} = {}^tXAY$$

quindi $\langle f(v), w \rangle = \langle v, f(w) \rangle \Leftrightarrow {}^tX\,{}^tAY = {}^tXAY \Leftrightarrow {}^tA = A$. □

Cor. *Data $A \in M_n(\mathbb{R})$, $L_A \colon \mathbb{R}^n \to \mathbb{R}^n$ autoaggiunto rispetto al prodotto scalare canonico $\Leftrightarrow {}^tA = A$.*

Dim. $\mathcal{E}_n = (e_1, \dots, e_n)$ ortonormale e $M_{\mathcal{E}_n}^{\mathcal{E}_n}(L_A) = A$. □

Esempio. Consideriamo $\mathbb{R}^2$ col prodotto scalare canonico.
$L_A \colon \mathbb{R}^2 \to \mathbb{R}^2$ con

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

autoaggiunto perché A simmetrica.

$L_B \colon \mathbb{R}^2 \to \mathbb{R}^2$ con

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

non è autoaggiunto.

Prop. *$f \in \mathrm{End}(V)$ autoaggiunto, $\lambda \neq \mu$ autovalori $\Rightarrow \mathrm{Aut}(\lambda) \perp \mathrm{Aut}(\mu)$.*

Dim. $\forall v \in \mathrm{Aut}(\lambda)$ e $\forall w \in \mathrm{Aut}(\mu)$ facciamo vedere che $v \perp w$

$$\lambda\langle v, w \rangle = \langle \lambda v, w \rangle = \langle f(v), w \rangle = \langle v, f(w) \rangle = \mu\langle v, w \rangle$$
$$(\lambda - \mu)\langle v, w \rangle = 0 \ \Rightarrow \ \langle v, w \rangle = 0.$$ □

Oss. Per il Teorema Fondamentale dell'Algebra, ogni $A \in M_n(\mathbb{C})$ ammette n autovalori contati con la loro molteplicità algebrica.

Lem. *Ogni matrice simmetrica reale ha tutti gli autovalori reali.*

Cor. *Se* $\dim V < \infty$, *ogni* $f \in \mathrm{End}(V)$ *autoaggiunto ammette autovalori.*

Teorema spettrale. *Se* $\dim V < \infty$, *ogni* $f \in \mathrm{End}(V)$ *autoaggiunto ammette una base ortonormale diagonalizzante.*

Dim. Per induzione su $n = \dim V$.

Base dell'induzione: $n = 1$ $v \in V - \{0_V\} \rightsquigarrow u = \dfrac{v}{\|v\|}$ base ortonormale diagonalizzante.

Ipotesi induttiva Supponiamo vero l'enunciato in dimensione $n - 1 \geqslant 1$ e dimostriamolo in dimensione n.
$\exists\, \lambda \in \mathbb{R}$ autovalore di $f \rightsquigarrow v \in V - \{0_V\}$ autovettore $\rightsquigarrow U := \mathrm{span}(v)^\perp$ complemento ortogonale di $\mathrm{span}(v) \Rightarrow \dim U = n - 1$.
$\forall\, u \in U \Rightarrow \langle f(u), v\rangle = \langle u, f(v)\rangle = \lambda\langle u, v\rangle = 0 \Rightarrow f(u) \in U$.
Ipotesi induttiva $\rightsquigarrow (u_1, \ldots, u_{n-1})$ base ortonormale per U che diagonalizza la restrizione $f|_U : U \to U$. Posto $u_n = \dfrac{v}{\|v\|} \Rightarrow (u_1, \ldots, u_n)$ base ortonormale che diagonalizza $f : V \to V$. □

In pratica si procede nel modo seguente: si sceglie una base ortonormale $\mathcal{B} = (b_1, \ldots, b_n)$ per V. Si calcola la matrice simmetrica reale $M_{\mathcal{B}}^{\mathcal{B}}(f)$ e si procede con la diagonalizzazione, calcolando gli autovalori distinti $\lambda_1, \ldots, \lambda_k$ e una base ortonormale $\mathcal{C}_i$ per ciascun autospazio $\mathrm{Aut}(\lambda_i)$ (se $m_g(\lambda_i) > 1$ occorre applicare Gram-Schmidt e normalizzazione ad una base di $\mathrm{Aut}(\lambda_i)$ se questa non è ortonormale). $\mathcal{C} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_k$ è base ortonormale diagonalizzante perché gli autospazi sono tra loro ortogonali.

