Corso di Laurea specialistica in Ingegneria Edile – 2007–2008
Matematica 3

Giulia Furioli

11 gennaio 2008

Questi sono gli appunti delle lezioni di Matematica 3 tenute durante l'a.a. 2007–2008. Non intendono in alcun modo essere esaustivi e non sono sufficienti alla preparazione dell'esame, per il quale è necessario consultare i libri di testo che verranno indicati per ogni argomento.

Indice

Capitolo 1

Algebra lineare

Testo consigliato: M. Bramanti, C. D. Pagani e S. Salsa, MATEMATICA, Calcolo infinitesimale e algebra lineare, Zanichelli, nuova edizione 2004

1.1 Spazi vettoriali

Ricordiamo innanzi tutto la struttura dell'insieme dei vettori a tre componenti reali $\mathbb{R}^3$, la cui generalizzazione darà luogo alla definizione di spazio vettoriale.

Lo spazio dei vettori dello spazio $\mathbb{R}^3$ è l'insieme

$$\mathbb{R}^3 = \{(x, y, z), x, y, z \in \mathbb{R}\}$$

dotato delle due leggi seguenti:

1) **somma:**

$$\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$
$$(x, y, z) + (x', y', z') := (x + x', y + y', z + z')$$

2) **prodotto per uno scalare:**

$$\mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$
$$\lambda \cdot (x, y, z) := (\lambda x, \lambda y, \lambda z)$$

(Attenzione: non si tratta di un PRODOTTO SCALARE, di cui parleremo in seguito, ma di un prodotto PER uno scalare, cioè del prodotto tra un VETTORE (un elemento di $\mathbb{R}^3$) e un numero reale, che viene definito uno SCALARE per distinguerlo da un vettore).

Quali proprietà sono verificate da queste leggi?

- Per la somma:

  (a) proprietà commutativa: per ogni v, $w \in \mathbb{R}^3$ si ha $v + w = w + v$;

  (b) proprietà associativa: per ogni v, w e $t \in \mathbb{R}^3$ si ha $(v + w) + t = v + (w + t)$;

  (c) esistenza dell'elemento neutro: esiste un elemento $\underline{0} \in \mathbb{R}^3$ tale che per ogni $v \in \mathbb{R}^3$ si ha $v + \underline{0} = v$ (tale elemento è il vettore $(0, 0, 0)$);

  (d) esistenza dell'opposto per ogni elemento: per ogni $v \in \mathbb{R}^3$ esiste un elemento $-v \in \mathbb{R}^3$ tale che $v + (-v) = \underline{0}$ (se $v = (x, y, z)$ allora $-v = (-x, -y, -z)$).

  La verifica di queste proprietà è immediata e proviene dalle analoghe proprietà che sono verificate in $\mathbb{R}$.

- Per il prodotto per uno scalare:

(a') per ogni $v \in \mathbb{R}^3$ si ha $1 \cdot v = v$;

(b') proprietà associativa: per ogni $v \in \mathbb{R}^3$ e per ogni $\lambda, \mu \in \mathbb{R}$ si ha $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$;

(c') proprietà distributiva del prodotto per uno scalare rispetto alla somma di vettori: per ogni $v, w \in \mathbb{R}^3$ e per ogni $\lambda \in \mathbb{R}$ si ha $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$

(d') proprietà distributiva del prodotto per uno scalare rispetto alla somma di scalari: per ogni $\lambda, \mu \in \mathbb{R}$ e per ogni $v \in \mathbb{R}^3$ si ha $(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$;

Anche la verifica di queste proprietà è immediata.

D'ora in poi non scriveremo più $\lambda \cdot v$ per indicare il prodotto per uno scalare, ma più semplicemente λv.

Ci chiediamo ora: esistono altri insiemi che si possono dotare di due leggi che verificano le stesse proprietà? Sì, si tratta degli spazi vettoriali.

Definizione 1

Sia V un insieme munito di due leggi di somma e prodotto per uno scalare:

$$+ : V \times V \longrightarrow V$$
$$\cdot : \mathbb{R} \times V \longrightarrow V$$

che godono delle proprietà (a), (b), (c) e (d) per la somma, (a'), (b'), (c') e (d') per il prodotto per uno scalare. La terna $(V, +, \cdot)$ verrà chiamata uno **spazio vettoriale reale**.

Se è definito un prodotto per uno scalare $\cdot : \mathbb{C} \times V \longrightarrow V$ con le stesse proprietà, la terna $(V, +, \cdot)$ sarà detta uno **spazio vettoriale complesso**.

La maggior parte degli spazi vettoriali che tratteremo in seguito saranno reali.

Osservazioni:

1. In uno spazio vettoriale l'elemento neutro rispetto alla somma è unico: infatti, se per assurdo esistessero due elementi $\underline{0} \neq \tilde{\underline{0}}$ entrambi neutri, allora si avrebbe
$$\underline{0} + \tilde{\underline{0}} = \underline{0}$$
ma anche
$$\underline{0} + \tilde{\underline{0}} = \tilde{\underline{0}}$$
da cui
$$\underline{0} = \tilde{\underline{0}}.$$
Inoltre, per ogni elemento, l'opposto è unico: infatti, dato un elemento $v \in V$, se esistessero due opposti $(-v) \neq w$ si avrebbe
$$v + (-v) = \underline{0}$$
e quindi
$$w + (v + (-v)) = w + \underline{0} = w$$
ma allora, poiché abbiamo supposto che anche w sia un opposto di v si avrebbe $w + v = \underline{0}$ e quindi (grazie alla proprietà associativa della somma)
$$(w + v) + (-v) = w \iff -v = w.$$

2. Valgono le proprietà:
$$\lambda\underline{0} = \underline{0}, \quad \forall \lambda \in \mathbb{R},$$
$$0v = \underline{0}, \quad \forall v \in V$$
(e inoltre $\lambda v = \underline{0} \iff \lambda = 0$ oppure $v = \underline{0}$).

Osservazione importante: In uno spazio vettoriale **non** è in generale definito un prodotto tra vettori.

Esempi:

1. Lo spazio dei vettori a n componenti reali $\mathbb{R}^n$, munito dell'usuale somma componente per componente e prodotto per uno scalare componente per componente, è uno spazio vettoriale reale. In particolare, $\mathbb{R}$ è uno spazio vettoriale. Lo spazio dei vettori a n componenti complesse $\mathbb{C}^n$, munito dell'usuale somma componente per componente e prodotto per uno scalare complesso componente per componente, è uno spazio vettoriale complesso.

2. Lo spazio delle matrici a m righe e n colonne a coefficienti reali $M_{m\times n}(\mathbb{R})$, dotato dell'usuale somma e prodotto per uno scalare componente per componente è uno spazio vettoriale reale. L'elemento neutro è la matrice con tutti i coefficienti nulli e per ogni matrice
$$M = [a_{i,j}]_{i=1,\dots,m,\ j=1,\dots,n}$$
la matrice opposta è la matrice
$$-M = [-a_{i,j}]_{i=1,\dots,m,\ j=1,\dots,n}.$$

3. È uno spazio vettoriale reale lo spazio $\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ di tutte le funzioni reali di variabile reale munito di somma e prodotto per uno scalare usuali: date due funzioni f e g e un numero reale λ, si pone
$$\begin{aligned}(f+g)(x) &:= f(x) + g(x), \quad \forall x \in \mathbb{R}\\ (\lambda f)(x) &:= \lambda f(x), \quad \forall x \in \mathbb{R}\end{aligned}$$

4. Consideriamo l'insieme dei polinomi di grado al più 3 a coefficienti reali:
$$\mathbb{R}_3[X] = \{aX^3 + bX^2 + cX + d, \quad a, b, c, d \in \mathbb{R}\}$$
dotato delle usuali leggi di somma e prodotto per uno scalare tra funzioni definite come prima: per ogni P_1, $P_2 \in \mathbb{R}_3[X]$, per ogni $\lambda \in \mathbb{R}$ poniamo
$$\begin{aligned}(P_1 + P_2)(x) &:= P_1(x) + P_2(x), \quad \forall x \in \mathbb{R}\\ (\lambda P_1)(x) &:= \lambda P_1(x), \quad \forall x \in \mathbb{R}\end{aligned}$$
Si tratta di uno spazio vettoriale reale. Infatti, le proprietà (a) ... (d), (a') ... (d') di somma e prodotto per uno scalare sono banalmente verificate perché sono verificate in $\mathbb{R}$; l'unica cosa da mostrare è che la somma e il prodotto per uno scalare sono ben definite, cioè che per ogni P_1, $P_2 \in \mathbb{R}_3[X]$, per ogni $\lambda \in \mathbb{R}$ si ha $P_1 + P_2 \in \mathbb{R}_3[X]$ e $\lambda P_1 \in \mathbb{R}_3[X]$ (cioè, si tratta sicuramente di due funzioni ma siamo sicuri che si tratta ancora di due polinomi di grado al più 3?) Basta scrivere somma e prodotto secondo la definizione, ponendo $P_1 = aX^3 + bX^2 + cX + d$ e $P_2 = \alpha X^3 + \beta X^2 + \gamma X + \delta$ con $a, b, c, d, \alpha.\beta, \gamma$ e $\delta \in \mathbb{R}$. Si avrà, per ogni $x \in \mathbb{R}$:
$$\begin{aligned}&\big((aX^3 + bX^2 + cX + d) + (\alpha X^3 + \beta X^2 + \gamma X + \delta)\big)(x)\\ &= ax^3 + bx^2 + cx + d + \alpha x^3 + \beta x^2 + \gamma x + \delta\\ &= (a+\alpha)x^3 + (b+\beta)x^2 + (c+\gamma)x + (d+\delta)\\ &= \big((a+\alpha)X^3 + (b+\beta)X^2 + (c+\gamma)X + (d+\delta)\big)(x) \in \mathbb{R}_3[X]\end{aligned}$$
e per ogni $\lambda \in \mathbb{R}$:
$$\begin{aligned}&\big(\lambda(aX^3 + bX^2 + cX + d)\big)(x)\\ &= \lambda(ax^3 + bx^2 + cx + d)\\ &= (\lambda a)x^3 + (\lambda b)x^2 + (\lambda c)x + (\lambda d)\\ &= \big((\lambda a)X^3 + (\lambda b)X^2 + (\lambda c)X + (\lambda d)\big)(x) \in \mathbb{R}_3[X].\end{aligned}$$
Lo spazio dei polinomi di grado al più n a coefficienti reali si denota spesso con $\mathbb{R}_n[X]$.

5. Sono spazi vettoriali reali anche:

$$C(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R},\ f \text{ continua su } \mathbb{R}\}$$
$$C^1(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R},\ f \text{ derivabile su } \mathbb{R} \text{ e con derivata continua}\}$$
$$C^n(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R},\ f \text{ derivabile } n \text{ volte su } \mathbb{R} \text{ e con tutte le derivate continue}\}$$
$$C^\infty(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R},\ f \text{ derivabile infinite volte su } \mathbb{R} \text{ e con tutte le derivate continue}\}$$

6. Lo spazio delle funzioni $\underline{f} : \mathbb{R} \longrightarrow \mathbb{R}^2$ a 2 componenti è uno spazio vettoriale reale munito della somma e prodotto per uno scalare componente per componente: per $\underline{f} = (f_1, f_2)$, $\underline{g} = (g_1, g_2)$ si ha

$$(\underline{f} + \underline{g})(x) = [(f_1, f_2) + (g_1, g_2)](x) := (f_1(x) + g_1(x), f_2(x) + g_2(x)), \quad x \in \mathbb{R}$$
$$(\lambda \underline{f})(x) = [\lambda(f_1, f_2)](x) := (\lambda f_1(x), \lambda f_2(x)), \quad x \in \mathbb{R},\ \lambda \in \mathbb{R}.$$

   Analogamente, lo spazio delle funzioni $\underline{f} : \mathbb{R} \longrightarrow \mathbb{R}^n$ a n componenti è uno spazio vettoriale reale.

Sia ora $(V, +, \cdot)$ uno spazio vettoriale.

Definizione 2
Diremo che un sottoinsieme $W \subseteq V$ è un **sottospazio vettoriale** *di V se è a sua volta uno spazio vettoriale rispetto alle leggi di somma e prodotto per uno scalare definite in V.*

Osservazione: Di fatto, W è un sottospazio vettoriale di V se e solo se per ogni $w_1, w_2 \in W$ e per ogni $\lambda \in \mathbb{R}$ si ha $w_1 + w_2 \in W$ e $\lambda w_1 \in W$. Inoltre, **condizione necessaria (ma non sufficiente) perché W sia un sottospazio vettoriale è che $\underline{0} \in W$.**

Esempi:

1. In $\mathbb{R}^2$ tutti i possibili sottospazi vettoriali sono $\{(0,0)\}$, tutte le rette passanti per l'origine e $\mathbb{R}^2$ stesso (quest'ultimo è un sottospazio banale, nel senso che non è un sottoinsieme proprio).

2. Sempre in $\mathbb{R}^2$ la retta $r : \{(x, y) : y = 2x + 3\}$ non è un sottospazio vettoriale. Perché?

3. Nello spazio dei polinomi $\mathbb{R}_3[X] = \{aX^3 + bX^2 + cX + d, \quad a, b, c, d \in \mathbb{R}\}$, il sottoinsieme dei polinomi dispari è un sottospazio vettoriale: infatti la somma di due polinomi dispari è dispari e il prodotto di un polinomio dispari per un numero reale resta un polinomio dispari. (**Esercizio**: dimostrare che anche il sottoinsieme dei polinomi pari è sottospazio vettoriale).

4. Lo spazio $V = \{aX^3+bX^2+cX+d, \quad a, b, c, d \in \mathbb{R}, a \neq 0\}$ dei polinomi di grado esattamente 3 NON è un sottospazio vettoriale. Infatti non contiene l'elemento neutro rispetto alla somma di funzioni, che è il polinomio identicamente nullo.

Definizione 3
Si dice **combinazione lineare di n vettori** *ogni vettore $w \in V$ della forma*

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

ove $\alpha_i \in \mathbb{R}$ e $v_i \in V$ per $i = 1, \ldots, n$.

Esempi:

1. In $\mathbb{R}^2$ il vettore $w = 3(1, 0) + 4(0, 1) = (3, 4)$ è combinazione lineare di $(1, 0)$ e $(0, 1)$.

2. Nello spazio $\mathbb{R}_3[X] = \{aX^3 + bX^2 + cX + d, \quad a, b, c, d \in \mathbb{R}\}$ dei polinomi di grado al più 3, il polinomio $P = 3X^3 - 2$ è combinazione lineare di X^3 e 1.

Definizione 4
Dati n vettori $v_1, \dots, v_n$ in V, diremo che sono **linearmente indipendenti** *se l'unica combinazione lineare di $v_1, \dots, v_n$ che dà il vettore nullo è quella con tutti i coefficienti nulli, cioè:*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \underline{0} \iff \alpha_i = 0, \quad \forall i = 1, \dots, n.$$

Diremo che sono **linearmente dipendenti** *se ciò non accade, cioè se esistono n coefficienti $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ non tutti nulli tali che $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \underline{0}$.*

Osservazione: Se l'elemento neutro dello spazio vettoriale appartiene alla n-pla di vettori $v_1, \dots, v_n$, essi sono senz'altro dipendenti.

Esempi:

1. In $\mathbb{R}^2$ i vettori $v_1 = (1, 0)$ e $v_2 = (0, 1)$ sono linearmente indipendenti. Infatti

$$\alpha_1 v_1 + \alpha_2 v_2 = \underline{0} \iff \alpha_1(1, 0) + \alpha_2(0, 1) = (0, 0) \iff \begin{cases} \alpha_1 = 0 \\ \alpha_2 = 0 \end{cases}$$

2. Sempre in $\mathbb{R}^2$ i vettori $v_1 = (1, 2)$ e $v_2 = (3, 1)$ sono linearmente indipendenti. (**esercizio**)

3. Ancora in $\mathbb{R}^2$, i vettori $v_1 = (1, 0)$, $v_2 = (2, 1)$ e $v_3 = (3, 2)$ sono linearmente dipendenti. Infatti
$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \underline{0} \iff \alpha_1(1, 0) + \alpha_2(2, 1) + \alpha_3(3, 2) = (0, 0)$$
e quindi
$$\begin{cases} \alpha_1 + 2\alpha_2 + 3\alpha_3 = 0 \\ \alpha_2 + 2\alpha_3 = 0 \end{cases}$$
Questo sistema ammette sicuramente la soluzione $(0, 0, 0)$, ma anche ogni soluzione della forma $(\alpha_3, -2\alpha_3, \alpha_3)$ ove $\alpha_3 \in \mathbb{R}$ è arbitrario. Per esempio, una soluzione possibile è $(1, -2, 1)$.

4. Nello spazio $\mathbb{R}_1[X] = \{aX + b, \quad a, b \in \mathbb{R}\}$ dei polinomi di grado al più 1, i polinomi $P_1 = X$, $P_2 = 2X + 1$ e $P_3 = 3X + 2$ sono linearmente dipendenti. Infatti
$$\begin{aligned} \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 = \underline{0} &\iff \alpha_1 x + \alpha_2(2x + 1) + \alpha_3(3x + 2) = 0 \quad \forall x \in \mathbb{R} \\ &\iff (\alpha_1 + 2\alpha_2 + 3\alpha_3)x + (\alpha_2 + 2\alpha_3) = 0 \quad \forall x \in \mathbb{R}. \end{aligned}$$
Ora, osserviamo che ciò è possibile se e solo se entrambi i coefficienti sono nulli: infatti, se $\alpha_2 + 2\alpha_3 \neq 0$, allora avremmo che $(\alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3)(0) \neq 0$ e se $\alpha_1 + 2\alpha_2 + 3\alpha_3 \neq 0$ avremmo (ad esempio) che $(\alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3)(1) \neq 0$. Dunque, $\alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 = \underline{0}$ se e solo se α_1, α_2 e α_3 soddisfano il sistema dell'esempio precedente. Ma abbiamo visto che tale sistema ammette soluzioni non tutte nulle e quindi i due polinomi sono dipendenti.

5. Sempre in $\mathbb{R}_1[X]$, i polinomi $P_1 = X$ e $P_2 = 1 + X$ sono indipendenti. Infatti
$$\alpha_1 P_1 + \alpha_2 P_2 = \underline{0} \iff \alpha_1 x + \alpha_2(1 + x) = 0 \quad \forall x \in \mathbb{R}$$
da cui deduciamo che, per ogni $x \in \mathbb{R}$ si ha
$$(\alpha_1 + \alpha_2)x + \alpha_2 = 0.$$

In particolare, per $x = 0$ si ottiene

$$\alpha_2 = 0$$

e quindi deve essere

$$\alpha_1 x = 0 \quad \forall x \in \mathbb{R}$$

cioè $\alpha_1 = 0$.

6. Nell'insieme di tutte le funzioni reali di variabile reale, le due funzioni $f(x) = e^x$, $g(x) = \sin x$ sono indipendenti: infatti

$$\alpha e^x + \beta \sin x = 0, \quad \forall x \in \mathbb{R} \iff \alpha = \beta = 0.$$

Osservazione importante: Vogliamo analizzare più precisamente la nozione di dipendenza o indipendenza lineare in spazi di funzioni. Infatti, bisogna avere chiaro che esistono varie nozioni di indipendenza lineare quando si ha a che fare con funzioni: indipendenza lineare *puntuale in ogni punto*, *puntuale in un punto preciso* oppure indipendenza *come funzioni.*

Vale la seguente

Proposizione 5
Siano $f_1,\ f_2, \dots, f_m$ funzioni da $\mathbb{R}$ a valori in $\mathbb{R}^n$ (supponiamo $m \leq n$). Allora valgono le implicazioni seguenti:

(1) Per ogni $t \in \mathbb{R}$ in vettori $f_1(t),\ f_2(t), \dots, f_m(t)$ sono indipendenti in $\mathbb{R}^n$

$$\Downarrow$$

(2) Esiste $\bar{t} \in \mathbb{R}$ tale che i vettori $f_1(\bar{t}),\ f_2(\bar{t}), \dots, f_m(\bar{t})$ sono indipendenti in $\mathbb{R}^n$

$$\Downarrow$$

(3) Le funzioni $f_1,\ f_2, \dots, f_m$ sono funzioni indipendenti.

Corollario 6
Siano $f_1,\ f_2, \dots, f_m$ funzioni da $\mathbb{R}$ a valori in $\mathbb{R}^n$ (supponiamo $m \leq n$). Allora valgono le implicazioni seguenti:

(1') Le funzioni $f_1,\ f_2, \dots, f_m$ sono funzioni dipendenti.

$$\Downarrow$$

(2') Per ogni $t \in \mathbb{R}$ in vettori $f_1(t),\ f_2(t), \dots, f_m(t)$ sono dipendenti in $\mathbb{R}^n$

$$\Downarrow$$

(3') Esiste $\bar{t} \in \mathbb{R}$ tale che i vettori $f_1(\bar{t}),\ f_2(\bar{t}), \dots, f_m(\bar{t})$ sono dipendenti in $\mathbb{R}^n$

Dimostrazione: Dimostriamo solo la proposizione poiché il corollario è immediato a partire dalla proposizione.

L'implicazione (1) $\Rightarrow$ (2) è evidente. Mostriamo che (2) $\Rightarrow$ (3): per assurdo, supponiamo che le funzioni $f_1,\ f_2, \dots, f_m$ siano dipendenti. Allora, esistono m coefficienti $\alpha_1, \alpha_2, \dots, \alpha_m$ non tutti nulli tali che $\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_m f_m = \underline{0}$ (ove $\underline{0}$ è la funzione nulla). Ma allora in particolare in $\bar{t}$ si ha $\alpha_1 f_1(\bar{t}) + \alpha_2 f_2(\bar{t}) + \cdots + \alpha_m f_m(\bar{t}) = \underline{0}$ (ove $\underline{0}$ è il vettore nullo in $\mathbb{R}^n$) e quindi i vettori $f_1(\bar{t}),\ f_2(\bar{t}), \dots, f_m(\bar{t})$ sono dipendenti in $\mathbb{R}^n$, contrariamente all'ipotesi (2). •

Tuttavia, in generale, le implicazioni opposte sono false. Prendiamo come esempio $m = n = 2$.

Esempi:

1. Mostriamo che la proprietà *le due funzioni f_1 e f_2 sono funzioni indipendenti* **non implica** *esiste $\bar{t} \in I$ tale che i vettori $f_1(\bar{t})$ e $f_2(\bar{t})$ sono indipendenti in $\mathbb{R}^2$* .

   Consideriamo infatti le due funzioni $f_1,\ f_2 : \mathbb{R} \longrightarrow \mathbb{R}^2$ così definite:

   $$f_1(t) = (t, t), \quad f_2(t) = (e^t, e^t).$$

   Si tratta di due funzioni indipendenti, poiché non è vero che $f_1 = \alpha f_2$ con $\alpha \in \mathbb{R}$. Tuttavia, per ogni $t \in \mathbb{R}$ i vettori (t, t) e (e^t, e^t) sono proporzionali, dunque sono dipendenti (la costante di proporzionalità **varia con t**!).

2. Mostriamo ora che la proprietà *esiste $\bar{t} \in I$ tale che i vettori $f_1(\bar{t})$ e $f_2(\bar{t})$ sono indipendenti in $\mathbb{R}^2$* **non implica** *per ogni $t \in I$ in vettori $f_1(t)$ e $f_2(t)$ sono indipendenti in $\mathbb{R}^2$*.

   Consideriamo le due funzioni $f_1,\ f_2 : \mathbb{R} \longrightarrow \mathbb{R}^2$ così definite:

   $$f_1(t) = (t, 0), \quad f_2(t) = (0, t).$$

   Si ha chiaramente che per ogni $t \neq 0$ i due vettori $(t, 0)$ e $(0, t)$ sono indipendenti. Tuttavia in $t = 0$ non sono indipendenti, poiché entrambi risultano uguali a $(0, 0)$.

Torneremo su queste osservazioni fondamentali più avanti nel corso, nell'ambito delle soluzioni di un sistema di equazioni differenziali lineari.

Definizione 7
Siano $v_1, \ldots, v_n \in V$. Diremo che $\{v_1, \ldots, v_n\}$ è una **famiglia di generatori di** *V se per ogni $w \in V$ esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che*

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

Osservazione: In generale, lo stesso vettore può essere scritto in infiniti modi come combinazione lineare dei vettori $v_1, \ldots, v_n$.

Esempi:

1. In $\mathbb{R}^2$ la famiglia $\{(1, 0), (1, 1), (1, 2)\}$ è una famiglia di generatori. Infatti, per ogni $(x, y) \in \mathbb{R}^2$ è possibile trovare $(\alpha_1, \alpha_2, \alpha_3)$ tali che

   $$(x, y) = \alpha_1(1, 0) + \alpha_2(1, 1) + \alpha_3(1, 2).$$

   Basta risolvere il sistema

   $$\begin{cases} x = \alpha_1 + \alpha_2 + \alpha_3 \\ y = \alpha_2 + 2\alpha_3 \end{cases}$$

   che ammette infinite soluzioni del tipo $(x - y + \alpha_3, y - 2\alpha_3, \alpha_3)$ ove $\alpha_3 \in \mathbb{R}$ è arbitrario.

2. In $\mathbb{R}^2$ la famiglia $\{(1, 0), (0, 1)\}$ è generatrice e ogni vettore (x, y) ammette un'unica decomposizione nei vettori della famiglia.

3. In $\mathbb{R}^2$, l'unico vettore $(1, 1)$ non costituisce una famiglia generatrice poiché ad esempio il vettore $(1, 0)$ non si può scrivere come $\alpha(1, 1)$, con $\alpha \in \mathbb{R}$.

4. Nello spazio di polinomi $\mathbb{R}_1[X] = \{aX + b, \quad a, b \in \mathbb{R}\}$, la famiglia $\{X+1, X-1\}$ è generatrice (ESERCIZIO).

Definizione 8
Diremo che $B = \{v_1, \ldots, v_n\}$ è una **base** *di V se*

i) *$v_1, \ldots, v_n$ sono linearmente indipendenti* *ii)* $v_1, \ldots, v_n$ *costituiscono una famiglia di generatori.*

Osservazioni:

1. Non tutti gli spazi vettoriali ammettono una base formata da un numero finito di vettori.

   Esempio: Lo spazio di tutti i polinomi di ogni grado a coefficienti reali, che si denota con $\mathbb{R}[X]$, è uno spazio vettoriale che non ammette una base finita. Infatti, se per assurdo $B = \{P_1, P_2, \ldots, P_n\}$ fosse una base di V, allora denotiamo con $N = \max_{i=1,\ldots,n}\{\deg P_i\}$ e mostriamo che il polinomio $P = X^{N+1}$ non può essere scritto come combinazione lineare dei vettori di B. Infatti se fosse

   $$X^{N+1} = \alpha_1 P_1 + \cdots + \alpha_n P_n$$

   allora (indicando con D^{N+1} la derivata di ordine $N+1$):

   $$D^{N+1}X^{N+1} = D^{N+1}(\alpha_1 P_1 + \cdots + \alpha_n P_n)$$

   ma poiché $\deg(\alpha_1 P_1 + \cdots + \alpha_n P_n) \leq N$ la derivata $N+1$–esima del termine di destra è identicamente nulla, mentre

   $$D^{N+1}X^{N+1} = (N+1)!$$

2. Se V ammette una base di n vettori, allora si può dimostrare che ogni altra base ha esattamente n vettori. Chiameremo quindi **dimensione di** V il numero dei vettori di ogni sua base. Se V non ammette una base finita, diremo che V ha dimensione infinita. Ad esempio, $\mathbb{R}[X]$ ha dimensione infinita.

3. Se V è uno spazio di dimensione n e $W \subseteq V$ è un sottospazio, allora $\dim W \leq n$.

Proposizione 9
Se $B = \{v_1, \ldots, v_n\}$ *è una base per* V*, allora ogni vettore* $w \in V$ *ammette un'unica decomposizione nei vettori* $v_1, \ldots, v_n$*, cioè* $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ *ove i coefficienti* α_i *sono univocamente determinati.*

Dimostrazione: Infatti

$$\begin{aligned} & \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 \cdots + \beta_n v_n \\ \iff & (\alpha_1 - \beta_1)v_1 + \ldots (\alpha_n - \beta_n)v_n = \underline{0} \\ \iff & \begin{cases} \alpha_1 - \beta_1 = 0 \\ \alpha_2 - \beta_2 = 0 \\ \vdots \\ \alpha_n - \beta_n = 0 \end{cases} \end{aligned}$$

•

Osservazione importante:

1. Se $B = \{v_1, \dots, v_n\}$ è tale che per ogni $v \in V$ si ha $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ in modo unico, allora B è una base;
2. se $\dim V = n$ e $B = \{v_1, \dots, v_n\}$ è generatrice, allora B è una base;
3. se $\dim V = n$ e $B = \{v_1, \dots, v_n\}$ è formata da n vettori indipendenti, allora B è una base;
4. se $\dim V = n$, ogni famiglia generatrice è formata da almeno n vettori;
5. se $\dim V = n$, ogni famiglia indipendente è formata da al più n vettori.

Definizione 10
Sia $B = \{v_1, \dots, v_n\}$ una base per V. Allora, poiché per ogni $w \in V$ si ha $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ in modo unico, diremo che $\alpha_1, \alpha_2, \dots, \alpha_n$ sono le **componenti** *o* **coordinate** *di w nella base B.*

Dunque, fissata una base di V, un vettore è individuato univocamente dalle sue coordinate.

Esempi:

1. In $\mathbb{R}^n$, si definisce **base canonica** la base $B = \{(1, 0, \dots, 0), \dots, (0, 0, \dots, 1)\}$.
2. Nello spazio $\mathbb{R}_n[X] = \{a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0, \quad a_i \in \mathbb{R}, \quad i = 0, \dots, n\}$ dei polinomi di grado al più n si definisce **base canonica** la base $B = \{1, X, X^2, \dots, X^n\}$. Tale spazio ha dimensione $n+1$. Ad esempio, se $n = 3$ lo spazio ha dimensione 4 e il polinomio $P = X^3 - 2X^2 + \pi X - \sqrt{2}$ ha coordinate $(-\sqrt{2}, \pi, -2, 1)$ nella base canonica.

   Attenzione: L'ordine dei vettori nella base è importante per individuare le coordinate!

3. Lo spazio $V = \{aX^2 + b, \quad a, b \in \mathbb{R}\}$ dei polinomi pari di grado al più 3 ha dimensione 2 e una base è data da $B = \{1, X^2\}$. Un'altra base possibile è $B' = \{X^2 + 1, X^2 - 1\}$. Infatti i due polinomi sono linearmente indipendenti e generano lo spazio (ESERCIZIO).

Definizione 11
Sia V uno spazio vettoriale e siano $v_1, \dots, v_n \in V$. Chiameremo **sottospazio vettoriale generato da** *$v_1, \dots, v_n$ l'insieme di tutte le combinazioni lineari di $v_1, \dots, v_n$ e lo indicheremo con* $\operatorname{span}\{v_1, \dots, v_n\}$ *oppure* $\operatorname{vect}\{v_1, \dots, v_n\}$ *oppure con $\langle v_1, \dots, v_n \rangle$. Dunque*

$$\operatorname{span}\{v_1, \dots, v_n\} = \{\alpha_1 v_1 + \cdots + \alpha_n v_n, \quad \alpha_i \in \mathbb{R}, \quad i = 1, \dots, n\}$$

Esempio: In $\mathbb{R}^2$, consideriamo il vettore $v_1 = (1, 1)$. Allora

$$\operatorname{span}\{v_1\} = \{(\alpha, \alpha), \alpha \in \mathbb{R}\}$$

cioè è la retta passante per l'origine di vettore direttore $(1, 1)$.

Veniamo ora ad altre leggi definite in $\mathbb{R}^3$ che vorremmo estendere ad uno spazio vettoriale reale (o complesso) qualsiasi.

Definizione 12
Chiamiamo **prodotto scalare euclideo** *in $\mathbb{R}^3$ l'applicazione seguente:*

$$\begin{aligned} &\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R} \\ &(v, w) := \|v\| \|w\| \cos \alpha \end{aligned}$$

ove $\alpha \in [0, \pi]$ è l'angolo formato dai due vettori.

Osservazioni:

1. In componenti, se $v = (x, y, z)$ e $w = (x', y', z')$, si può dimostrare che
$$(v, w) = xx' + yy' + zz'.$$
2. Il prodotto scalare euclideo in $\mathbb{R}^3$ verifica le seguenti proprietà:
   (A) proprietà commutativa: per ogni $v, w \in \mathbb{R}^3$ si ha $(v, w) = (w, v)$;
   (B) positività: per ogni $v \in \mathbb{R}^3$ si ha $(v, v) = \|v\|^2$ dunque $(v, v) \geq 0$;
   (C) additività: per ogni $v, w, t \in \mathbb{R}^3$ si ha $(v + w, t) = (v, t) + (w, t)$;
   (D) omogeneità: per ogni $v, w \in \mathbb{R}^3$ e per ogni $\lambda \in \mathbb{R}$ si ha $(\lambda v, w) = \lambda(v, w)$.
3. Ricordiamo inoltre che vale l'importante proprietà geometrica: $(v, w) = 0$ se e solo se v e w sono ortogonali.

Definizione 13
Chiamiamo **norma (o modulo)** *in* $\mathbb{R}^3$ *l'applicazione seguente:*
$$\mathbb{R}^3 \longrightarrow \mathbb{R}$$
$$\|v\| := (v, v)^{\frac{1}{2}}$$
(ove (v, v) è il prodotto scalare euclideo di v per se stesso).

Osservazioni:

1. In componenti, se $v = (x, y, z)$ si ha $\|v\| = \sqrt{x^2 + y^2 + z^2}$.
2. La norma verifica le seguenti proprietà:
   - positività e legge di annullamento: per ogni $v \in \mathbb{R}^3$ si ha $\|v\| \geq 0$; inoltre $\|v\| = 0 \iff v = \underline{0}$;
   - per ogni $v \in \mathbb{R}^3$ e per ogni $\lambda \in \mathbb{R}$ si ha $\|\lambda v\| = |\lambda|\|v\|$;
   - disuguaglianza triangolare: per ogni $v, w \in \mathbb{R}^3$ si ha $\|v + w\| \leq \|v\| + \|w\|$.
3. La norma di un vettore in $\mathbb{R}^3$ non è altro che la lunghezza del vettore stesso o anche la distanza del punto che esso individua nello spazio dall'origine.

Osservazione: Dato il prodotto scalare euclideo e quindi la norma in $\mathbb{R}^3$, si ha seguente relazione: $\text{dist}(u, v) = \|u - v\|$ (se identifichiamo i vettori di $\mathbb{R}^3$ con i punti dello spazio).

Veniamo ora alla definizione di un prodotto scalare in uno spazio vettoriale reale astratto, che ci permetterà di definire una nozione di perpendicolarità, una norma e quindi una distanza.

Definizione 14
Sia $(V, +, \cdot)$ *uno spazio vettoriale reale. Definiamo* **prodotto scalare** *in* V *ogni legge*
$$(\cdot, \cdot) : V \times V \longrightarrow \mathbb{R}$$
che soddisfa le seguenti proprietà:

1. *(positività) per ogni $v \in V$ si ha $(v, v) \geq 0$; inoltre $(v, v) = 0 \iff v = \underline{0}$;*
2. *(simmetria) per ogni $v, w \in V$ si ha $(v, w) = (w, v)$*
3. *(additività) per ogni $v, w, t \in V$ si ha $(v + w, t) = (v, t) + (w, t)$;*

4. *(omogeneità) per ogni $v, w \in V$ ed ogni $\lambda \in \mathbb{R}$ si ha $(\lambda v, w) = \lambda(v, w)$.*

Diremo che $v, w \in V$ sono **ortogonali** *se $(v, w) = 0$.*

Definizione 15
In uno spazio vettoriale dotato di un prodotto scalare $(\cdot, \cdot)$, definiremo la **norma** *di un vettore $\|v\|$ nel modo seguente:*

$$\|v\| := (v, v)^{\frac{1}{2}}.$$

Definiremo poi la **distanza** *fra due vettori $\mathrm{dist}(v, w)$ come*

$$\mathrm{dist}(v, w) := \|v - w\|.$$

Osservazione: Si può dimostrare che la norma definita a partire da un prodotto scalare qualsiasi soddisfa le proprietà vere per la norma euclidea.