Cor. $\forall\, A \in M_n(\mathbb{R})$ *simmetrica* $\Rightarrow \exists\, P \in \mathrm{O}(n)$ *t.c.* $D = {}^tPAP$ *diagonale.*

Dim. $L_A : \mathbb{R}^n \to \mathbb{R}^n$ autoaggiunto $\Rightarrow \exists\, \mathcal{C} = (c_1, \ldots, c_n)$ base ortonormale diagonalizzante $\rightsquigarrow P = M_{\mathcal{C}}^{\mathcal{E}_n}(\mathrm{id}_{\mathbb{R}^n}) = (c_1 \cdots c_n) \in \mathrm{O}(n) \Rightarrow P^{-1} = {}^tP$ e $D = P^{-1}AP = {}^tPAP$ diagonale. □

Oss. A meno di cambiare segno a c_1 possiamo assumere $P \in \mathrm{SO}(n)$.

Sottospazi affini

V spazio vettoriale reale o complesso ($\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$).

Def. La *traslazione di vettore* $v \in V$ è la funzione

$$t_v : V \to V$$
$$t_v(u) = u + v.$$

Oss. $t_v(0_V) = v$.

Prop. *Valgono le seguenti proprietà:*
$t_{0_V} = \mathrm{id}_V$
$t_v \circ t_w = t_{v+w}$
$t_v^{-1} = t_{-v} \Rightarrow t_v$ *biiettiva.*

Dato $A \subset V$ e $v \in V$ definiamo il traslato

$$A + v \stackrel{\text{def}}{=} t_v(A) = \{a + v \mid a \in A\} \subset V.$$

Def. Un sottoinsieme $A \subset V$ è detto *sottospazio affine di* V se $\exists\, A_0 \subset V$ sottospazio vettoriale e $\exists\, v \in V$ t.c. $A = A_0 + v$. A_0 è detto *giacitura* di A. Poniamo $\dim A \stackrel{\text{def}}{=} \dim A_0$. Gli elementi di A si chiamano *punti*.
In particolare V stesso è sottospazio affine e si chiama *spazio affine*.

Oss. I sottospazi affini sono non vuoti.

$\underline{\dim A = 0}$ *Punto*.
$\underline{\dim A = 1}$ *Retta affine*.
$\underline{\dim A = 2}$ *Piano affine*.

Oss. I sottospazi affini sono i traslati dei sottospazi vettoriali.
I sottospazi vettoriali sono anche sottospazi affini. Un sottospazio affine $A \subset V$ è vettoriale $\Leftrightarrow 0_V \in A$.

Oss. Ogni punto $a \in A = A_0 + v$ si può scrivere come $a = a_0 + v$ per un unico $a_0 \in A_0$.

Dati due punti qualsiasi $a = a_0 + v$ e $b = b_0 + v$ di un sottospazio affine $A \subset V$ con giacitura $A_0 \subset V$ si ha

$$b - a = b_0 - a_0 \in A_0.$$

Dato un vettore $u \in A_0$ si ha

$$a + u = a_0 + u + v \in A$$

perché $a_0 + u \in A_0$. Dunque la differenza di due punti di A è un vettore di A_0, e la somma di un punto di A e di un vettore di A_0 è un punto di A.

Def. $b - a \in A_0$ è detto *vettore* ab.