Esempi:

1. In $\mathbb{R}^n$ il **prodotto scalare euclideo** è il seguente:

$$((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=1}^{n} x_i y_i$$

   e la **norma euclidea** che ne consegue è la classica

$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

2. In $\mathbb{R}^3$, un esempio di prodotto scalare (non euclideo) il seguente:

$$((x_1, x_2, x_3), (y_1, y_2, y_3)) := x_1 y_1 + 2x_2 y_2 + 5x_3 y_3.$$

   (ESERCIZIO)

3. In $\mathbb{R}^2$, un esempio di prodotto scalare (non euclideo) è il seguente:

$$((x, y), (x', y')) := 5xx' + 2xy' + 2x'y + 4yy'$$

   (ESERCIZIO)

4. In $\mathbb{R}_1[X]$ un esempio di prodotto scalare è il seguente: per ogni $P_1, P_2 \in \mathbb{R}_1[X]$ poniamo

$$(P_1, P_2) := P_1(0)P_2(0) + P_1(1)P_2(1).$$

5. In $C([0, 1])$ un esempio di prodotto scalare è il seguente: per ogni $f, g \in C([0, 1])$ poniamo

$$(f, g) := \int_0^1 f(t)g(t)\,\mathrm{d}t.$$

La nozione di ortogonalità permette di individuare tra tutte le possibili basi di uno spazio vettoriale dotato di un prodotto scalare, quelle che meglio sono adattate alla struttura dello spazio.

Definizione 16
Sia $(V, +, \cdot)$ uno spazio vettoriale reale dotato di un prodotto scalare $(\cdot, \cdot)$ e sia $B = \{v_1, \ldots, v_n\}$ una base. Diremo che B è **ortogonale** *se i vettori v_i sono ortogonali a due a due (cioè $(v_i, v_j) = 0$ se $i \neq j$).*

Diremo che B è una **base ortonormale** *se:*

i) i vettori v_i sono ortogonali a due a due;

ii) per ogni $i = 1, \dots, n$ si ha $\|v_i\| = 1$,

cioè

$$(v_i, v_j) = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j. \end{cases}$$

Esempi:

1. In $\mathbb{R}^2$, con il prodotto scalare euclideo, la base canonica è ortonormale.

2. In $\mathbb{R}^2$ con il prodotto scalare

$$((x, y), (x', y')) := xx' + 5yy'$$

   la base canonica è ortogonale ma non è ortonormale, poiché

$$\begin{aligned} \|(1, 0)\| &= \sqrt{1^2 + 5 \cdot 0^2} = 1 \\ \|(0, 1)\| &= \sqrt{0^2 + 5 \cdot 1^2} = \sqrt{5}. \end{aligned}$$

   Una base ortonormale è invece $B = \{(1, 0), (0, \frac{1}{\sqrt{5}})\}$.

3. In $\mathbb{R}^2$ con il prodotto scalare definito in un esercizio precedente

$$((x, y), (x', y')) := 5xx' + 2xy' + 2x'y + 4yy'$$

   la base canonica non è neanche ortogonale (ESERCIZIO).

4. <u>**Base ortogonale nello spazio dei polinomi trigonometrici**</u>.

   Sia
$$T_n = \text{vect}\{1, \cos(kx), \sin(kx)\}_{k=1,\dots,n}.$$

   Lo spazio vettoriale T_n è un sottospazio dello spazio delle funzioni continue su $\mathbb{R}$, periodiche di periodo 2π (dotato della solita somma e prodotto per uno scalare delle funzioni). Per definizione di spazio vettoriale generato da una famiglia di vettori, si ha

$$f \in T_n \iff f = a_0 + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \sin(kx)), \quad a_k \in \mathbb{R},\ b_k \in \mathbb{R}.$$

   e chiameremo ogni funzione di questo tipo **polinomio trigonometrico di grado minore o uguale a n**.

   Possiamo considerare su T_n (o anche sullo spazio delle funzioni continue su $\mathbb{R}$, periodiche di periodo 2π) il prodotto scalare seguente: per ogni coppia di funzioni f, g si pone

$$(f, g) := \int_0^{2\pi} f(x) g(x) \, \mathrm{d}x.$$

   Abbiamo già visto in un esercizio precedente che questo è un prodotto scalare (bisogna sottolineare che le funzioni in questione sono periodiche di periodo 2π e quindi se si annullano su $[0, 2\pi]$, si annullano ovunque e quindi è verificata la legge di annullamento).

   Abbiamo quindi anche una norma:

$$\|f\| = \int_0^{2\pi} (f(x))^2 \, \mathrm{d}x.$$

Vogliamo mostrare che

$$B = \{1, \cos(kx), \sin(kx)\}_{k=1,\dots,n}$$

è una base ortogonale di T_n.

Per semplicità, consideriamo solo il caso $n = 2$, dunque

$$B = \{1, \cos x, \cos(2x), \sin x, \sin(2x)\},$$

(il caso generale si tratterà in modo totalmente analogo).

Il fatto che B sia generatrice di T_2 è evidente, per costruzione di T_2. Per mostrare che è una base, basta quindi mostrare che i vettori sono indipendenti. Poniamo

$$a + b\cos x + c\cos(2x) + d\sin x + e\sin(2x) = \underline{0}, \quad a, b, c, d, e \in \mathbb{R}$$

e vogliamo mostrare che allora $a = b = c = d = e = 0$. Ponendo successivamente $x = 0, \frac{\pi}{2}, \pi, -\frac{\pi}{2}, \frac{\pi}{4}$ otteniamo le seguenti 5 condizioni necessarie sui coefficenti:

$$\begin{cases} a + b + c = 0 \\ a - c + d = 0 \\ a - b + c = 0 \\ a - c - d = 0 \\ a + \dfrac{\sqrt{2}}{2}b + \dfrac{\sqrt{2}}{2}d + e = 0 \end{cases}$$

e poiché si tratta di un sistema lineare di 5 equazioni in 5 incognite con determinante dei coefficienti diverso da zero, l'unica soluzione è quella tutta nulla. Abbiamo mostrato che i vettori sono indipendenti.

Ora, per mostrare che sono ortogonali, basta osservare le seguenti uguaglianze:

$$\begin{aligned}
&\int_0^{2\pi} 1 \cdot \cos x \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} 1 \cdot \sin x \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} 1 \cdot \cos(2x) \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} 1 \cdot \sin(2x) \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \cos x \cos(2x) \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \cos x \sin x \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \cos x \sin(2x) \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \cos(2x) \sin x \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \cos(2x) \sin(2x) \,\mathrm{d}x = 0, \\
&\int_0^{2\pi} \sin x \sin(2x) \,\mathrm{d}x = 0.
\end{aligned}$$

Osserviamo che la base non è ortonormale: infatti:

$$\int_0^{2\pi} 1 \mathrm{d}x = 2\pi,$$
$$\int_0^{2\pi} (\cos x)^2 \,\mathrm{d}x = \pi,$$
$$\int_0^{2\pi} (\sin x)^2 \,\mathrm{d}x = \pi,$$
$$\int_0^{2\pi} (\cos(2x))^2 \,\mathrm{d}x = \pi,$$
$$\int_0^{2\pi} (\sin(2x))^2 \,\mathrm{d}x = \pi;$$

dunque

$$\|1\| = \sqrt{2\pi},$$
$$\|\cos x\| = \sqrt{\pi},$$
$$\|\sin x\| = \sqrt{\pi} i,$$
$$\|\cos(2x)\| = \sqrt{\pi},$$
$$\|\sin(2x)\| = \sqrt{\pi}$$

e quindi una base ortonormale di T_2 è

$$\tilde{B} = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\cos(2x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}} \right\}.$$

In generale, una base ortonormale di T_n è

$$\tilde{B} = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(kx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}} \right\}_{k=1,\dots,n}.$$

Osservazione: Se $(V, +, \cdot)$ è uno spazio vettoriale dotato di un prodotto scalare $(\cdot, \cdot)$ e $B = \{v_1, \dots, v_n\}$ è una base ortonormale, allora il prodotto scalare ha un'espressione particolarmente semplice in termini delle coordinate dei vettori nella base B. Infatti, siano $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e $w = \beta_1 v_1 + \cdots + \beta_n v_n$; utilizzando le proprietà di additività e omogeneità del prodotto scalare e il fatto che la base è ortonormale, si ha:

$$\begin{aligned}(v, w) &= (\alpha_1 v_1 + \cdots + \alpha_n v_n, \beta_1 v_1 + \cdots + \beta_n v_n) \\ &= \sum_{i=1}^n \alpha_i (v_i,\ \beta_1 v_1 + \cdots + \beta_n v_n) = \sum_{i=1}^n \alpha_i \sum_{j=1}^n \beta_j (v_i, v_j) \quad = \sum_{i=1}^n \alpha_i \beta_i\end{aligned}$$

cioè il prodotto scalare ha la stessa espressione del prodotto scalare euclideo in $\mathbb{R}^n$ quando i vettori sono decomposti nella base canonica.

Senza fare una verifica diretta, potevamo dedurre anche da questo che negli esempi precedenti 2 e 3 di prodotto scalare non euclideo in $\mathbb{R}^3$ e $\mathbb{R}^2$, la base canonica NON è ortonormale, poiché abbiamo definito tali prodotti proprio tramite le coordinate dei vettori nella base canonica.

Vale il seguente

Teorema 17 (procedimento di ortonormalizzazione di Gram–Schmidt)
In ogni spazio vettoriale di dimensione finita dotato di un prodotto scalare, esiste sempre una base ortonormale.

(Non lo dimostriamo).

(ESERCIZIO: trovare una base ortonormale in $\mathbb{R}^2$ rispetto al prodotto scalare non euclideo $((x,y),(x',y')) := 5xx' + 2xy' + 2x'y + 4yy'$).

Ricordiamo infine che in $\mathbb{R}^3$ è definita un'ultima applicazione: il prodotto vettoriale.

Definizione 18
Chiamiamo **prodotto vettoriale** *in* $\mathbb{R}^3$ *l'applicazione seguente:*

$$\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$
$$v \wedge w = z$$

ove z è definito nel modo seguente:

i) $\|z\| = \|v\|\|w\| \sin\alpha$*, ove* $\alpha \in [0,\pi]$ *è l'angolo formato dai due vettori (osserviamo che* $\sin\alpha \geq 0$ *per* $\alpha \in [0,\pi]$ *e dunque la norma è un numero maggiore o uguale a zero, coerentemente con la definizione di norma);*

ii) la direzione di z è perpendicolare al piano formato dai due vettori;

iii) il verso di z si trova tramite la regola della mano destra.

Osservazioni:

1. In componenti, $v = (x,y,z)$ e $w = (x',y',z')$, si può dimostrare (ma non è banale) che
$$v \wedge w = (yz' - y'z, x'z - xz', xy' - x'y).$$

2. Il prodotto vettoriale gode delle seguenti proprietà:
   - proprietà anti–commutativa: per ogni $v, w \in \mathbb{R}^3$ si ha $v \wedge w = -w \wedge v$;
   - proprietà distributiva della somma rispetto al prodotto vettoriale: per ogni $v, w, t \in \mathbb{R}^3$ si ha $(v + w) \wedge t = v \wedge t + w \wedge t$;
   - per ogni $v, w \in \mathbb{R}^3$ e per ogni $\lambda \in \mathbb{R}$ si ha $\lambda(v \wedge w) = (\lambda v) \wedge w$;
   - per ogni $v \in \mathbb{R}^3$ si ha $v \wedge v = \underline{0}$.

3. Ricordiamo inoltre che vale l'importante proprietà geometrica: $v \wedge w = \underline{0}$ se e solo se v e w sono paralleli.

Sulla possibilità di estendere la definizione di prodotto vettoriale ad un generico spazio vettoriale astratto non ci dilunghiamo, poiché l'argomento non si può trattare rapidamente. Possiamo però dire che in ogni spazio vettoriale V di dimensione n l'analogo dell'operatore di prodotto vettoriale sarebbe un operatore definito su V^{n-1}, dunque se $n \neq 3$, non sarebbe un prodotto vettoriale fra due vettori, ma su $n-1$ vettori.

1.2 Applicazioni lineari

Veniamo ora alle funzioni definite fra spazi vettoriali che sono "ben adattate" alla struttura vettoriale.

Definizione 19
Siano V, W due spazi vettoriali reali e sia $L : V \longrightarrow W$ una funzione. Diremo che L è una **applicazione o funzione lineare** *se per ogni $\lambda \in \mathbb{R}$, per ogni $v_1, v_2 \in V$ si ha:*

1. *additività:* $L(v_1 + v_2) = L(v_1) + L(v_2)$*;*

2. *omogeneità:* $L(\lambda v_1) = \lambda L(v_1)$.

In altre parole, per ogni $v_1, v_2 \in V$ e per ogni $\lambda_1, \lambda_2 \in \mathbb{R}$ si ha

$$L(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 L(v_1) + \lambda_2 L(v_2)$$

(l'immagine di ogni combinazione lineare è la combinazione lineare delle immagini).

Proposizione 20
Condizione necessaria affinché $L : V \longrightarrow W$ sia lineare è $L(\underline{0}_V) = \underline{0}_W$.

Dimostrazione: Infatti, preso $v \in V$, si ha $L(\underline{0}_V) = L(v - v) = L(v) - L(v) = \underline{0}_W$. •

Esempi:

1. $\mathbb{R}$ è uno spazio vettoriale reale di dimensione 1. Ci chiediamo quali sono le applicazioni lineari $L : \mathbb{R} \longrightarrow \mathbb{R}$. Sappiamo che L sarà lineare $\iff$ $L(x\lambda) = xL(\lambda)$ per ogni $\lambda, x \in \mathbb{R}$ (questa volta abbiamo indicato con λ il vettore e con x lo scalare). Ma allora $L(x) = L(x \cdot 1) = xL(1)$, dunque l'applicazione è completamente definita dal suo valore in $x = 1$. Se $L(1) = \alpha$, allora

$$L(x) = x\alpha, \quad \alpha \in \mathbb{R}.$$

   Per esempio, la funzione $L(x) = 3x$ è lineare, mentre le funzioni $g(x) = x^2$ oppure $h(x) = e^x$ non sono lineari.

2. La funzione $f(x) = 3x + 4$ **non** è lineare, poiché $f(0) = 4 \neq 0$. Tuttavia, possiamo dire che "cresce linearmente" oppure "ha un andamento lineare". Infatti, se chiamiamo Δx l'incremento della variabile indipendente, avremo che l'incremento della funzione $(\Delta f)_x$ definita da $(\Delta f)_x(\Delta x) = f(x + \Delta x) - f(x)$ è una funzione lineare dell'incremento, poiché

$$f(x + \Delta x) - f(x) = 3(x + \Delta x) + 4 - 3x - 4 = 3\,\Delta x.$$

3. Una funzione $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ lineare è la seguente:

$$L(x, y) = (2x - y, x - 2y)$$

   mentre una funzione NON lineare è la seguente

$$g(x, y) = (xy, \sin x).$$

   (ESERCIZIO)

4. Se $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f \in C^1(\mathbb{R})\}$ e $W = \{g : \mathbb{R} \longrightarrow \mathbb{R}, \quad g \in C(\mathbb{R})\}$ (sono spazi vettoriali di dimensione infinita), allora

$$\begin{array}{rcl} D : V & \longrightarrow & W \\ f & \longmapsto & f' \end{array}$$

   è un'applicazione lineare, poiché

$$\begin{array}{l} (f_1 + f_2)' = f_1' + f_2' \\ (\lambda f)' = \lambda f'. \end{array}$$

5. Se $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f \in C(\mathbb{R})\}$ e $W = \{g : \mathbb{R} \longrightarrow \mathbb{R}, \quad g \in C^1(\mathbb{R})\}$, allora

$$\begin{array}{rcl} I : V & \longrightarrow & W \\ f & \longmapsto & \int_0^x f(t)dt \end{array}$$

   è un'applicazione lineare, poiché

$$\begin{array}{l} \displaystyle\int_0^x (f_1 + f_2)(t)dt = \int_0^x f_1(t)dt + \int_0^x f_2(t)dt \\ \displaystyle\int_0^x (\lambda f)(t)dt = \lambda \int_0^x f(t)dt. \end{array}$$

6. Se $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f \in C^2(\mathbb{R})\}$ e $W = \{g : \mathbb{R} \longrightarrow \mathbb{R}, \quad g \in C(\mathbb{R})\}$, le applicazioni
$$\begin{array}{rcl} L_1 : V & \longrightarrow & W \\ f & \longmapsto & f'' \\ L_2 : V & \longrightarrow & W \\ f & \longmapsto & xf' \\ L_3 : V & \longrightarrow & W \\ f & \longmapsto & e^x f \end{array}$$
sono lineari (ESERCIZIO) e quindi anche la loro somma
$$\begin{array}{rcl} L : V & \longrightarrow & W \\ f & \longmapsto & f'' + xf + e^x f \end{array}$$
è lineare.

Ricordiamo alcune definizioni valide per ogni funzione $F : A \longrightarrow B$ definita tra due insiemi di natura qualsiasi:

- si dice che F è iniettiva se per ogni $a_1, a_2 \in A$ si ha $F(a_1) = F(a_2) \iff a_1 = a_2$;
- si dice che F è suriettiva se per ogni $b \in B$ esiste $a \in A$ tale che $F(a) = b$;
- si dice che F è biunivoca se è iniettiva e suriettiva.

Siano ora V e W spazi vettoriali e sia $L : V \longrightarrow W$ lineare.

Definizione 21
L'insieme $\{u \in V : L(u) = \underline{0}\}$ *si dice* **nucleo di** L *e si indica con* $\mathrm{Ker}L$.
L'insieme $\{w \in W : \exists v \in V \ t.c.\ L(v) = w\}$ *si dice* **immagine di** L *e si indica con* $\mathrm{Im}L$.

Esempi:

1. Se $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ è l'applicazione lineare definita da $L(x, y) = (2x - y, x - 2y)$, allora
$$\mathrm{Ker}\, L = \{(x, y) : L(x, y) = (0, 0)\} = \{(x, y) : 2x - y = 0, x - 2y = 0\} = \{(0, 0)\}.$$
Inoltre
$$\begin{array}{rcl} (X, Y) \in \mathrm{Im}\, L & \iff & \exists (x, y) \in \mathbb{R}^2 \ t.c.\ (2x - y, x - 2y) = (X, Y) \\ & \iff & \exists (x, y) \in \mathbb{R}^2 \ t.c.\ x(2, 1) + y(-1, -2) = (X, Y) \\ & \iff & (X, Y) \in \mathrm{vect}\{(2, 1), (-1, -2)\} \end{array}$$
dunque $\mathrm{Im}\, L = \mathrm{vect}\{(2, 1), (-1, -2)\}$, ma poiché $\{(2, 1), (-1, -2)\}$ è una base di $\mathbb{R}^2$, allora si deduce che $\mathrm{Im}\, L = \mathbb{R}^2$, cioè L è suriettiva. In particolare, si ha $(x, y) = (\frac{2X-Y}{3}, \frac{X-2Y}{3})$.

2. Se $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f \in C^2(\mathbb{R})\}$ e $W = \{g : \mathbb{R} \longrightarrow \mathbb{R}, \quad g \in C(\mathbb{R})\}$ e L è l'applicazione lineare definita in un esempio precedente:
$$\begin{array}{rcl} L : V & \longrightarrow & W \\ f & \longmapsto & f'' + xf' + e^x f \end{array}$$
allora
$$\mathrm{Ker}\, L = \{f : f'' + xf' + e^x f = \underline{0}\},$$
cioè è l'integrale generale dell'equazione differenziale lineare omogenea. Per quanto riguarda $\mathrm{Im}\, L$, dalla teoria delle equazioni differenziali lineari sappiamo che, poiché i coefficienti x e e^x sono funzioni continue definite su $\mathbb{R}$, allora per ogni $g \in C(\mathbb{R})$ esiste almeno una soluzione dell'equazione differenziale
$$f'' + xf' + e^x f = g$$
(di fatto, come ricorderemo fra poco, ne esistono infinite e non ci si deve confondere con il teorema di esistenza ed unicità della soluzione di un problema di Cauchy, che riguarda quante soluzioni passano per un punto $(x_0, y_0) \in \mathbb{R}^2$ fissato!) Dunque $\mathrm{Im}\, L = W$.

Osservazione: Osserviamo che $\underline{0}_V \in \operatorname{Ker} L$ e $\underline{0}_W \in \operatorname{Im} L$. Infatti, la condizione necessaria affinché L sia lineare è $L(\underline{0}_V) = \underline{0}_W$.

Proposizione 22
Se $L : V \longrightarrow W$ è lineare, allora $\operatorname{Ker} L \subset V$ è un sottospazio vettoriale e $\operatorname{Im} L \subset W$ è un sottospazio vettoriale.

Dimostrazione: Abbiamo già osservato che la condizione necessaria affinché $\operatorname{Ker} L$ e $\operatorname{Im} L$ siano sottospazi vettoriali è verificata. Poiché tale condizione non è sufficiente, procediamo ora con la dimostrazione. Mostriamo che $\operatorname{Ker} L$ è sottospazio vettoriale di V. Infatti, per ogni $u_1, u_2 \in \operatorname{Ker} L$ si ha $L(u_1 + u_2) = L(u_1) + L(u_2) = \underline{0} + \underline{0} = \underline{0}$, dunque $u_1 + u_2 \in \operatorname{Ker} L$. Inoltre, per ogni $\lambda \in \mathbb{R}$ si ha $L(\lambda u_1) = \lambda L(u_1) = \lambda \underline{0} = \underline{0}$, dunque anche $\lambda u_1 \in \operatorname{Ker} L$.

Mostriamo ora che $\operatorname{Im} L$ è sottospazio di W. Siano $w_1, w_2 \in \operatorname{Im} L$ e sia $\lambda \in \mathbb{R}$. Siano quindi $v_1, v_2 \in V$ controimmagini di w_1, w_2, cioè $L(v_1) = w_1$, $L(v_2) = w_2$. Ma allora $L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2$, dunque $w_1 + w_2 \in \operatorname{Im} L$ e anche $L(\lambda v_1) = \lambda L(v_1) = \lambda w_1$, dunque anche $\lambda w_1 \in \operatorname{Im} L$. •

Esempio: Grazie a questa proposizione, se $V = C^2(\mathbb{R})$ e $W = C(\mathbb{R})$ e L è l'applicazione definita nell'esempio precedente:

$$\begin{array}{rcl} L : V & \longrightarrow & W \\ f & \longmapsto & f'' + xf' + e^x f \end{array}$$

allora

$$\operatorname{Ker} L = \{f : f'' + xf' + e^x f = \underline{0}\}$$

è uno spazio vettoriale. Dunque, l'integrale generale delle soluzioni dell'equazione differenziale lineare omogenea

$$y'' + xy' + e^x y = 0$$

(che abbiamo riscritto con le notazioni più abituali in equazioni differenziali) è uno spazio vettoriale, sottospazio delle funzioni $C(\mathbb{R})$. Avevamo già osservato questa proprietà con una dimostrazione diretta nel corso delle esercitazioni.

Proposizione 23
Un'applicazione $L : V \longrightarrow W$ lineare è iniettiva $\iff \operatorname{Ker} L = \{\underline{0}\}$.

Dimostrazione: Se L è iniettiva, dalla definizione di iniettività è immediato che $\operatorname{Ker} L = \{\underline{0}\}$, dato che essendo $L(\underline{0}) = \underline{0}$, non ci può essere alcun altro elemento di V che ha come immagine $\underline{0} \in W$. Viceversa, se $\operatorname{Ker} L = \{\underline{0}\}$, allora supponiamo che v_1 e $v_2 \in V$ verifichino $L(v_1) = L(v_2)$. Per la linearità di L si ha $L(v_1 - v_2) = \underline{0}$ ma allora $v_1 - v_2 \in \operatorname{Ker} L$ dunque per ipotesi $v_1 - v_2 = \underline{0}$, cioè $v_1 = v_2$ e quindi L è iniettiva. •

Vogliamo ora caratterizzare l'insieme delle controimmagini di un uno stesso elemento $w \in \operatorname{Im} L$. Abbiamo già risolto il problema nel caso in cui $w = \underline{0}$; abbiamo visto che l'insieme delle controimmagini (che è il nucleo di L) è un sottospazio vettoriale. Se ora $w \in \operatorname{Im} L$ e $w \neq \underline{0}$, cosa possiamo dire dell'insieme delle sue controimmagini? Vediamo ora che tale insieme NON è un sottospazio vettoriale, ma ha una struttura molto particolare.

Proposizione 24 (teorema di struttura)
Sia $L : V \longrightarrow W$ lineare e sia $w \in \operatorname{Im} L \subset W$ con $w = L(z)$, $z \in V$. Allora

$$\{v \in V : L(v) = w\} = \{v \in V : v = z + u, u \in \operatorname{Ker} L\}.$$

Dimostrazione: Mostriamo dapprima che

$$\{v \in V : v = z + u, u \in \operatorname{Ker} L\} \subseteq \{v \in V : L(v) = w\}.$$

Infatti, se $v = z + u$ con $u \in \operatorname{Ker} L$, si ha $L(v) = L(z) + L(u) = w + \underline{0} = w$.

Mostriamo ora che

$$\{v \in V : L(v) = w\} \subseteq \{v \in V : v = z + u, u \in \operatorname{Ker} L\}.$$

Infatti, se $v \in V$ è tale che $L(v) = w$, il vettore $v - z$ appartiene a $\operatorname{Ker} L$, poiché $L(v - z) = L(v) - L(z) = w - w = \underline{0}$. Dunque $v - z = u$, con $u \in \operatorname{Ker} L$, cioè $v = z + u$. •

Esempio: La proposizione precedente ci permette di ritrovare il teorema di struttura dell'integrale generale di una equazione differenziale lineare. Infatti, se consideriamo ancora l'esempio precedente:

$$\begin{array}{rcl} L : C^2(\mathbb{R}) & \longrightarrow & C(\mathbb{R}) \\ f & \longmapsto & f'' + xf' + e^x f \end{array}$$

si può dedurre che, data $g_0 \in \operatorname{Im} L = C(\mathbb{R})$ (vedere l'esempio 2 della pagina precedente), si ha

$$\{f \in V : L(f) = g_0\} = \{f \in V : f = f_0 + u,\ u \in \operatorname{Ker} L,\ L(f_0) = g_0\},$$

e questo è proprio il teorema di struttura dell'integrale generale: ogni soluzione dell'equazione non omogenea si trova sommando ad una soluzione particolare dell'equazione non omogenea (f_0) una soluzione dell'equazione omogenea associata (u).

1.3 Applicazioni lineari tra spazi di dimensione finita

Proposizione 25
Sia $L : V \longrightarrow W$ un'applicazione lineare, $\dim V = n$, $\dim W = m$ *e sia $B_V = \{v_1, \ldots, v_n\}$ una base di V. Allora $\operatorname{Im} L$ è generato dai vettori $\{L(v_1), \ldots, L(v_n)\}$*

Dimostrazione: Poiché $\{v_1, \ldots, v_n\}$ è una base di V, per ogni $v \in V$ si ha $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e quindi per la linearità di L si ha $L(v) = \alpha_1 L(v_1) + \cdots + \alpha_n L(v_n)$, cioè $\operatorname{Im} L \subset \operatorname{vect}\{L(v_1), \ldots, L(v_n)\}$; viceversa se $w = \alpha_1 L(v_1) + \cdots + \alpha_n L(v_n) \in \operatorname{vect}\{L(v_1), \ldots, L(v_n)\}$, allora $w \in \operatorname{Im} L$ poiché $w = L(\alpha_1 v_1 + \cdots + \alpha_n v_n)$. •

Osservazione: I vettori $\{L(v_1), \ldots, L(v_n)\}$ non costituiscono in generale una famiglia indipendente!

Esempio: Se

$$\begin{array}{rcl} L : \mathbb{R}^2 & \longrightarrow & \mathbb{R}_1[X] \\ (a, b) & \longmapsto & aX \end{array}$$

allora

$$\{L(1, 0), L(0, 1)\} = \{X, \underline{0}\}$$

e sappiamo che ogni famiglia di vettori che contiene il vettore nullo è sempre dipendente. Tuttavia, la famiglia precedente genera $\operatorname{Im} L$ e quindi

$$\operatorname{Im} L = \operatorname{vect}\{X, \underline{0}\} = \{\alpha X, \alpha \in \mathbb{R}\}.$$

Teorema 26
Sia $L : V \longrightarrow W$ un'applicazione lineare e $\dim V = n$, $\dim W = m$. *Allora*

$$\dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L) = \dim V = n.$$

Dimostrazione: (**facoltativa**) Sia $k = \dim(\operatorname{Ker} L)$ con $0 \leq k \leq n$ e sia $B_{\operatorname{Ker} L} = \{u_1, \ldots, u_k\}$ una base di $\operatorname{Ker} L$. Sia poi $B_V = \{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$ una base di V ottenuta completando la base del nucleo con $n-k$ vettori. Chiaramente questi ultimi sono in particolare linearmente indipendenti e non appartenenti a $\operatorname{Ker} L$. Basta dimostrare che $\dim(\operatorname{Im} L) = n-k$ e per fare ciò dimostriamo che la famiglia $\{L(v_{k+1}), \ldots, L(v_n)\}$ è una base di $\operatorname{Im} L$.

Mostriamo che è una famiglia generatrice: sia $w \in \operatorname{Im} L$; allora esiste $x \in V$ tale che $L(x) = w$. Ma poiché B_V è una base per V, si ha

$$x = \alpha_1 u_1 + \cdots + \alpha_k u_k + \beta_{k+1} v_{k+1} + \cdots + \beta_n v_n,$$

da cui, per la linearità di L si ha

$$w = L(x) = \alpha_1 L(u_1) + \cdots + \alpha_k L(u_k) + \beta_{k+1} L(v_{k+1}) + \cdots + \beta_n L(v_n) = \beta_{k+1} L(v_{k+1}) + \cdots + \beta_n L(v_n)$$

poiché i vettori u_i appartengono al nucleo di L. Dunque, $w \in \operatorname{vect}\{L(v_{k+1}), \ldots, L(v_n)\}$. Viceversa, ogni vettore $w = \beta_{k+1} L(v_{k+1}) + \cdots + \beta_n L(v_n) \in \operatorname{vect}\{L(v_{k+1}), \ldots, L(v_n)\}$ appartiene a $\operatorname{Im} L$, poiché si ha $w = L(\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n)$ e $\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n \in V$.

Mostriamo ora che la famiglia $\{L(v_{k+1}), \ldots, L(v_n)\}$ è linearmente indipendente.

Siano $\beta_{k+1}, \ldots, \beta_n \in \mathbb{R}$ tali che

$$\beta_{k+1} L(v_{k+1}) + \cdots + \beta_n L(v_n) = \underline{0};$$

allora per la linearità di L si ha

$$L(\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n) = \underline{0}$$

cioè

$$\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n \in \operatorname{Ker} L$$

e ciò implica che esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tali che

$$\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n = \alpha_1 u_1 + \cdots + \alpha_k u_k$$

e quindi

$$\beta_{k+1} v_{k+1} + \cdots + \beta_n v_n - \alpha_1 u_1 - \cdots - \alpha_k u_k = \underline{0}$$

da cui tutti i coefficienti sono nulli (poiché i vettori sono linearmente indipendenti per ipotesi). •

Osservazioni:

1. Se L è iniettiva, allora $\dim(\operatorname{Ker} L) = 0$ e quindi $\dim(\operatorname{Im} L) = n$.
2. Se L è suriettiva, allora $\dim(\operatorname{Im} L) = m$ e quindi $\dim(\operatorname{Ker} L) = n - m$.

Consideriamo dunque i vari casi $n > m$, $n < m$ e $n = m$.

- Se $n > m$, allora L non può essere iniettiva né biunivoca.
- Se $n < m$, allora L non può essere suriettiva né biunivoca.
- Se $n = m$, allora L è iniettiva $\iff$ L è suriettiva $\iff$ L è biunivoca, cioè $\dim(\operatorname{Ker} L) = 0 \iff \dim(\operatorname{Im} L) = n$.

Esempi:

1. Consideriamo
$$\begin{array}{rcl} L : \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x, y, z) & \longmapsto & (2x - y, z + y); \end{array}$$
è un'applicazione lineare. Poiché la dimensione dello spazio di partenza è superiore allo spazio di arrivo, tale applicazione non potrà essere iniettiva. Viceversa, potrebbe essere o meno suriettiva. Si ha
$$\begin{aligned} \operatorname{Im} L &= \{(X, Y) \in \mathbb{R}^2 : \exists (x, y, z) \in \mathbb{R}^3 \ t.c.\ (X, Y) = (2x - y, z + y)\} \\ &= \{(X, Y) \in \mathbb{R}^2 : \exists (x, y, z) \in \mathbb{R}^3 \ t.c.\ (X, Y) = x(2, 0) + y(-1, 1) + z(0, 1)\} \\ &= \operatorname{vect}\{(2, 0), (-1, 1), (0, 1)\}. \end{aligned}$$
Ora, i vettori $\{(2, 0), (-1, 1), (0, 1)\}$ costituiscono una famiglia generatrice di $\mathbb{R}^2$ ma chiaramente non sono indipendenti. Dunque, per il teorema (poiché $\dim V = 3$):
$$3 = \dim \operatorname{Ker} L + 2 \iff \dim \operatorname{Ker} L = 1.$$
D'altra parte, se ricerchiamo esplicitamente il nucleo di L avremo:
$$\begin{aligned} \operatorname{Ker} L &= \{(x, y, z) \in \mathbb{R}^3 : (2x - y, z + y) = (0, 0)\} = \{(x, y, z) \in \mathbb{R}^3 : y = 2x, z = -2x, x \in \mathbb{R}\} \\ &= \{\alpha(1, 2, -2), \alpha \in \mathbb{R}\}. \end{aligned}$$

2. Consideriamo ora
$$\begin{array}{rcl} L : \mathbb{R}^2 & \longrightarrow & \operatorname{vect}\{e^x, \sin x, \cos x\} \\ (a, b) & \longmapsto & (a + b)e^x + (a - b)\sin x + 2b \cos x; \end{array}$$
è un'applicazione lineare. Poiché la dimensione dello spazio di partenza è più piccola della dimensione dello spazio di arrivo, sicuramente questa applicazione non potrà essere suriettiva. Potrebbe essere o meno iniettiva. Per determinarne il nucleo, determiniamo quali vettori $(a, b) \in \mathbb{R}^2$ sono tali che $(a + b)e^x + (a - b)\sin x + 2b \cos x = \underline{0}$. Poiché sappiamo che le funzioni $e^x, \sin x, \cos x$ sono indipendenti, l'unica possibilità è che
$$\begin{cases} a + b = 0 \\ a - b = 0 \\ 2b = 0 \end{cases}$$
che ha come unica soluzione la soluzione nulla $(0, 0)$. Dunque, $\operatorname{Ker} L = \{\underline{0}\}$. Per il teorema di nullità più rango, si ha $2 = 0 + \dim \operatorname{Im} L$ da cui $\dim \operatorname{Im} L = 2$. Per caratterizzare il sottospazio $\operatorname{Im} L$ di $\operatorname{vect}\{e^x, \sin x, \cos x\}$, scegliamo in $\mathbb{R}^2$ la base canonica; sappiamo quindi che $\operatorname{Im} L = \left\{ L \begin{pmatrix} 1 \\ 0 \end{pmatrix}, L \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \operatorname{vect}\{e^x + \sin x, e^x - \sin x + 2 \cos x\}$.

1.3.1 Richiami sul calcolo matriciale

Come abbiamo già ricordato nel paragrafo sugli spazi vettoriali, l'insieme $M_{m\times n}(\mathbb{R})$, munito di somma e prodotto per un numero reale è uno spazio vettoriale reale (così come l'insieme $M_{m\times n}(\mathbb{C})$, munito di somma e prodotto per un numero complesso è uno spazio vettoriale complesso).

Ricordiamo che date due matrici $A \in M_{m\times n}$ e $B \in M_{s\times p}$, diremo che A **è conformabile a** B se $n = s$ (cioè A conformabile a B se il numero di colonne di A è uguale al numero di righe di B). Attenzione che la nozione di conformabilità non è simmetrica.

Tra matrici conformabili è possibile definire un **prodotto righe per colonne** nel modo seguente:

$$\begin{array}{rcl} M_{m\times n} \times M_{n\times p} & \longrightarrow & M_{m\times p} \\ [a_{ij}][b_{jk}] & \longmapsto & [c_{ik}] \end{array}$$

con la regola:

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}, \quad i = 1, \ldots, m, \quad k = 1, \ldots, p.$$

Osservazione importante: In generale, se è possibile eseguire il prodotto AB non è detto che si possa eseguire anche il prodotto BA e anche in questo caso (cioè quando $A \in M_{m\times n}$ e $B \in M_{n\times m}$, non è in generale vero che $AB = BA$ (se $m \neq n$ si ha $AB \in M_{m\times m}$ mentre $BA \in M_{n\times n}$; se anche A e B sono matrici quadrate, i prodotti AB e BA non sono in generale uguali).