Prop. *Sia $A_0 \subset V$ un sottospazio vettoriale e sia $a \in V$. Allora esiste un unico sottospazio affine $A \subset V$ passante per a con giacitura A_0.*

Dim. $\underline{\text{Esistenza}}$. Basta porre $A := A_0 + a \Rightarrow a = 0_V + a \in A$.

$\underline{\text{Unicità}}$. Se $A' = A_0 + v \subset V$ è un qualunque sottospazio affine passante per a con giacitura $A_0 \Rightarrow a = a_0 + v$ per un certo $a_0 \in A_0 \Rightarrow$
$A = A_0 + a = A_0 + a_0 + v = A'$ perché $A_0 + a_0 = A_0$. □

Teorema di struttura per sistemi lineari. *Sia $S: AX = B$ un sistema lineare compatibile, con $A \in M_{m,n}(\mathbb{K})$, $B \in \mathbb{K}^m$. Allora $\Sigma_S \subset \mathbb{K}^n$ è un sottospazio affine con giacitura lo spazio delle soluzioni Σ_{S_0} del sistema omogeneo associato $S_0: AX = 0_{\mathbb{K}^m}$.* $\dim \Sigma_S = \dim \Sigma_{S_0} = n - \mathrm{rg}(A)$.

Dim. $\Sigma_{S_0} = \ker L_A \subset \mathbb{K}^n$ sottospazio vettoriale, $\dim \Sigma_{S_0} = n - \mathrm{rg}\, A$.
Scegliamo una soluzione $s \in \Sigma_S$ (esiste perché S compatibile) $\Rightarrow As = B$.
$v \in \Sigma_S \Leftrightarrow Av = B = As \Leftrightarrow 0_{\mathbb{K}^m} = Av - As = A(v - s) \Leftrightarrow v - s \in \Sigma_{S_0}$
$\Leftrightarrow v \in s + \Sigma_{S_0} \Rightarrow \Sigma_S = s + \Sigma_{S_0}$. □

Oss.
S incompatibile $\Leftrightarrow \Sigma_S = \emptyset$ non è sottospazio affine.
Σ_S sottospazio vettoriale $\Leftrightarrow S$ omogeneo.

Equazioni di sottospazi affini

$A \subset V$ sottospazio affine passante per $a \in V$ e con giacitura A_0

$$A = A_0 + a.$$

Scegliamo una base $(u_1, \dots, u_k)$ per A_0, quindi $\dim A = \dim A_0 = k$.

Equazione vettoriale.

$$A : v = t_1 u_1 + \cdots + t_k u_k + a$$

è detta *equazione vettoriale di* A. Al variare dei *parametri* $t_1, \dots, t_k \in \mathbb{K}$ si ottengono tutti i punti di A. La giacitura A_0 ha equazione vettoriale

$$A_0 : v = t_1 u_1 + \cdots + t_k u_k$$

Equazione parametrica. $\mathcal{B} = (b_1, \dots, b_n)$ base per V.
L'equazione vettoriale di A si scrive in coordinate

$$A : X = t_1 C_1 + \cdots + t_k C_k + Q$$

dove X è il vettore delle coordinate di v, C_j di u_j e Q di a, rispetto a $\mathcal{B}$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \; C_j = \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix}, \; Q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$$

$$A : \begin{cases} x_1 = c_{11} t_1 + \cdots + c_{1k} t_k + q_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n1} t_1 + \cdots + c_{nk} t_k + q_n \end{cases}$$

sono dette *equazioni parametriche* di A. A_0 ha equazioni parametriche

$$A_0 : \begin{cases} x_1 = c_{11} t_1 + \cdots + c_{1k} t_k \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n1} t_1 + \cdots + c_{nk} t_k \end{cases}$$

Equazione cartesiana. Eliminando i parametri $t_1, \dots, t_k$ dalle equazioni parametriche si ottengono le *equazioni cartesiane*

$$A : \begin{cases} \alpha_{11} x_1 + \cdots + \alpha_{1n} x_n = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m1} x_1 + \cdots + \alpha_{mn} x_n = \beta_m \end{cases}$$