Valgono le due proprietà:

- **proprietà associativa del prodotto righe per colonne:** per ogni A, B e C conformabili, si ha $(AB)C = A(BC)$;
- **proprietà distributiva della somma rispetto al prodotto righe per colonne:** per ogni A, B e C si ha $(A + B)C = AC + BC$ e $A(B + C) = AB + AC$, quando tali operazioni sono possibili.

Ricordiamo che nello spazio delle matrici quadrate di ordine n è definita la **matrice identità** I_n

$$I_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Tale matrice soddisfa la proprietà $AI_n = I_n A = A$ per ogni $A \in M_{n\times n}$.

Data una matrice $A \in M_{m\times n}$, chiameremo **matrice trasposta** $A^T \in M_{n\times m}$ la matrice che si ottiene scambiando le righe con le colonne di A

Esempio:

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & \pi & -2 \end{pmatrix} \quad A^T = \begin{pmatrix} 1 & 0 \\ -2 & \pi \\ 3 & -2 \end{pmatrix}$$

Se $A = A^T$ (e quindi deve essere per forza $m = n$), la matrice A si dice **simmetrica** (cioè gli elementi della matrice sono simmetrici rispetto alla diagonale principale).

Esempio:

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -2 & \pi & 1 \\ 3 & 1 & e \end{pmatrix}$$

Una matrice $A \in M_{n\times n}$ si dice **triangolare alta (o bassa)** se gli elementi al di sotto (o sopra) della diagonale principale sono nulli.

Esempi:

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & \pi & 1 \\ 0 & 0 & e \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ -2 & \pi & 0 \\ 3 & -1 & 2 \end{pmatrix}$$

1.3.2 Matrici e applicazioni lineari

Consideriamo ora una matrice $A \in M_{m\times n}(\mathbb{R})$ e costruiamo un'applicazione lineare $L : V \longrightarrow W$ ove $\dim V = n$, $\dim W = m$ (per ora considereremo solo matrici a coefficienti reali e spazi vettoriali reali). Prendiamo dapprima un esempio numerico; sia $A \in M_{2\times 3}$ così definita:

$$A = \begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix}$$

e definiamo un'applicazione lineare $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$. Per fare questo, dobbiamo dapprima scegliere una base in $\mathbb{R}^3$ e una base in $\mathbb{R}^2$ e consideriamo in $\mathbb{R}^3$ la base canonica $B_3 = \{(1,0,0),(0,1,0),(0,0,1)\}$ e in $\mathbb{R}^2$ la base canonica $B_2 = \{(1,0),(0,1)\}$. Si ha quindi $u \in \mathbb{R}^2 \iff u = x(1,0)+y(0,1) = (x,y)$ e $v \in \mathbb{R}^3 \iff v = X(1,0,0) + Y(0,1,0) + Z(0,0,1) = (X,Y,Z)$.

Definiamo tramite il prodotto righe per colonne un'applicazione $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ nel modo seguente: per ogni $v = (X,Y,Z)$ poniamo

$$L(v) := A\,v = \begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 3X + \pi Y + (\log 2)Z \\ -X + eY - 2Z \end{pmatrix}$$

Tale applicazione è lineare (ESERCIZIO). Dunque, data una matrice $A \in M_{2\times 3}$ e date due basi, una nello spazio di partenza $\mathbb{R}^3$ e una nello spazio di arrivo $\mathbb{R}^2$, il prodotto righe per colonne permette di definire un'applicazione lineare da $\mathbb{R}^3$ in $\mathbb{R}^2$ ove i vettori di $\mathbb{R}^3$ sono espressi in coordinate nella base canonica e i vettori di $\mathbb{R}^2$ che si ottengono dopo il prodotto sono espressi in coordinate nella base canonica di $\mathbb{R}^2$. Osserviamo che le colonne della matrice sono le immagini dei vettori della base canonica dello spazio di partenza $\mathbb{R}^3$, espresse in coordinate nella base canonica dello spazio di arrivo $\mathbb{R}^2$. Infatti:

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \pi \\ e \end{pmatrix},$$

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \log 2 \\ -2 \end{pmatrix}.$$

Osservazione: In generale, ogni matrice $A \in M_{m\times n}$ permette di definire un'applicazione lineare

$$\begin{array}{ccc} L : \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

ove si sottointende che si sono scelte nello spazio di partenza e nello spazio di arrivo le basi canoniche.

Tuttavia, la stessa matrice permette di definire infinite applicazioni lineari poiché è possibile scegliere un numero infinito di spazi vettoriali sia in partenza che in arrivo e in ogni spazio si possono scegliere infinite basi.

Esempio: Riprendiamo l'esempio precedente:

$$A = \begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix}$$

Consideriamo ora $V = \mathrm{vect}\{1, e^x, e^{-x}\}$ e $W = \mathbb{R}_1[X]$. Si ha $\dim V = 3$ e $\dim W = 2$. Ora, scegliamo $B_V = \{1, e^x, e^{-x}\}$ come base in V e $B_W = \{1, X\}$ come base in W. Dunque, per ogni $v \in V$ si ha $v = \alpha 1 + \beta e^x + \gamma e^{-x}$ e il vettore v è individuato univocamente dalle coordinate (α, β, γ) nella base B_V. Analogamente, ogni $w \in W$ è tale che $w = a + bX$ e quindi w è individuato univocamente dalle coordinate (a, b) nella base B_W. Ora, definiamo l'applicazione lineare

$$L : V \longrightarrow W$$

nel modo seguente:

$$L(\alpha 1 + \beta e^x + \gamma e^{-x}) = a + bX$$

dove

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 3\alpha + \pi\beta + (\log 2)\gamma \\ -\alpha + e\beta - 2\gamma \end{pmatrix}$$

Dunque l'applicazione è:

$$L(\alpha 1 + \beta e^x + \gamma e^{-x}) = (3\alpha + \pi\beta + (\log 2)\gamma)1 + (-\alpha + e\beta - 2\gamma)X \tag{1.1}$$

Anche in questo caso, le colonne della matrice sono le coordinate delle immagini tramite l'applicazione L dei vettori della base B_V nella base B_W. Infatti: $1 = 1 + 0e^x + 0e^{-x}$, dunque le coordinate di 1 nella base B_V sono $(1, 0, 0)$. Ora

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

e tramite l'espressione trovata in (1.1)

$$L(1) = 3 - X.$$

Analogamente e^x ha coordinate $(0, 1, 0)$,

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \pi \\ e \end{pmatrix}$$

e tramite l'espressione trovata in (1.1)

$$L(e^x) = \pi + eX;$$

infine e^{-x} ha coordinate $(0, 0, 1)$,

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

e tramite l'espressione trovata in (1.1)

$$L(e^x) = 2 - 2X.$$

Ora, ad esempio, quanto vale $L(2-4e^x)$? Basta trovare le coordinate di $2-4e^x$ nella base B_V, che sono $(2,-4,0)$, calcolare

$$\begin{pmatrix} 3 & \pi & \log 2 \\ -1 & e & -2 \end{pmatrix} \begin{pmatrix} 2 \\ -4 \\ 0 \end{pmatrix} = \begin{pmatrix} 6-4\pi \\ -2-4e \end{pmatrix}$$

e quindi si ha $L(2-4e^x) = (6-4\pi)1 + (-2-4e)X$.

Vale il seguente:

Teorema 27 (di rappresentazione delle applicazioni lineari)
Siano V, W spazi vettoriali reali, sia $L : V \longrightarrow W$ lineare e siano $\dim V = n$ e $\dim W = m$. Siano poi $B_v = \{v_1, \dots, v_n\}$ una base di V e $B_W = \{w_1, \dots, w_m\}$ una base di W. Allora esiste un'unica matrice $A \in M_{m\times n}(\mathbb{R})$ che rappresenta L, cioè se

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

e

$$L(v) = \beta_1 w_1 + \cdots + \beta_m w_m$$

allora

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} = A \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Il teorema vale in modo totalmente analogo tra spazi vettoriali complessi con una matrice di rappresentazione a coefficienti complessi.

Non dimostriamo con precisione questo teorema, ma diamo semplicemente l'espressione della matrice di rappresentazione.

Se

$$\begin{aligned} L(v_1) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\ L(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\ &\vdots \\ L(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m \end{aligned}$$

allora si ha

$$A = \Big(L(v_1)\, L(v_2) \, \dots \, L(v_n) \Big) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Esempi:

1. Consideriamo

$$\begin{aligned} L : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto \begin{pmatrix} 2x-3y \\ x \\ 3y \end{pmatrix} \end{aligned}$$

Vogliamo determinare la matrice che rappresenta l'applicazione L scegliendo le basi canoniche sia in $\mathbb{R}^2$ che in $\mathbb{R}^3$. Dobbiamo determinare $L(1,0)$ e $L(0,1)$, in coordinate nella base canonica di $\mathbb{R}^3$. Si ha

$$L\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad L\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 3 \end{pmatrix}$$

da cui la matrice A risulta

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

Effettivamente si ha per ogni $(x, y) \in \mathbb{R}^2$:

$$\begin{pmatrix} 2 & -3 \\ 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ x \\ 3y \end{pmatrix} = L\left(\begin{pmatrix} x \\ y \end{pmatrix}\right).$$

2. Consideriamo ora $P_2 = \{a + bX + cX^2, \quad a, b, c \in \mathbb{R}\}$ e $P_1 = \{\alpha + \beta X, \quad \alpha, \beta \in \mathbb{R}\}$ e l'applicazione

$$\begin{aligned} L : P_2 &\longrightarrow P_1 \\ a + bX + cX^2 &\longmapsto (a + 2b) + (b + c)X. \end{aligned}$$

L'applicazione L è lineare (ESERCIZIO). Vogliamo determinare la matrice A che rappresenta l'applicazione L avendo scelto le basi canoniche $B_2 = \{1, X, X^2\}$ in P_2 e $B_1 = \{1, X\}$ in P_1. Determiniamo quindi $L(1)$, $L(X)$ e $L(X^2)$ in coordinate rispetto alla base canonica B_1. Si ha

$$L(1) = 1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad L(X) = 2 + X = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad L(X^2) = X = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

da cui

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

Infatti, per ogni polinomio $a + bX + cX^2 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, si ha in coordinate rispetto alla base canonica B_1

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + 2b \\ b + c \end{pmatrix}$$

che corrisponde al polinomio

$$(a + 2b) + (b + c)X.$$

3. **ESERCIZIO** Sia $V = \text{vect}\{e^x, e^{2x}\}$ e sia $W = \mathbb{R}^3$. Consideriamo l'applicazione

$$\begin{aligned} L : V &\longrightarrow W \\ ae^x + be^{2x} &\longmapsto \begin{pmatrix} a + b \\ a - b \\ a \end{pmatrix} \end{aligned}$$

   a) Mostrare che $B_V = \{e^x, e^{2x}\}$ è una base di V;
   b) mostrare che L è lineare;
   c) determinare $\operatorname{Ker} L$ e $\operatorname{Im} L$;
   d) determinare la matrice ${}_{B_3}[\mathrm{L}]_{B_V}$ che rappresenta l'applicazione lineare L nelle basi B_V e nella base canonica B_3 di $\mathbb{R}^3$;
   e) mostrare che $\tilde{B}_V = \{e^x + e^{2x}, e^x - e^{2x}\}$ è una base di V;
   f) determinare la matrice ${}_{B_3}[L]_{\tilde{B}_V}$ che rappresenta l'applicazione lineare L nelle basi $\tilde{B}_V$ e nella base canonica di $\mathbb{R}^3$.

1.3.3 Composizione di applicazioni lineari. Applicazioni invertibili.

Siano V, W e Z tre spazi vettoriali reali e siano $L_1 : V \longrightarrow W$, $L_2 : W \longrightarrow Z$ applicazioni lineari. Possiamo fare agire queste applicazioni una dopo l'altra, costruendo una nuova applicazione:

$$\begin{array}{rcl} L_2 \circ L_1 : V & \longrightarrow W \longrightarrow & Z \\ v & \longmapsto L_1(v) \longmapsto & L_2(L_1(v)) \end{array}$$

Vale la seguente

Proposizione 28
Se L_1 e L_2 sono lineari, anche $L_2 \circ L_1$ è lineare.

Dimostrazione: Siano v_1, $v_2 \in V$ e $\lambda_1, \lambda_2 \in \mathbb{R}$ e consideriamo la combinazione lineare $\lambda_1 v_1 + \lambda_2 v_2 \in V$. Per la linearità di L_2 e di L_1 si ha

$$\begin{aligned} L_2 \circ L_1(\lambda_1 v_1 + \lambda_2 v_2) = L_2(L_1(\lambda_1 v_1 + \lambda_2 v_2)) &= L_2(\lambda_1 L_1(v_1) + \lambda_2 L_1(v_2)) \\ &= \lambda_1 L_2(L_1(v_1)) + \lambda_2 L_2(L_1(v_2)) \\ &= \lambda_1 L_2 \circ L_1(v_1) + \lambda_2 L_2 \circ L_1(v_2) \end{aligned}$$

dunque $L_2 \circ L_1$ è lineare. •

Esempi:

1. Siano

$$\begin{array}{rcl} L_1 : \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} 2x+y \\ x-y \\ 2x \end{pmatrix} \end{array}$$

e

$$\begin{array}{rcl} L_2 : \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} & \longmapsto & X+Y+Z \end{array}$$

Le applicazioni L_1 e L_2 sono lineari (ESERCIZIO). Consideriamo ora

$$\begin{array}{rcl} L_2 \circ L_1 : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} \overset{L_1}{\mapsto} \begin{pmatrix} 2x+y \\ x-y \\ 2x \end{pmatrix} & \overset{L_2}{\mapsto} & 2x+y+x-y+2x = 5x; \end{array}$$

è ancora un'applicazione lineare (ESERCIZIO).

2. Siano:

$$\begin{array}{rcl} L_1 : \mathbb{R}_1[X] & \longrightarrow & M_{2\times 2}(\mathbb{R}) \\ a+bX & \longmapsto & \begin{pmatrix} a & 2b \\ a+b & b \end{pmatrix} \end{array}$$

e

$$\begin{array}{rcl} L_2 : M_{2\times 2}(\mathbb{R}) & \longrightarrow & \mathbb{R}_3[X] \\ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} & \longmapsto & \alpha + \beta X + \gamma X^2 + \delta X^3 \end{array}$$

Le applicazioni L_1 e L_2 sono lineari (ESERCIZIO). Consideriamo ora

$$\begin{array}{rcl} L_2 \circ L_1 : \mathbb{R}_1[X] & \longrightarrow & \mathbb{R}_3[X] \\ a+bX \overset{L_1}{\mapsto} \begin{pmatrix} a & 2b \\ a+b & b \end{pmatrix} & \overset{L_2}{\mapsto} & a+(2b)X+(a+b)X^2+bX^3; \end{array}$$

è ancora un'applicazione lineare (ESERCIZIO).

Supponiamo ora che V, W e Z abbiano dimensione finita, sia $\dim V = n$, $\dim W = m$ e $\dim Z = p$ e siano B_V, B_W e B_Z basi scelte rispettivamente in V, W e Z. Consideriamo ora

$$_{B_W}[L_1]_{B_V}, \quad _{B_Z}[L_2]_{B_W}, \quad _{B_Z}[L_2 \circ L_1]_{B_V}$$

le matrici che rappresentano L_1, L_2 e $L_2 \circ L_1$ tra le basi scelte. Si ha il risultato seguente:

Proposizione 29
Si ha

$$_{B_Z}[L_2 \circ L_1]_{B_V} = {}_{B_Z}[L_2]_{B_W\,B_W}[L_1]_{B_V}.$$

Dimostrazione: (Non fatta a lezione – facoltativa) La dimostrazione è immediata, utilizzando l'associatività del prodotto righe per colonne. Infatti, sia $v \in V$ e consideriamo $[v]_{B_V}$ il vettore delle coordinate di v nella base B_V; allora sappiamo che il prodotto righe per colonne $_{B_W}[L_1]_{B_V}[v]_{B_V}$ fornisce le coordinate del vettore $L_1(v)$ nella base B_W e che per ogni $w \in W$, denotando con $[w]_{B_W}$ il vettore delle coordinate di w nella base W, il prodotto righe per colonne $_{B_Z}[L_2]_{B_W}[w]_{B_W}$ fornisce le coordinate del vettore $L_2(w)$ nella base B_Z. Allora:

$$\begin{aligned}
[(L_2 \circ L_1)(v)]_{B_Z} &= [L_2(L_1(v))]_{B_Z} \\
&= {}_{B_Z}[L_2]_{B_W}[L_1(v)]_{B_W} \\
&= {}_{B_Z}[L_2]_{B_W}\left({}_{B_W}[L_1]_{B_V}[v]_{B_V}\right) \\
&= \left({}_{B_Z}[L_2]_{B_W\,B_W}[L_1]_{B_V}\right)[v]_{B_V}
\end{aligned}$$

e per l'unicità della matrice che rappresenta un'applicazione lineare tra basi scelte, si ha che

$$_{B_Z}[L_2 \circ L_1]_{B_V} = {}_{B_Z}[L_2]_{B_W\,B_W}[L_1]_{B_V}.$$

•

Esempi:

1. Tornando all'esempio precedente, scegliamo in $\mathbb{R}^2$, $\mathbb{R}^3$ e $\mathbb{R}$ le basi canoniche. Si ha quindi, per $L_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$:
$$_3[L_1]_2 = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \in M_{3\times 2}(\mathbb{R}),$$
per $L_2 : \mathbb{R}^3 \longrightarrow \mathbb{R}$:
$$_1[L_2]_3 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \in M_{1\times 3}(\mathbb{R})$$
e per $L_2 \circ L_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$:
$$_3[L_2 \circ L_1]_2 = \begin{pmatrix} 5 & 0 \end{pmatrix} \in M_{1\times 2}(\mathbb{R}).$$
Eseguendo il prodotto righe per colonne:
$$_3[L_2]_{21}[L_1]_2 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 \end{pmatrix} =_1 [L_2 \circ L_1]_2.$$

2. Se consideriamo le applicazioni lineari dell'esempio 2 precedente e scegliamo $B_1 = \{1, X\}$ come base in $\mathbb{R}_1[X]$, $B_{2\times 2} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ come base in $M_{2\times 2}$ e

$B_3 = \{1, X, X^2, X^3\}$ come base in $\mathbb{R}_3[X]$, allora

$$_{B_{2\times 2}}[L_1]_{B_1} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$_{B_3}[L_2]_{B_{2\times 2}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$_{B_3}[L_2 \circ L_1]_{B_1} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

e infatti:

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

3. **ESERCIZIO:** Si considerino i tre spazi vettoriali reali:

$$P_1 = \{a + bX, \quad a, b \in \mathbb{R}\}, \quad P_2 = \{\alpha + \beta X + \gamma X^2, \quad \alpha, \beta, \gamma \in \mathbb{R}\}, \quad \mathbb{R}^2,$$

muniti delle basi canoniche $B_1 = \{1, X\}$, $B_2 = \{1, X, X^2\}$ e $B_c \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ e si considerino le due applicazioni lineari:

$$\begin{array}{rcl} L_1 : P_1 & \longrightarrow & P_2 \\ a + bX & \longmapsto & (a+b) + (a-b)X + 2bX^2 \end{array}, \qquad \begin{array}{rcl} L_2 : P_2 & \longrightarrow & \mathbb{R}^2 \\ \alpha + \beta X + \gamma X^2 & \longmapsto & \begin{pmatrix} \alpha + \beta \\ \alpha + \gamma \end{pmatrix} \end{array}.$$

Si costruisca poi l'applicazione composta

$$\begin{array}{rcl} L_2 \circ L_1 : P_1 & \xrightarrow{\hspace{6cm}} & \mathbb{R}^2 \\ a + bX & \xmapsto{L_1} (a+b) + (a-b)X + 2bX^2 \xmapsto{L_2} \begin{pmatrix} a+b+a-b \\ a+b+2b \end{pmatrix} = & \begin{pmatrix} 2a \\ a+3b \end{pmatrix} \end{array}$$

(a) Determinare le matrici

$$_{B_2}[L_1]_{B_1}, \quad _{B_c}[L_2]_{B_2}, \quad _{B_c}[L_2 \circ L_1]_{B_1}$$

delle applicazioni L_1, L_2 e $(L_2 \circ L_1)$ nelle basi canoniche;

(b) mostrare tramite il prodotto righe per colonne che vale la proposizione precedente.