A è l'insieme dei punti le cui coordinate soddisfano questo sistema lineare. A_0 ha equazioni cartesiane

$$A_0 : \begin{cases} \alpha_{11} x_1 + \cdots + \alpha_{1n} x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m1} x_1 + \cdots + \alpha_{mn} x_n = 0 \end{cases}$$

Pertanto otteniamo l'inverso del Teorema di struttura: *ogni sottospazio affine di* $\mathbb{K}^n$ *è lo spazio delle soluzioni di un sistema lineare compatibile.*

Il passaggio tra i vari tipi di equazioni è analogo a quanto visto nel caso dei sottospazi vettoriali.

Posizione reciproca di due sottospazi affini

Intersezione di sottospazi vettoriali. $U, W \subset V$ sottospazi vettoriali $\Rightarrow$

$$U \cap W = \{v \in V \mid v \in U \text{ e } v \in W\} \subset V$$

sottospazio vettoriale. Infatti:
$0_V \in U \cap W$ perché 0_V appartiene sia a U che a W;
$\forall \alpha_1, \alpha_2 \in \mathbb{K}$, $\forall v_1, v_2 \in U \cap W \Rightarrow \alpha_1 v_1 + \alpha_2 v_2 \in U \cap W$.

In termini di equazioni cartesiane abbiamo

$$A : FX = 0 \qquad B : GX = 0 \qquad A \cap B : \begin{cases} FX = 0 \\ GX = 0 \end{cases}$$

dove F e G sono matrici.

Intersezione di sottospazi affini. $A, B \subset V$ sottospazi affini $\Rightarrow$

$$A \cap B = \{v \in V \mid v \in A \text{ e } v \in B\} \subset V$$

sottospazio affine oppure vuoto. Infatti se $A \cap B \neq \emptyset$ scegliamo un punto $a \in A \cap B$ e abbiamo

$$A \cap B = (A_0 \cap B_0) + a.$$

Quindi la giacitura di $A \cap B$ è l'intersezione delle rispettive giaciture

$$(A \cap B)_0 = A_0 \cap B_0.$$

In termini di equazioni cartesiane abbiamo

$$A : FX = P \qquad B : GX = Q \qquad A \cap B : \begin{cases} FX = P \\ GX = Q \end{cases}$$

Def. Due sottospazi affini A e $B \subset V$ con giaciture risp. A_0 e B_0 sono

1) *incidenti* se $A \cap B \neq \emptyset$;

2) *disgiunti* se $A \cap B = \emptyset$;

3) *paralleli* se $A_0 \subset B_0$ oppure $B_0 \subset A_0$, e scriviamo $A \parallel B$;

4) *sghembi* se $A \cap B = \emptyset$ e $A \nparallel B$ (A e B disgiunti e non paralleli).

Oss. Se $\dim A = \dim B < \infty$ allora $A \parallel B \Leftrightarrow A_0 = B_0$.

Teor ("Postulato delle parallele")**.** *Sia $A \subset V$ un sottospazio affine di dimensione finita e sia $b \in V$ un punto qualsiasi. Allora $\exists!$ $B \subset V$ sottospazio affine passante per b t.c. $B \parallel A$ e $\dim B = \dim A$.*

Dim. L'unico sottospazio affine che soddisfa la tesi è $B = A_0 + b$. □

Per capire la posizione reciproca di due sottospazi affini si scrivono le equazioni cartesiane e si studia il sistema formato da tutte le equazioni messe insieme. Si applica Rouché-Capelli per capire la compatibilità e la dimensione dell'intersezione (se non vuota).

Nel seguito indichiamo con A la matrice dei coefficienti del sistema formato da tutte le equazioni coinvolte e con $\tilde{A}$ la matrice completa. Si osservi che in generale

$$\operatorname{rg}\tilde{A} = \begin{cases} \operatorname{rg} A \\ \operatorname{rg} A + 1 \quad \text{(se i ranghi sono diversi)} \end{cases}$$

perché $\tilde{A}$ si ottiene da A aggiungendo una colonna.

In $\mathbb{R}^2$

Due rette $r, s \subset \mathbb{R}^2$.

$$\boxed{\begin{array}{l} r: \ ax + by = c \\ s: \ dx + ey = f \end{array}}$$

$$\boxed{\begin{array}{l} r = s \Leftrightarrow \operatorname{rg}\tilde{A} = 1 \\ r \parallel s \text{ e } r \cap s = \emptyset \Leftrightarrow \operatorname{rg} A = 1 \text{ e } \operatorname{rg}\tilde{A} = 2 \\ r \cap s = \text{punto} \Leftrightarrow \operatorname{rg} A = 2. \end{array}}$$

In $\mathbb{R}^3$

Due piani $p, q \subset \mathbb{R}^3$.

$$\boxed{\begin{array}{l} p: \ ax + by + cz = d \\ q: \ ex + fy + gz = h \end{array}}$$

$$\boxed{\begin{array}{l} p = q \Leftrightarrow \operatorname{rg}\tilde{A} = 1 \\ p \parallel q \text{ e } p \cap q = \emptyset \Leftrightarrow \operatorname{rg} A = 1 \text{ e } \operatorname{rg}\tilde{A} = 2 \\ p \cap q = \text{retta} \Leftrightarrow \operatorname{rg} A = 2. \end{array}}$$

Retta e piano $r, p \subset \mathbb{R}^3$.

$$\boxed{\begin{array}{l} r: \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases} \\ p: ex + fy + gz = h \end{array}}$$

$$\boxed{\begin{array}{l} r \subset p \Leftrightarrow \operatorname{rg}\tilde{A} = 2 \\ r \parallel p \text{ e } r \cap L = \emptyset \Leftrightarrow \operatorname{rg} A = 2 \text{ e } \operatorname{rg}\tilde{A} = 3 \\ r \cap p = \text{punto} \Leftrightarrow \operatorname{rg} A = 3. \end{array}}$$

Due rette $r, s \subset \mathbb{R}^3$.

$$\boxed{\begin{array}{l} r: \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases} \\ s: \begin{cases} e_1x + f_1y + g_1z = h_1 \\ e_2x + f_2y + g_2z = h_2 \end{cases} \end{array}}$$

$$\boxed{\begin{array}{l} r = s \Leftrightarrow \operatorname{rg}\tilde{A} = 2 \\ r \parallel s \text{ e } r \cap s = \emptyset \Leftrightarrow \operatorname{rg} A = 2 \text{ e } \operatorname{rg}\tilde{A} = 3 \\ r \cap s = \text{punto} \Leftrightarrow \operatorname{rg} A = \operatorname{rg}\tilde{A} = 3 \\ r \text{ e } s \text{ sghembe} \Leftrightarrow \operatorname{rg}\tilde{A} = 4. \end{array}}$$

Sottospazio affine generato da un sottoinsieme

$P_0, \dots, P_k \in V \rightsquigarrow A(P_0, \dots, P_k) \stackrel{\text{def}}{=} \text{span}(P_1 - P_0, \dots, P_k - P_0) + P_0$

$P_1 - P_0, \dots, P_k - P_0 \in V$ lin. indip. $\Leftrightarrow \dim A(P_0, \dots, P_k) = k$.
Equazione vettoriale di $A(P_0, \dots, P_k)$

$$v = t_1(P_1 - P_0) + \cdots + t_k(P_k - P_0) + P_0$$

Oss. La giacitura di $A(P_0, \dots, P_k)$ è $\text{span}(P_1 - P_0, \dots, P_k - P_0)$.