Supponiamo ora che V e W siano spazi vettoriali della stessa dimensione (per esempio $\dim V = \dim W = n$) e siano B_V, B_W basi scelte in V e W. Supponiamo che $L_1 : V \longrightarrow W$ e $L_2 : W \longrightarrow V$ siano applicazioni lineari tali che $L_2 \circ L_1 = \mathrm{Id}_V$ e $L_1 \circ L_2 = \mathrm{Id}_W$ (ove Id_V è l'applicazione identità su V e Id_W è l'applicazione identità su W). Si ha, per definizione di applicazione inversa, $L_2 = L_1^{-1}$. Poiché l'applicazione Id è lineare chiaramente esiste una matrice che la rappresenta tra due basi scelte. Ora, poiché V e W hanno la stessa dimensione n, si ha

$$_{B_V}[\mathrm{Id}]_{B_V} =_{B_W} [\mathrm{Id}]_{B_W} = I_n.$$

Dunque, se ${}_{B_W}[L_1]_{B_V}$ è la matrice che rappresenta L_1 e ${}_{B_V}[L_2]_{B_W}$ è la matrice che rappresenta L_2 nelle basi scelte, si ha

$$I_n = {}_{B_W}[\mathrm{Id}]_{B_W} = {}_{B_W}[L_1]_{B_V}\,{}_{B_V}[L_2]_{B_W}$$

cioè, essendo ${}_{B_V}[L_2]_{B_W} = {}_{B_V}[L_1^{-1}]_{B_W}$

$${}_{B_V}[L_1^{-1}]_{B_W} = \left({}_{B_W}[L_1]_{B_V}\right)^{-1}.$$

1.3.4 Cambiamenti di coordinate.

Consideriamo ora uno spazio vettoriale V di dimensione n e due basi diverse $B = \{e_1, \dots, e_n\}$ e $\tilde{B} = \{\tilde{e}_1, \dots, \tilde{e}_n\}$; consideriamo poi l'applicazione identità su V:

$$\begin{array}{rcl} \mathrm{Id}_V : V & \longrightarrow & V \\ v & \longmapsto & v. \end{array}$$

Poiché è un'applicazione lineare, fissata una base nello spazio di partenza ed una base nello spazio di arrivo, è possibile rappresentare Id_V tramite una matrice. Se si sceglie la stessa base (sia essa B oppure $\tilde{B}$) sia in partenza che in arrivo, la matrice rappresentativa (${}_B[\mathrm{Id}]_B$ oppure ${}_{\tilde{B}}[\mathrm{Id}]_{\tilde{B}}$) è chiaramente la matrice I_n. Se però si scelgono due basi diverse in partenza ed in arrivo, allora la matrice rappresentativa cambia. Infatti:

$${}_{\tilde{B}}[\mathrm{Id}]_B = \begin{pmatrix} [e_1]_{\tilde{B}} & [e_2]_{\tilde{B}} & \dots & [e_n]_{\tilde{B}} \end{pmatrix}$$

ove $[e_j]_{\tilde{B}}$ è il vettore di $\mathbb{R}^n$ delle coordinate dell'elemento $e_j \in B$ nella nuova base $\tilde{B}$.
Questa matrice fornisce il cambiamento di coordinate di un vettore di V dalla base B alla base $\tilde{B}$.

Analogamente

$${}_B[\mathrm{Id}]_{\tilde{B}} = \begin{pmatrix} [\tilde{e}_1]_B & [\tilde{e}_2]_B & \dots & [\tilde{e}_n]_B \end{pmatrix}$$

ove $[\tilde{e}_j]_B$ è il vettore di $\mathbb{R}^n$ delle coordinate dell'elemento $\tilde{e}_j \in \tilde{B}$ nella base B.
Questa matrice fornisce il cambiamento di coordinate di un vettore di V dalla base $\tilde{B}$ alla base B.

Poiché

$${}_B[\mathrm{Id}]_{\tilde{B}}\,{}_{\tilde{B}}[\mathrm{Id}]_B = {}_B[\mathrm{Id}]_B = I_n$$

si ha che

$${}_{\tilde{B}}[\mathrm{Id}]_B = \left({}_B[\mathrm{Id}]_{\tilde{B}}\right)^{-1}.$$

Dunque, è sufficiente determinare direttamente una delle due matrici ed invertirla per ottenere l'altra.

Esempio: In $\mathbb{R}^2$ scegliamo le due basi $B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ e $\tilde{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$. Ricerchiamo la matrice che rappresenta l'applicazione identità scegliendo $\tilde{B}$ nello spazio $\mathbb{R}^2$ di partenza e B nello spazio $\mathbb{R}^2$ di arrivo:

$${}_B[\mathrm{Id}]_{\tilde{B}} = \begin{pmatrix} [\tilde{e}_1]_B & [\tilde{e}_2]_B \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B & \begin{pmatrix} 2 \\ 1 \end{pmatrix}_B \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Questa matrice è facile da trovare, poiché basta allineare come colonne i vettori della nuova base $\tilde{B}$ nelle coordinate canoniche. Determiniamo ora ${}_{\tilde{B}}[\mathrm{Id}]_B$ dapprima direttamente, calcolando le coordinate dei vettori della base canonica B nella nuova base $\tilde{B}$ e poi utilizzando la proprietà

$${}_{\tilde{B}}[\mathrm{Id}]_B = \left({}_B[\mathrm{Id}]_{\tilde{B}}\right)^{-1}.$$

Si ha

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \end{pmatrix} \iff \begin{cases} \alpha = -\dfrac{1}{3} \\ \beta = \dfrac{2}{3} \end{cases}, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \end{pmatrix} \iff \begin{cases} \alpha = \dfrac{2}{3} \\ \beta = -\dfrac{1}{3} \end{cases},$$

da cui

$$_{\tilde{B}}[\mathrm{Id}]_B = \big([e_1]_{\tilde{B}} \quad [e_2]_{\tilde{B}}\big) = \left(\begin{pmatrix}1\\0\end{pmatrix}_{\tilde{B}} \quad \begin{pmatrix}0\\1\end{pmatrix}_{\tilde{B}}\right) = \begin{pmatrix}-\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{pmatrix}.$$

Verifichiamo ora la proprietà $_{\tilde{B}}[\mathrm{Id}]_B = ({}_B[\mathrm{Id}]_{\tilde{B}})^{-1}$:

$$({}_B[\mathrm{Id}]_{\tilde{B}})^{-1} = \frac{1}{(-3)}\begin{pmatrix}1 & -2\\ -2 & 1\end{pmatrix} = \begin{pmatrix}-\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{pmatrix} =_{\tilde{B}} [\mathrm{Id}]_B.$$

Osservazione importante: Se in $\mathbb{R}^n$ consideriamo la base canonica B e un'altra base $\tilde{B}$, è facile costruire la matrice $_B[\mathrm{Id}]_{\tilde{B}}$ del cambiamento di coordinate dalla nuova base $\tilde{B}$ alla base canonica B che si ottiene allineando come colonne i vettori della base $\tilde{B}$. La matrice $_{\tilde{B}}[\mathrm{Id}]_B$ del cambiamento di coordinate dalla base canonica B alla nuova base $\tilde{B}$ si ottiene invertendo la precedente matrice $_B[\mathrm{Id}]_{\tilde{B}}$.

Ora torniamo alle applicazioni lineari e consideriamo solo $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Chiaramente, la matrice che rappresenta l'applicazione L cambia se scegliamo basi diverse nello spazio di partenza e in quello di arrivo. Se chiamiamo B_n e $\tilde{B}_n$ due basi di $\mathbb{R}^n$, B_m e $\tilde{B}_m$ due basi di $\mathbb{R}^m$ e $_{B_m}[L]_{B_n}$, $_{\tilde{B}_m}[L]_{\tilde{B}_n}$ le matrici che rappresentano la stessa applicazione L tra le basi B (sia in partenza che in arrivo) e $\tilde{B}$ (sia in partenza che in arrivo) rispettivamente, avremo grazie ai cambiamenti di coordinate:

$$_{\tilde{B}_m}[L]_{\tilde{B}_n} = {}_{\tilde{B}_m}[\mathrm{Id}]_{B_m\,B_m}[L]_{B_n\,B_n}[\mathrm{Id}]_{\tilde{B}_n}$$

Esempio: Consideriamo l'applicazione

$$\begin{aligned} L : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix}x\\y\end{pmatrix} &\longmapsto \begin{pmatrix}3x+2y\\ x-2y\end{pmatrix}\end{aligned}$$

e scegliamo in $\mathbb{R}^2$ le due basi dell'esempio precedente:

$$B = \{e_1, e_2\} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\},$$

$$\tilde{B} = \{\tilde{e}_1, \tilde{e}_2\} = \left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}\right\}.$$

Determiniamo dapprima direttamente le matrici $_B[L]_{\tilde{B}}$ e $_{\tilde{B}}[L]_{\tilde{B}}$ e successivamente verificheremo che

$$_{\tilde{B}}[L]_{\tilde{B}} = {}_{\tilde{B}}[\mathrm{Id}]_{B\,B}[L]_{B\,B}[\mathrm{Id}]_{\tilde{B}}.$$

Facendo i conti, si ha

$$\begin{aligned}
_B[L]_B &= \big([L(e_1)]_B \quad [L(e_2)]_B\big) = \begin{pmatrix}3 & 2\\ 1 & -2\end{pmatrix}\\
_B[L]_{\tilde{B}} &= \big([L(\tilde{e}_1)]_B \quad [L(\tilde{e}_2)]_B\big) = \begin{pmatrix}7 & 8\\ -3 & 0\end{pmatrix}\\
_{\tilde{B}}[L]_{\tilde{B}} &= \big([L(\tilde{e}_1)]_{\tilde{B}} \quad [L(\tilde{e}_2)]_{\tilde{B}}\big) = \begin{pmatrix}-\frac{13}{3} & -\frac{8}{3}\\ \frac{17}{3} & \frac{16}{3}\end{pmatrix}.
\end{aligned}$$

Nell'esempio precedente abbiamo calcolato:

$$_B[\mathrm{Id}]_{\tilde{B}} = \begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}, \qquad _{\tilde{B}}[\mathrm{Id}]_B = \begin{pmatrix}-\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{pmatrix}.$$

Come previsto, si ha:

$$_{\tilde{B}}[\mathrm{Id}]_{B\,B}[L]_{B\,B}[\mathrm{Id}]_{\tilde{B}} = \begin{pmatrix}-\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{pmatrix}\begin{pmatrix}3 & 2\\ 1 & -2\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix} = \begin{pmatrix}-\frac{13}{3} & -\frac{8}{3}\\ \frac{17}{3} & \frac{16}{3}\end{pmatrix} =_{\tilde{B}} [L]_{\tilde{B}}.$$

1.3.5 Matrici simili, diagonalizzabili.

Consideriamo ora solo matrici quadrate (reali o complesse) ed applicazioni lineari $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, oppure $L : \mathbb{C}^n \longrightarrow \mathbb{C}^n$.

Definizione 30
Diremo che due matrici A e $\tilde{A} \in M_{n\times n}(\mathbb{R})$ sono **simili** *o* **equivalenti** *se esiste una matrice $S \in M_{n\times n}(\mathbb{R})$ invertibile tale che*

$$\tilde{A} = S^{-1}AS.$$

Analogamente, due matrici A e $\tilde{A} \in M_{n\times n}(\mathbb{C})$ sono **simili** *o* **equivalenti** *se esiste una matrice $S \in M_{n\times n}(\mathbb{C})$ invertibile tale che*

$$\tilde{A} = S^{-1}AS.$$

Osservazione importante: Chiaramente, le matrici che rappresentano una stessa applicazione lineare tra $\mathbb{R}^n$ e $\mathbb{R}^n$ (oppure tra $\mathbb{C}^n$ e $\mathbb{C}^n$) tra due basi diverse (con la stessa base in partenza ed in arrivo) sono simili. Reciprocamente, date due matrici simili A e $\tilde{A}$, possiamo sempre pensarle come matrici che rappresentano la stessa applicazione lineare tra due basi diverse. Basta considerare ad esempio l'applicazione lineare

$$\begin{array}{rcl} L : \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ v & \longmapsto & A\,v \end{array}$$

la cui matrice rappresentativa nella base canonica B_c (sia in partenza che in arrivo) è proprio

$$_{B_c}[L]_{B_c} = A$$

(si noti che un vettore $v \in \mathbb{R}^n$ coincide con il vettore delle sue coordinate nella base canonica $[v]_{B_c}$!). Allora, poiché

$$\tilde{A} = S^{-1}AS$$

si ha

$$\tilde{A} =_{\tilde{B}} [L]_{\tilde{B}}$$

ove abbiamo definito $\tilde{B}$ la base in $\mathbb{R}^n$ per la quale si ha

$$S =_{B_c} [\mathrm{Id}]_{\tilde{B}}.$$

In questo modo

$$\tilde{A} = S^{-1}AS \iff {}_{\tilde{B}}[L]_{\tilde{B}} = {}_{\tilde{B}}[\mathrm{Id}]_{B_c B_c}[L]_{B_c B_c}[\mathrm{Id}]_{\tilde{B}}.$$

Definizione 31
Diremo che una matrice $A \in M_{n\times n}(\mathbb{R})$ è **diagonalizzabile su** $\mathbb{R}$ *se esistono una matrice $S \in M_{n\times n}(\mathbb{R})$ invertibile ed una matrice diagonale $\Lambda \in M_{n\times n}(\mathbb{R})$ tali che*

$$A = S\Lambda S^{-1}$$

Diremo che $A \in M_{n\times n}(\mathbb{C})$ è **diagonalizzabile su** $\mathbb{C}$ *se esistono una matrice $S \in M_{n\times n}(\mathbb{C})$ invertibile ed una matrice diagonale $\Lambda \in M_{n\times n}(\mathbb{C})$ tali che*

$$A = S\Lambda S^{-1}.$$

La matrice S verrà detta **matrice di passaggio**.

Osservazioni:

1. Poiché $\mathbb{R} \subset \mathbb{C}$ ogni matrice $A \in M_{n\times n}(\mathbb{R})$ è anche in $M_{n\times n}(\mathbb{C})$.

2. La proprietà "A è diagonalizzabile su $\mathbb{R}$" si può rileggere in termini dell'applicazione lineare $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ di cui A è la matrice rappresentativa nelle basi canoniche: A è diagonalizzabile su $\mathbb{R}$ se esiste una base $\tilde{B}$ in $\mathbb{R}^n$ tale che la matrice che rappresenta L nella base $\tilde{B}$ sia in partenza che in arrivo è diagonale. Analogamente, la proprietà "A è diagonalizzabile su $\mathbb{C}$" si può rileggere in termini dell'applicazione lineare $L : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ di cui A è la matrice rappresentativa nelle basi canoniche: A è diagonalizzabile su $\mathbb{C}$ se esiste una base $\tilde{B}$ in $\mathbb{C}^n$ tale che la matrice che rappresenta L nella base $\tilde{B}$ sia in partenza che in arrivo è diagonale. (Si osservi che la definizione di $L : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ è la stessa che nel caso reale:

$$\begin{array}{rcl} L : \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ v & \longmapsto & A\,v \end{array}$$

dove però il vettore v ha coordinate complesse).

1.3.6 Autovalori, autovettori. Diagonalizzabilità.

Cominciamo con l'osservazione seguente: se un'applicazione lineare $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ è rappresentata in una determinata base $\tilde{B}$ (in partenza ed in arrivo) da una matrice diagonale

$$_{\tilde{B}}[L]_{\tilde{B}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

ove i $\lambda_j \in \mathbb{R}$ non sono necessariamente tutti distinti né non nulli, ciò significa che i vettori $\tilde{e}_j$ della base $\tilde{B}$ verificano la proprietà seguente:

$$L(\tilde{e}_j) = \lambda_j \tilde{e}_j$$

cioè ogni vettore della base $\tilde{B}$ è trasformato dall'applicazione L in un vettore ad esso parallelo. Questa proprietà NON dipende dalla matrice che rappresenta l'applicazione lineare ma dall'applicazione stessa.

Osservazione: Se la matrice che rappresenta l'applicazione L è diagonale, l'applicazione assume un aspetto particolarmente semplice in coordinate rispetto alla base $\tilde{B}$: se

$$[v]_{\tilde{B}} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{pmatrix},$$

allora

$$[L\,(v)]_{\tilde{B}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\tilde{x}_1 \\ \lambda_2\tilde{x}_2 \\ \vdots \\ \lambda_n\tilde{x}_n \end{pmatrix}.$$

Vedremo che una matrice $A \in M_{n\times n}(\mathbb{R})$ può non essere diagonalizzabile su $\mathbb{R}$ ma esserlo su $\mathbb{C}$, oppure può non esserlo neanche in $\mathbb{C}$.