Oss. $P_i \in A(P_0, \dots, P_k)$, $\forall\, i = 0, \dots, k$. Infatti $P_i = (P_i - P_0) + P_0$.

Oss. $A(P_0, \dots, P_k)$ non dipende dall'ordine dei punti $P_0, \dots, P_k$. Infatti

$$A(P_0, \dots, P_k) = \{t_0P_0 + \cdots + t_kP_k \mid t_i \in \mathbb{K}\ \forall i \text{ e } t_0 + \cdots + t_k = 1\}$$

ottenuta dall'equazione vettoriale ponendo $t_0 = 1 - t_1 - \cdots - t_k$

Retta per due punti distinti. $P_0 \neq P_1 \in V \Leftrightarrow \dim A(P_0, P_1) = 1$

$$A(P_0, P_1) = \text{span}(P_1 - P_0) + P_0$$
$$A(P_0, P_1) \colon\ v = t(P_1 - P_0) + P_0$$

Oss. $P_0, \dots, P_k \in V$ allineati $\Leftrightarrow$ $P_1 - P_0, \dots, P_k - P_0 \in \mathbb{K}^n$ proporzionali.

Piano per tre punti non allineati. $P_0, P_1, P_2 \in V$ non allineati $\Leftrightarrow$ $P_1 - P_0, P_2 - P_0$ lin. indip. $\Leftrightarrow \dim A(P_0, P_1, P_2) = 2$

$$A(P_0, P_1, P_2) \colon\ v = t_1(P_1 - P_0) + t_2(P_2 - P_0) + P_0$$

Affinità

Def. $f : V \to V$ è detta *affinità* se $\exists\ \varphi : V \to V$ isomorfismo t.c.

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

φ si chiama *parte lineare* di f.

Oss. $f(v) = \varphi(v) + f(0_V)$, $\forall\, v \in V$, ottenuta ponendo $w = 0_V$.

$f : \mathbb{K}^n \to \mathbb{K}^n$ affinità $\Leftrightarrow \exists\, G \in \text{GL}_n(\mathbb{K})$ e $\exists\, a \in \mathbb{K}^n$ t.c.

$$f(X) = GX + a, \quad \forall X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

Oss. $a = 0_{\mathbb{K}^n} \Rightarrow f$ isomorfismo. $G = I_n \Rightarrow f = t_a$ traslazione.

$f : \mathbb{K}^n \to \mathbb{K}^n$ affinità $\Rightarrow f^{-1}$ affinità ottenuta risolvendo $Y = GX + a$

$$X = G^{-1}Y - G^{-1}a$$
$$f^{-1}(Y) = G^{-1}Y - G^{-1}a.$$

Le affinità mandano sottospazi affini in sottospazi affini della stessa dimensione e preservano la posizione relativa, in particolare il parallelismo.

Elementi di Geometria Euclidea

V spazio vettoriale Euclideo con prodotto scalare $\langle\ ,\ \rangle$.

Distanza in coordinate. $\mathcal{B} = (b_1, \dots, b_n)$ base *ortonormale* per V.

$$u \text{ di coordinate } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ v \text{ di coordinate } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$d(u, v) = d(X, Y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}}.$$

Vale anche in $\mathbb{R}^n$ perché la base canonica è ortonormale.

Esempio. In $\mathbb{R}^2$ col prodotto scalare canonico abbiamo

$$d\left(\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right) = \sqrt{(1-3)^2 + (-2-4)^2} = 2\sqrt{10}.$$

Oss. La formula per la distanza generalizza il Teorema di Pitagora.

Def. Un vettore di norma 1 è detto *versore*.

Def. $u \in V$ è *normale* ad un sottospazio affine $L \subset V$ se $u \in L_0^{\perp}$ e $u \neq 0_V$, dove $L_0 \subset V$ è la giacitura di L. u è *parallelo* a L se $u \in L_0$.

Oss. u versore normale a L se $u \in L_0^{\perp}$ e $\|u\| = 1$.

Esempio. Dato il piano $L \subset \mathbb{R}^3$ di equazione

$$L:\ 2x - y - 3z = 1$$

un versore normale è

$$u = \frac{1}{\sqrt{14}} \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}$$

ottenuto normalizzando il vettore dei coefficienti.

Distanza tra sottoinsiemi. La *distanza* tra A e $B \subset V$, $A, B \neq \emptyset$, è

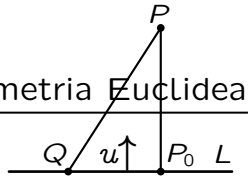
$$d(A, B) \stackrel{\text{def}}{=} \inf\{d(a, b) \mid a \in A, b \in B\}$$

Oss. $A \cap B \neq \emptyset \Rightarrow d(A, B) = 0$. Vale viceversa se A, B sottospazi affini.

Distanza tra punto e iperpiano affine. In $\mathbb{R}^n$ consideriamo l'iperpiano

$$L:\ a_1x_1 + \cdots + a_nx_n + b = 0, \quad A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \neq 0_{\mathbb{R}^n}$$

e il punto $P = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{R}^n$. Vogliamo calcolare $d(P, L)$.



Un versore normale a L è $u = \dfrac{A}{\|A\|}$.

Il punto $P_0 \in L$ di minima distanza da P si ottiene come intersezione tra L e la retta ortogonale a L passante per P. Sia $Q \in L \Rightarrow \langle Q, A\rangle + b = 0$.

$$d(P,L) = d(P,P_0) = |\langle P - P_0, u\rangle| = |\langle P - Q + Q - P_0, u\rangle| =$$
$$= |\langle P - Q, u\rangle| = \frac{|\langle P, A\rangle - \langle Q, A\rangle|}{\|A\|} =$$
$$= \frac{|\langle P, A\rangle + b|}{\|A\|} = \frac{|a_1 p_1 + \cdots + a_n p_n + b|}{\sqrt{a_1^2 + \cdots + a_n^2}}$$

dove abbiamo usato il fatto che $Q - P_0 \in L_0 \Rightarrow \langle Q - P_0, u\rangle = 0$.

Esempio. Calcoliamo in $\mathbb{R}^2$ la distanza tra la retta r e il punto P con

$$r:\ 2x - y + 2 = 0, \quad P = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad d(P,r) = \frac{|2 - 3 + 2|}{\sqrt{2^2 + (-1)^2}} = \frac{1}{\sqrt{5}}.$$

Angolo tra due rette in $\mathbb{R}^n$. $r, s \subset \mathbb{R}^n$ rette con *vettori direzionali* risp. v_r, v_s, ossia $r_0 = \text{span}(v_r)$, $s_0 = \text{span}(v_s)$. L'*angolo* tra r e s è

$$\widehat{rs} \stackrel{\text{def}}{=} \arccos \frac{|\langle v_r, v_s\rangle|}{\|v_r\|\|v_s\|} \in \left[0, \frac{\pi}{2}\right]$$

Oss. Il valore assoluto nel numeratore è dovuto al fatto che i vettori direzionali non sono univocamente determinati: se v_r è un vettore direzionale di r, anche $-v_r$ lo è.

Def. r e s sono *perpendicolari* se $\widehat{rs} = \dfrac{\pi}{2}$.

Angolo tra due iperpiani in $\mathbb{R}^n$. $L, T \subset \mathbb{R}^n$ iperpiani con *vettori normali* risp. u_L, u_T, ossia $L_0^\perp = \text{span}(u_L)$, $T_0^\perp = \text{span}(u_T)$. L'*angolo* tra L e T è

$$\widehat{LT} \stackrel{\text{def}}{=} \arccos \frac{|\langle u_L, u_T\rangle|}{\|u_L\|\|u_T\|} \in \left[0, \frac{\pi}{2}\right]$$

Def. L e T sono *perpendicolari* se $\widehat{LT} = \dfrac{\pi}{2}$.