Ora, sia $A \in M_{n\times n}(\mathbb{R})$ e sia L l'applicazione lineare da $\mathbb{R}^n$ in $\mathbb{R}^n$ di cui A è la matrice rappresentativa nella base canonica; una **condizione necessaria e sufficiente** perché la matrice A sia diagonalizzabile su $\mathbb{R}$ (cioè sia simile a una matrice diagonale a coefficienti reali) è che esista una base $\tilde{B}$ di $\mathbb{R}^n$ tale che per ogni $v \in \tilde{B}$ esista $\lambda \in \mathbb{R}$ tale che

$$L(v) = \lambda v$$

cioè
$$Av = \lambda v.$$
(Analogamente, sia $A \in M_{n\times n}(\mathbb{C})$ e sia $L : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ l'applicazione lineare da $\mathbb{C}^n$ in $\mathbb{C}^n$ di cui A è la matrice rappresentativa nella base canonica. Una **condizione necessaria e sufficiente** perché la matrice A sia diagonalizzabile su $\mathbb{C}$ (cioè sia simile a una matrice diagonale a coefficienti complessi) è che esista una base $\tilde{B}$ di $\mathbb{C}^n$ tale che per ogni $v \in \tilde{B}$ esista $\lambda \in \mathbb{C}$ tale che
$$L(v) = \lambda v$$
cioè
$$Av = \lambda v.)$$
Sia dunque $A \in M_{n\times n}(\mathbb{R})$; cerchiamo di caratterizzare i numeri reali λ tali che esistano dei vettori $v \in \mathbb{R}^n$, <u>non nulli</u> tali che
$$Av = \lambda v$$
o equivalentemente, essendo $v = I_n v$,
$$(A - \lambda I_n)v = \underline{0}$$
(ci faremo contemporaneamente anche la stessa domanda in $\mathbb{C}$).

Osservazione: Data una matrice quadrata $A \in M_{n\times n}(\mathbb{R})$, la funzione $\det(A - \lambda I_n)$ di variabile λ è un polinomio a coefficienti reali di grado n e analogamente, data una matrice quadrata $A \in M_{n\times n}(\mathbb{C})$, la funzione $\det(A - \lambda I_n)$ di variabile λ è un polinomio a coefficienti complessi di grado n.

Esempio: Se
$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$$
si ha
$$A - \lambda I_2 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{pmatrix}$$
per cui
$$\det(A - \lambda I_n) = (1-\lambda)^2 + 4 = \lambda^2 - 2\lambda + 5.$$

Definizione 32
Il polinomio $\det(A - \lambda I_n)$ *viene detto* **polinomio caratteristico** *e l'equazione* $\det(A - \lambda I_n) = 0$ *viene detta* **equazione caratteristica** *della matrice* A.

Osservazione: Poiché il polinomio caratteristico ha grado n, sappiamo che esso ha sempre n radici in $\mathbb{C}$ se contate con la loro molteplicità; inoltre, se ha coefficienti reali, le radici complesse sono coniugate a due a due e se il grado del polinomio è dispari esiste sempre almeno una radice reale.

Osservazione: Come varia il polinomio caratteristico nel passaggio da una matrice A ad una matrice $\tilde{A}$ ad essa equivalente?

Se due matrici $\tilde{A}$ e A sono simili, allora $\det \tilde{A} = \det A$. Infatti, per il teorema di Binet, si ha:
$$\det \tilde{A} = \det(S^{-1}AS) = \det(S^{-1}) \det A \det S = (\det S)^{-1} \det A \det S = \det A. \qquad (1.2)$$

Proposizione 33
Il polinomio caratteristico è invariante nel passaggio da una matrice ad una equivalente.

Dimostrazione: Poiché
$$\tilde{A} = S^{-1}AS$$
si ha
$$\begin{aligned}\det(\tilde{A} - \lambda I_n) &= \det(S^{-1}AS - S^{-1}\lambda I_n S)\\ &= \det(S^{-1}(A - \lambda I_n)S) = \det S^{-1}\det(A - \lambda I_n)\det S = \det(A - \lambda I_n).\end{aligned}$$

•

In particolare, se $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ si ha
$$\det(A - \lambda I_2) = \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - (\operatorname{Tr} A) + \det A,$$
da cui anche la traccia (che è la somma degli elementi sulla diagonale principale) è invariante per passaggio ad una matrice equivalente, così come il determinante (come abbiamo ricordato nell'osservazione precedente).

Veniamo alle seguenti definizioni fondamentali.

Definizione 34
Data una matrice $A \in M_{n\times n}(\mathbb{R})$ (oppure $A \in M_{n\times n}(\mathbb{C})$) diremo che $\lambda \in \mathbb{R}$ (oppure $\lambda \in \mathbb{C}$) è un **autovalore reale (o complesso)** *per A se soddisfa l'equazione caratteristica*
$$\det(A - \lambda I_n) = 0.$$
Diremo che un vettore $v \in \mathbb{R}^n$ (oppure $v \in \mathbb{C}^n$) non nullo è un **autovettore reale (o complesso)** *corrispondente all'autovalore λ se risolve il sistema omogeneo*
$$(A - \lambda I_n)v = \underline{0}.$$
Chiameremo **autospazio relativo all'autovalore** λ *il sottospazio vettoriale (reale o complesso) generato dagli autovettori relativi a λ.*

Osservazioni:

1. Dalla proposizione precedente, si deduce che matrici simili hanno gli stessi autovalori.
2. Si ha che v è un autovettore corrispondente all'autovalore λ se e solo se v è non nullo e $v \in \operatorname{Ker}(L - \lambda \operatorname{Id})$, ove L è l'applicazione lineare di cui A è la matrice rappresentativa tra le basi canoniche.

Abbiamo tutti gli elementi per dimostrare il

Teorema 35
La matrice $A \in M_{n\times n}(\mathbb{R})$ è diagonalizzabile su $\mathbb{R}$ se e solo se esiste una base di autovettori reali di A. Chiamati $\tilde{B} = \{h_1, h_2, \ldots, h_n\}$ tale base e $\lambda_1, \lambda_2, \ldots, \lambda_n$ gli autovalori relativi ($\lambda_i \in \mathbb{R}$ non necessariamente distinti) si ha:
$$A = S\,\Lambda\,S^{-1}$$
ove
$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$
e
$$S =_B [\operatorname{Id}]_{\tilde{B}} = \begin{pmatrix} h_1 & h_2 & \ldots & h_n \end{pmatrix}.$$
(Analogamente, $A \in M_{n\times n}(\mathbb{C})$ è diagonalizzabile su $\mathbb{C}$ se e solo se esiste una base di autovettori complessi di A.)

Dimostrazione: Sia $A \in M_{n\times n}(\mathbb{R})$ e sia $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'applicazione lineare di cui A è la matrice rappresentativa tra le basi canoniche. Sappiamo che A è diagonalizzabile su $\mathbb{R}$ se e solo se esiste una base di vettori $\tilde{B} = \{h_1, \dots, h_n\} \subset \mathbb{R}^n$ tali che esistano $\lambda_i \in \mathbb{R}$ con

$$L(h_i) = \lambda_i h_i, \quad i = 1, \dots, n.$$

Per definizione di autovalore ed autovettore, i numeri reali λ_i sono gli autovalori della matrice A e h_i è un autovettore relativo a λ_i. La matrice che rappresenta l'applicazione lineare L nella base di autovettori $\tilde{B}$ sia nello spazio di partenza che nello spazio di arrivo è proprio Λ e quindi avremo:

$$A =_B [\mathrm{Id}]_{\tilde{B}}\, \Lambda\, _{\tilde{B}}[\mathrm{Id}]_B$$

La dimostrazione è totalmente analoga per il caso complesso. •

Questo teorema non è operativo ma fornisce solo una rilettura della definizione di diagonalizzabilità. Ci chiediamo ora: esistono delle condizioni sulla matrice A che possano far prevedere *a priori* se la matrice è diagonalizzabile su $\mathbb{R}$ oppure su $\mathbb{C}$, cioè se esiste una base di autovettori reali oppure complessi?

Definizione 36
Dato un autovalore λ_j di A, chiameremo

- **molteplicità algebrica** *di λ_j (e la indicheremo m_j) la molteplicità di λ_j come radice del polinomio caratteristico;*

- **molteplicità geometrica** *di λ_j (e la indicheremo d_j) la dimensione dell'autospazio relativo a λ_j.*

Facciamo alcuni esempi.

Esempi:

1. Sia

   $$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R});$$

   il suo polinomio caratteristico è

   $$\begin{aligned} \det(A - \lambda I_3) &= \det \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -2 & 1 & 2-\lambda \end{pmatrix} \\ &= -\lambda(-2\lambda + \lambda^2 - 1) - 2(1) = -\lambda^3 + 2\lambda^2 + \lambda - 2 = (\lambda - 2)(\lambda - 1)(\lambda + 1) \end{aligned}$$

   e l'equazione caratteristica è

   $$(\lambda - 2)(\lambda - 1)(\lambda + 1) = 0.$$

   Esistono tre autovalori reali distinti:

   $$\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 2.$$

   Ognuno di essi ha moleplicità algebrica 1.

   Ricerchiamo gli autovettori relativi ai tre autovalori. Sappiamo già che per ogni λ troveremo un'infinità di soluzioni che formano uno spazio vettoriale (si tratta del nucleo dell'applicazione rappresentata dalla matrice $A - \lambda I_3$ che ha, per definizione di autovalore, determinante nullo!!). Consideriamo $\lambda_1 = 1$; ricerchiamo $v \in \mathbb{R}^3$ tale che

$$
\begin{aligned}
(A - \lambda_1 I_3)v = \underline{0} &\iff \begin{pmatrix} -\lambda_1 & 1 & 0 \\ 0 & -\lambda_1 & 1 \\ -2 & 1 & 2-\lambda_1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{cases} -x + y = 0 \\ -y + z = 0 \\ -2x + y + z = 0 \end{cases} \\
&\iff \begin{cases} x = y \\ z = y. \end{cases}
\end{aligned}
$$

Dunque

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ y \\ y \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad y \in \mathbb{R}.
$$

Abbiamo trovato uno spazio vettoriale reale di dimensione 1 (è l'autospazio relativo a λ_1), cioè la molteplicità geometrica di λ_1 è 1. Possiamo scegliere come autovettore relativo all'autovalore λ_1 ogni vettore non nullo di questo spazio, per esempio

$$
v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.
$$

Se consideriamo ora $\lambda_2 = -1$, ricerchiamo $v \in \mathbb{R}^3$ tale che

$$
\begin{aligned}
(A - \lambda_2 I_3)v = \underline{0} &\iff \begin{pmatrix} -\lambda_2 & 1 & 0 \\ 0 & -\lambda_2 & 1 \\ -2 & 1 & 2-\lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{cases} x + y = 0 \\ y + z = 0 \\ -2x + y + 3z = 0 \end{cases} \\
&\iff \begin{cases} x = -y \\ z = -y. \end{cases}
\end{aligned}
$$

Dunque

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y \\ y \\ -y \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad y \in \mathbb{R}.
$$

Anche in questo caso, la molteplicità geometrica di λ_2 è 1; possiamo scegliere come autovettore relativo all'autovalore λ_2 ogni vettore non nullo precedente, per esempio

$$
v_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.
$$

Se consideriamo infine $\lambda_3 = 2$, ricerchiamo $v \in \mathbb{R}^3$ tale che

$$
\begin{aligned}
(A - \lambda_3 I_3)v = \underline{0} &\iff \begin{pmatrix} -\lambda_3 & 1 & 0 \\ 0 & -\lambda_3 & 1 \\ -2 & 1 & 2-\lambda_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&\iff \begin{cases} -2x + y = 0 \\ -2y + z = 0 \\ -2x + y = 0 \end{cases} \\
&\iff \begin{cases} y = 2x \\ z = 4x. \end{cases}
\end{aligned}
$$

Dunque

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 2x \\ 4x \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \quad x \in \mathbb{R}.
$$

Anche questa volta la molteplicità geometrica è 1; possiamo scegliere come autovettore relativo all'autovalore λ_3 ogni vettore non nullo del tipo precendente, per esempio

$$
v_3 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.
$$

Avevamo tre autovalori (reali) distinti, ciascuno quindi con molteplicità algebrica 1. Abbiamo mostrato che la molteplicità geometrica di ognuno è 1 ed abbiamo trovato tre autovettori (reali) indipendenti v_1 e v_2 e v_3 (ESERCIZIO).

2. Sia

$$
A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R});
$$

il suo polinomio caratteristico è

$$
\det(A - \lambda I_2) = \det \begin{pmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 + 4
$$

e l'equazione caratteristica è

$$
(1-\lambda)^2 + 4 = 0.
$$

Chiaramente non esistono autovalori reali. Invece, esistono due autovalori complessi coniugati:

$$
\lambda_1 = 1 + 2i, \quad \lambda_2 = 1 - 2i.
$$

Ciascuno di essi ha molteplicità algebrica 1. Non esistono autovettori reali; ricerchiamo gli autovettori complessi. Consideriamo $\lambda_1 = 1 + 2i$; ricerchiamo $v \in \mathbb{C}^2$ tale che

$$
\begin{aligned}
(A - \lambda_1 I_2)v = \underline{0} &\iff \begin{pmatrix} 1-(1+2i) & 2 \\ -2 & 1-(1+2i) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\iff \begin{pmatrix} -2i & 2 \\ -2 & -2i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

$$
\begin{cases} (-2i)x + 2y = 0 \\ -2x + (-2i)y = 0 \end{cases}
$$

$$
\iff y = ix.
$$

Dunque
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ ix \end{pmatrix} = x \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad x \in \mathbb{C}.$$

Abbiamo trovato uno spazio vettoriale complesso di dimensione 1 (= molteplicità geometrica dell'autovalore); possiamo scegliere come autovettore relativo all'autovalore λ_1 ogni vettore non nullo di questo spazio, per esempio
$$v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

Se consideriamo ora $\lambda_2 = 1 - 2i$ avremo:
$$\begin{aligned} (A - \lambda_2 I_2)v = \underline{0} &\iff \begin{pmatrix} 1-(1-2i) & 2 \\ -2 & 1-(1-2i) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\iff \begin{pmatrix} 2i & 2 \\ -2 & 2i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$
$$\begin{cases} (2i)x + 2y = 0 \\ -2x + (2i)y = 0 \end{cases}$$
$$\iff y = -ix.$$

Dunque le soluzioni sono della forma:
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -ix \end{pmatrix} = x \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad x \in \mathbb{C}.$$

Abbiamo trovato uno spazio vettoriale complesso di dimensione 1 (= molteplicità geometrica dell'autovalore); possiamo scegliere come autovettore relativo all'autovalore λ_2 ogni vettore non nullo di questo spazio, per esempio
$$v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

Avevamo due autovalori (complessi) distinti ed abbiamo trovato due autovettori (complessi) indipendenti (ESERCIZIO) v_1 e v_2.

3. Sia
$$A = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \in M_{2\times 2}(\mathbb{R});$$
il suo polinomio caratteristico è
$$\det(A - \lambda I_2) = \det \begin{pmatrix} 2-\lambda & 0 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2$$
e l'equazione caratteristica è
$$(2-\lambda)^2 = 0 \iff \lambda = 2.$$
Abbiamo trovato un solo autovalore che ha molteplicità algebrica 2. Ricerchiamo gli autovettori relativi a $\lambda = 2$ cioè risolviamo il sistema
$$(A - \lambda I_2)v = \underline{0} \iff \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff x = 0.$$
Dunque le soluzioni sono della forma:
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} = y \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad y \in \mathbb{R}.$$

Abbiamo trovato uno spazio vettoriale reale di dimensione 1 (= molteplicità geometrica dell'autovalore); possiamo scegliere come autovettore relativo all'autovalore $\lambda = 2$ ogni vettore non nullo di questo spazio, per esempio

$$v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Avevamo due autovalori reali coincidenti (cioè un autovalore di molteplicità algebrica 2) ed abbiamo trovato un solo autovettore reale indipendente (nel senso che ogni altro autovettore è un suo multiplo), cioè la molteplicità geometrica dell'autovalore è 1.

La situazione degli esempi precedenti è coerente con il seguente risultato (che non dimostriamo):

Teorema 37
a) Autovettori relativi ad autovalori distinti sono indipendenti;

b) la molteplicità geometrica di un autovalore è sempre minore o uguale alla molteplicità algebrica:

$$d_j \leq m_j.$$

Definizione 38
Diremo che λ_j è un autovalore regolare se $d_j = m_j$.

Osservazione: Ogni autovalore semplice (cioè di molteplicità algebrica 1) è regolare.

Terminiamo con il seguente risultato fondamentale sulla diagonalizzazione.

Teorema 39
Una matrice $A \in M_{n\times n}(\mathbb{R})$ è diagonalizzabile su $\mathbb{R}$ se e solo se ha n autovalori reali (se contati con la loro molteplicità) e sono tutti regolari. Una matrice $A \in M_{n\times n}(\mathbb{C})$ è diagonalizzabile su $\mathbb{C}$ se e solo se tutti i suoi autovalori (reali e complessi) sono regolari. Data la matrice A, data la matrice diagonale Λ che ha sulla diagonale principale gli autovalori λ_i e data la matrice di passaggio S che ha come colonne gli autovettori h_i opportunamente ordinati secondo l'ordine degli autovalori, si ha

$$A = S\Lambda S^{-1}$$

Capitolo 2

Equazioni e sistemi di equazioni differenziali ordinarie

Testi consigliati: C. D. Pagani e S. Salsa: Analisi Matematica 2 (Ed. Zanichelli); A. Bacciotti e F. Ricci, Lezioni di Analisi Matematica 2 (Ed. Levrotto e Bella, Torino)

2.1 Introduzione e proprietà generali

Definizione 40
Una **equazione differenziale ordinaria** *(brevemente EDO) di ordine n è una relazione del tipo*

$$F(t, y(t), y'(t), \dots, y^{(n-1)}(t), y^{(n)}(t)) = 0$$

ove $F : U \subset \mathbb{R}^{n+2} \longrightarrow \mathbb{R}$ è una funzione di $n+2$ variabili.

Quindi, una EDO è una equazione in cui l'incognita è una funzione $y(t)$ di una variabile reale per la quale si richiede che per ogni t sia verificata una certa relazione fra t, $y(t)$ e tutte le derivate $y^{(k)}(t)$ fino a un certo ordine n.