Isometrie

V spazio vettoriale Euclideo.

Def. $f : V \to V$ *isometria* se $d(f(u), f(v)) = d(u, v)$, $\forall\, u, v \in V$.

Le isometrie di $\mathbb{R}^n$ sono affinità del tipo

$$f(X) = AX + B$$

con $A \in O(n)$ e $B \in \mathbb{R}^n$. f è detta *isometria diretta* se $A \in SO(n)$.
Verifichiamo che f è isometria:

$$\begin{aligned} d(f(X), f(Y)) &= \|(AX + B) - (AY + B)\| = \|A(X - Y)\| = \\ &= \sqrt{{}^t(A(X - Y))A(X - Y)} = \sqrt{{}^t(X - Y)\,{}^tAA(X - Y)} = \\ &= \|X - Y\| = d(X, Y). \end{aligned}$$

Se $A = I_n$, $f(X) = X + B$ traslazione $\Rightarrow$ isometria diretta.

Oss.

1) $\mathrm{id}_{\mathbb{R}^n}$ isometria;

2) $f, g : \mathbb{R}^n \to \mathbb{R}^n$ isometrie $\Rightarrow f \circ g$ isometria;

3) f isometria $\Rightarrow f^{-1}$ isometria.

4) isometrie preservano distanze e angoli.

Rotazioni del piano. *Matrice di rotazione di angolo* $\theta \in \mathbb{R}$

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2)$$

Oss. Si verificano facilmente le seguenti relazioni.
$R_0 = R_{2k\pi} = I_2$, $\forall\, k \in \mathbb{Z}$.
$R_\theta R_\varphi = R_{\theta+\varphi}$ (usando formule di addizione per sin e cos).
$R_\theta^{-1} = R_{-\theta}$ (usando le due precedenti).
$\det R_\theta = 1$.

Def. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(X) = R_\theta X + B$ è detta *rototraslazione*.

Se $\theta \neq 2k\pi$, $\forall\, k \in \mathbb{Z} \Rightarrow \exists!\, Q \in \mathbb{R}^2$ t.c. $f(Q) = Q$ (*punto fisso*), soluzione del sistema lineare

$$R_\theta X + B = X \;\Leftrightarrow\; (R_\theta - I_2)X = -B.$$

Ossia una rototraslazione di angolo non nullo è una rotazione attorno a Q (*centro di rotazione*). Infatti se $\theta \neq 2k\pi$, $\forall\, k \in \mathbb{Z}$

$$\det(R_\theta - I_2) = 2(1 - \cos\theta) \neq 0$$

$$Q = -(R_\theta - I_2)^{-1} B.$$

Prodotto vettoriale in $\mathbb{R}^3$

Dati due vettori

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \; Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$$

definiamo il *prodotto vettore* o *prodotto vettoriale* come il determinante

$$X \times Y \stackrel{\text{def}}{=} \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

È un determinante formale con la base canonica in prima riga. Si ha:

$$\begin{aligned} X \times Y &= \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3 = \\ &= (x_2y_3 - x_3y_2)e_1 - (x_1y_3 - x_3y_1)e_2 + (x_1y_2 - x_2y_1)e_3. \end{aligned}$$

Si può dimostrare l'identità seguente, sviluppando i calcoli a primo e secondo membro:

$$\|X \times Y\|^2 = \|X\|^2\|Y\|^2 - \langle X, Y\rangle^2$$

$$\|X \times Y\| = \|X\|\|Y\| \sin \widehat{XY}$$

$X \times Y = 0_{\mathbb{R}^3} \Leftrightarrow X$ e Y proporzionali.
X, Y linearmente indipendenti $\Rightarrow X \times Y \in \text{span}(X, Y)^\perp$.
Il verso si ricava con la "regola della mano destra": pollice verso X, indice verso Y e medio verso $X \times Y$ (come per i vettori della base canonica).