Definizione 41
*L'***ordine** *di una EDO è il massimo ordine di derivazione che compare nell'equazione.*

Osservazioni:

1. Si chiama equazione differenziale **ordinaria** per distinguerla da un'equazione differenziale **alle derivate parziali**, in cui la funzione incognita è una funzione di più variabili reali $y(t_1, t_2, \dots, t_d)$ per la quale si richiede che per ogni $(t_1, t_2, \dots, t_d)$ sia verificata una certa relazione fra $(t_1, t_2, \dots, t_d)$, $y(t_1, t_2, \dots, t_d)$ e tutte le derivate parziali fino a un certo ordine n.

2. Di solito nell'espressione di una EDO si omette la dipendenza dalla variabile indipendente t nella funzione incognita y e nelle sue derivate e si omettono le parentesi nell'ordine di derivazione, a meno che dal contesto nasca una ambiguità tra derivazione e potenza; si scrive cioè
$$F(t, y, y', \dots, y^n) = 0.$$

Definizione 42
Una EDO si dice **in forma normale** *se è esplicitata rispetto alla derivata di ordine massimo*

$$y^n = f(t, y, y', \dots, y^{n-1})$$

ove $f : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$.

Si ha cioè $F(t, y, \dots, y^{n-1}, y^n) = y^n - f(t, y, \dots, y^{n-1})$.

Definizione 43
Una EDO si dice **autonoma** *se non vi è dipendenza esplicita dalla variabile indipendente t, cioè si ha*

$$F(y, y', \dots, y^n) = 0.$$

Esempi:

1. $y' = \cos t$ è una EDO del primo ordine in forma normale. Si ha

$$F(t, y, y') = y' - \cos t$$

dunque non si ha dipendenza esplicita da y. Non è autonoma. Si tratta in realtà del problema della ricerca delle primitive della funzione $\cos t$, per cui si hanno le infinite soluzioni della forma

$$y = \sin t + C, \quad C \in \mathbb{R}.$$

2. $y' = y$ è una EDO del primo ordine in forma normale, autonoma. Si ha

$$F(t, y, y') = y' - y$$

e le infinite soluzioni sono della forma

$$y = Ke^t, \quad K \in \mathbb{R}.$$

3. $y'' + y = 0$ è una EDO del secondo ordine, in forma normale (anche se dovrebbe essere piuttosto scritta $y'' = -y$), autonoma. Si ha

$$F(t, y, y', y'') = y'' + y$$

e le ∞^2 soluzioni sono della forma

$$y = C_1 \sin t + C_2 \cos t, \quad C_1, C_2 \in \mathbb{R}.$$

4. $2y'y'' = 1$ è una EDO del secondo ordine, NON in forma normale, autonoma. Possiamo riscriverla come

$$\frac{d}{dt}((y')^2) = 1$$

da cui

$$\begin{aligned}
&(y')^2 = t + C, \quad C \in \mathbb{R} \\
&y' = \pm\sqrt{t + C}, \quad t \geq -C \\
&y = \pm \int \sqrt{t + C}\, dt \\
&y = \pm\frac{2}{3}(t + C)^{\frac{3}{2}} + D, \quad t \geq -C, \quad C, D \in \mathbb{R}.
\end{aligned}$$

Definizione 44
Data una EDO

$$F(t, y, y', \dots, y^n) = 0, \quad F : U \subset \mathbb{R}^{n+2} \longrightarrow \mathbb{R}$$

chiameremo **soluzione** *o* **integrale** *dell'EDO ogni funzione $\phi(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ tale che:*

i) I è un intervallo;

ii) ϕ è differenziabile n volte su I;

iii) per ogni $t \in I$ *si ha* $(t, \phi(t), \phi'(t), \ldots, \phi^n(t)) \in U \subset \mathbb{R}^{n+2}$*;*

iv) per ogni $t \in I$ *si ha* $F(t, \phi(t), \phi'(t), \ldots, \phi^n(t)) = 0$.

Definizione 45
Chiameremo **integrale generale** *di una equazione differenziale l'insieme di tutte le soluzioni.*

Invece di una equazione che abbia come incognita una funzione scalare $y(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ si può considerare una equazione che abbia come incognita una funzione vettoriale $\underline{y}(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$. Si tratta in questo caso di un sistema di equazioni, in cui le incognite sono le n componenti della funzione $\underline{y} = (y_1, y_2, \ldots, y_n)$. Ci occuperemo solo di sistemi del primo ordine e in forma normale.

Cominciamo con un esempio.

Esempio: Consideriamo il sistema:

$$\begin{cases} y_1' = y_1 y_2 + t e^{y_1} \\ y_2' = y_1 + \sin y_2 + t \end{cases}$$

In questo caso si ricercano le funzioni $(y_1(t), y_2(t))$ o meglio, si ricerca una funzione vettoriale $\underline{y}(t) = (y_1(t), y_2(t))$ tale che, per ogni t nel suo dominio di definizione verifichi:

$$\begin{cases} y_1'(t) = y_1(t) y_2(t) + t e^{y_1(t)} \\ y_2'(t) = y_1(t) + \sin y_2(t) + t. \end{cases}$$

È in forma normale in quanto la relazione è esplicitata rispetto al vettore delle derivate prime. La funzione vettoriale $\underline{f} = (f_1, f_2)$ che lega fra loro t, $y_1(t)$, $y_2(t)$ è

$$f_1(t, y_1, y_2) = y_1 y_2 + t e^{y_1} : \mathbb{R}^3 \longrightarrow \mathbb{R}$$
$$f_2(t, y_1, y_2) = y_1 + \sin y_2 + t : \mathbb{R}^3 \longrightarrow \mathbb{R}.$$

dunque $\underline{f} : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.

La definizione generale è la seguente.

Definizione 46
Un **sistema di equazioni differenziali ordinarie del primo ordine** *in forma normale è una equazione vettoriale*

$$\underline{y}' = \underline{f}(t, \underline{y})$$

ove $\underline{f} : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$*, cioè è un sistema del tipo*

$$\begin{cases} y_1' = f_1(t, y_1, y_2, \ldots, y_n) \\ y_2' = f_2(t, y_1, y_2, \ldots, y_n) \\ \vdots \\ y_n' = f_n(t, y_1, y_2, \ldots, y_n) \end{cases}$$

ove $\underline{y} = (y_1, y_2, \ldots, y_n)$ *e* $\underline{f} = (f_1, f_2, \ldots, f_n)$ *con* $f_i : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$.

Definizione 47
Un sistema si dice **autonomo** *se non vi è dipendenza esplicita dalla variabile indipendente* t.

Esempio:

$$\begin{cases} y_1' = y_2 \\ y_2' = -y_1. \end{cases}$$

È un sistema autonomo, $f_1(t, y_1, y_2) = y_2$ e $f_2(t, y_1, y_2) = -y_1$. Si verifica facilmente che la funzione vettoriale

$$\underline{\phi}(t) = (\phi_1(t), \phi_2(t)) = (\sin t, \cos t)$$

è soluzione su $(-\infty, +\infty)$.

Definizione 48
Dato un sistema in forma normale

$$\underline{y}' = \underline{f}(t, \underline{y}), \quad \underline{f} : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$$

chiameremo **soluzione** *o* **integrale** *del sistema ogni funzione vettoriale*

$$\underline{\phi}(t) = (\phi_1(t), \phi_2(t), \ldots, \phi_n(t)) : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$$

tale che:

i) I *è un intervallo;*

ii) $\underline{\phi}$ *è derivabile su* I*;*

iii) per ogni $t \in I$ *si ha* $(t, \phi_1(t), \phi_2(t), \ldots, \phi_n(t)) \in D \subset \mathbb{R}^{n+1}$*;*

iv) per ogni $t \in I$ *si ha* $\underline{\phi}'(t) = \underline{f}(t, \underline{\phi}(t))$*, cioè*

$$\begin{cases} \phi_1'(t) = f_1(t, \phi_1(t), \ldots, \phi_n(t)) \\ \phi_2'(t) = f_2(t, \phi_1(t), \ldots, \phi_n(t)) \\ \vdots \\ \phi_n'(t) = f_n(t, \phi_1(t), \ldots, \phi_n(t)). \end{cases}$$

Definizione 49
Chiameremo **integrale generale** *di un sistema di equazioni differenziali l'insieme di tutte le soluzioni.*

Definizione 50
Diremo che una EDO di ordine n

$$F(t, y, y', y'', \ldots, y^n) = 0$$

è **lineare** *se* $F(t, y, y', y'', \ldots, y^n)$ *è un polinomio di primo grado nelle variabili* $y, y', y'', \ldots, y^n$*, a coefficienti che sono funzioni della sola variabile* t*. Diremo che un sistema del primo ordine in forma normale*

$$\underline{y}' = \underline{f}(t, \underline{y})$$

è **lineare** *se* $\underline{f}(t, \underline{y})$ *è lineare in* $\underline{y}$*, cioè*

$$\underline{f}(t, \underline{y}) = A(t)\underline{y} + \underline{b}(t)$$

ove $A : I \subset \mathbb{R} \longrightarrow M_{n\times n}(\mathbb{R})$ *e* $\underline{b} : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$ *con* I *un intervallo aperto di* $\mathbb{R}$*.*

Esempi:

1.
$$(1+t^2)y''' + \sin t\, y'' + y' - e^t y + \cos t = 0$$
è una EDO del terzo ordine, non autonoma, lineare.

2.
$$y''' + ty''y' + t^2y^2 + 1 = 0$$
è una EDO del terzo ordine, non autonoma, non lineare.

3.
$$\begin{cases} y_1' = y_2 + t^2 \\ y_2' = -3t\, y_1 \end{cases}$$
è un sistema del primo ordine in forma normale, lineare; infatti possiamo riscriverlo come
$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -3t & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} t^2 \\ 0 \end{pmatrix}.$$

4.
$$\begin{cases} y_1' = y_1 y_2 + t^2 \\ y_2' = -3t\, y_1 \end{cases}$$
è un sistema del primo ordine in forma normale, non lineare.

Abbiamo visto le definizioni di una EDO di ordine n e di un sistema del primo ordine. Di fatto, c'è equivalenza tra i due, come afferma la prossima

Proposizione 51
L'EDO di ordine n
$$y^n = f(t, y, y', y'', \ldots, y^{n-1}) \tag{2.1}$$
è equivalente al sistema di n equazioni del primo ordine
$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \vdots \\ y_{n-1}' = y_n \\ y_n' = f(t, y_1, y_2, \ldots, y_n) \end{cases} \tag{2.2}$$
cioè:

a) se $\phi(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ è soluzione dell'equazione (2.1), allora la funzione vettoriale
$$\begin{cases} y_1 = \phi(t) \\ y_2 = \phi'(t) \\ \vdots \\ y_{n-1} = \phi^{n-2}(t) \\ y_n = \phi^{n-1}(t) \end{cases} : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$$
è soluzione del sistema (2.2);

b) se

$$\begin{cases} y_1 = \phi_1(t) \\ y_2 = \phi_2(t) \\ \vdots \\ y_{n-1} = \phi_{n-1}(t) \\ y_n = \phi_n(t) \end{cases} : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$$

è soluzione del sistema (2.2)*, con* $\phi_j(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ *per* $j = 1, \ldots, n$*, allora* $\phi_1(t)$ *è derivabile* n *volte su* I *ed è soluzione dell'equazione* (2.1).

Esempi:

1. Consideriamo l'equazione lineare del terzo ordine

$$y''' = t^2\, y'' - \sin t\, y' + e^t\, y + t.$$

Ponendo

$$\begin{cases} y_1 = y \\ y_2 = y' \\ y_3 = y'' \end{cases}$$

l'equazione si trasforma nel sistema lineare equivalente:

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ y_3' = t^2\, y_3 - \sin t\, y_2 + e^t\, y_1 + t. \end{cases}$$

2. L'equazione non lineare del secondo ordine

$$y'' = yy' + \cos t\, y$$

è equivalente al sistema non lineare

$$\begin{cases} y_1' = y_2 \\ y_2' = y_1 y_2 + \cos t\, y_1 \end{cases}$$

ottenuto ponendo

$$\begin{cases} y_1 = y \\ y_2 = y'. \end{cases}$$

Enunciamo infine un risultato sulla regolarità delle soluzioni di una EDO o di un sistema di EDO.

Proposizione 52 (Regolarità)
Si consideri il sistema

$$\underline{y}' = \underline{f}(t, \underline{y}), \quad \underline{f} : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n. \tag{2.3}$$

Se $\underline{f} \in C^k(D)$*, allora ogni soluzione del sistema* (2.3) *è di classe* C^{k+1} *sul suo dominio. In particolare, se* $\underline{f} \in C^\infty(D)$*, allora anche ogni soluzione è di classe* C^∞*.*

Analogamente, vale la proposizione seguente per le EDO di ordine n.

Proposizione 53
Sia

$$y^n = f(t, y, y', \ldots, y^{n-1}), \quad f : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}. \tag{2.4}$$

Se $f \in C^k(D)$*, allora ogni soluzione dell'EDO è di classe* C^{n+k} *sul suo dominio di definizione. In particolare, se* $f \in C^\infty(D)$*, allora anche ogni soluzione è di classe* C^∞*.*

2.2 Problema di Cauchy. Teoremi di esistenza ed unicità locale, globale e di dipendenza continua dai dati iniziali

2.2.1 Problema di Cauchy

Abbiamo visto che una EDO di ordine n (e equivalentemente un sistema di n equazioni del primo ordine) ha, in generale, infinite soluzioni dipendenti da un certo numero di parametri. Per restringere l'insieme delle soluzioni considerate, si possono imporre ad esse condizioni aggiuntive. Esistono principalmente due tipi di condizioni aggiuntive:

- condizioni iniziali, cioè si richiede che le soluzioni passino ad un dato istante t_0 per un dato punto ed abbiano in t_0 un certo numero di derivate assegnate, senza richiedere nulla sull'intervallo di definizione delle soluzioni;

- condizioni ai bordi, cioè si richiede che le soluzioni siano tutte definite in un certo intervallo $[a, b]$ ed abbiano valore prefissato per $t = a$ e $t = b$ (ed eventualmente anche valore di alcune derivate prefissato).

In queste note, tratteremo solo il problema delle condizioni iniziali, avvertendo però che nell'ambito delle equazioni differenziali alle derivate parziali le condizioni ai bordi (dette anche condizioni al contorno) portano a problematiche particolarmente importanti.

Definizione 54 (Problema di Cauchy)
Data una EDO di ordine n in forma normale

$$y^n = f(t, y, y', \dots, y^{n-1}), \quad f : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$$

e un vettore $(t_0, y_0, y_1, \dots, y_{n-1}) \in D$, chiameremo **soluzione del problema di Cauchy (o problema ai valori iniziali)**

$$\begin{cases} y^n = f(t, y, y', \dots, y^{n-1}) \\ y(t_0) = y_0 \\ y'(t_0) = y_1 \\ \vdots \\ y^{n-1}(t_0) = y_{n-1} \end{cases}$$

una soluzione $\phi(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}$ con I un intervallo, tale che $t_0 \in I$ e $\phi(t_0) = y_0$, $\phi'(t_0) = y_1$, $\dots$, $\phi^{n-1}(t_0) = y_{n-1}$.

Analogamente, dato un sistema

$$\underline{y}' = \underline{f}(t, \underline{y}), \quad \underline{f} : D \subset \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$$

ed un vettore $(t_0, \underline{y}_0) \in D$, chiameremo **soluzione del problema di Cauchy**

$$\begin{cases} \underline{y}' = \underline{f}(t, \underline{y}) \\ \underline{y}(t_0) = \underline{y}_0 \end{cases}$$

una soluzione $\underline{\phi}(t) : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$ con I un intervallo, tale che $t_0 \in I$ e $\underline{\phi}(t_0) = \underline{y}_0$.

Osservazione: È una soluzione di un problema di Cauchy ogni soluzione definita in un intervallo intorno al punto t_0, anche piccolo. Non si fa nessuna richiesta sull'ampiezza dell'intervallo di definizione; si tratta cioè di una nozione **locale**.

Esempio:

$$\begin{cases} y' = -y^2 \\ y(1) = 1 \end{cases}$$

Tutte le soluzioni dell'equazione proposta sono della forma:

$$y(t) = \frac{1}{t-c}, \quad c \in \mathbb{R}$$

e l'unica soluzione del problema di Cauchy proposto è

$$y(t) = \frac{1}{t}, \quad t \in (0, \infty).$$

Se avessimo considerato il problema di Cauchy

$$\begin{cases} y' = -y^2 \\ y(-1) = -1 \end{cases}$$

la soluzione sarebbe stata la stessa funzione, ma su un intervallo diverso:

$$y(t) = \frac{1}{t}, \quad t \in (-\infty, 0)$$

2.2.2 Teoremi di esistenza, di esistenza ed unicità locale, globale e di dipendenza continua dai dati iniziali

Grazie all'equivalenza tra una EDO di ordine n e un sistema del primo ordine, tratteremo solo il caso dei sistemi del primo ordine. Inoltre, per semplicità, enunceremo i risultati solo nel caso di una singola equazione scalare.

In questo paragrafo NON tratteremo il problema di descrivere metodi di risoluzione esplicita di equazioni differenziali o sistemi (come possono essere ad esempio i metodi di risoluzioni di equazioni a variabili separabili, lineari, di Bernoulli, di Riccati....), ma ci domandiamo invece quali informazioni possiamo dedurre *a priori* (cioè senza risolvere esplicitamente l'equazione o il sistema) riguardo alle seguenti questioni:

- Dato un problema di Cauchy, esiste almeno una soluzione?
- Ne esiste una sola?
- Possiamo stabilire *a priori* in quale intervallo è definita?
- Se variamo di poco il dato iniziale, la soluzione varia di poco?

Presentiamo ora tre risultati, validi per una EDO del primo ordine **in forma normale** $y' = f(t, y)$ (o per un sistema del primo ordine in forma normale $\underline{y}' = \underline{f}(t, \underline{y})$) :

- il primo fornisce una condizione sufficiente per garantire *a priori* l'**esistenza di almeno una soluzione** di un problema di Cauchy. Non esclude però che ci possano essere più soluzioni diverse dello stesso problema (teorema di esistenza di Peano);
- il secondo fornisce condizioni sufficienti per garantire *a priori* l'**esistenza di una sola soluzione locale** di un problema di Cauchy, senza però poter prevedere l'intervallo di definizione della soluzione (teorema di esistenza ed unicità locale);
- il terzo fornisce condizioni sufficienti per prevedere *a priori*, oltre all'**esistenza** ed unicità della soluzione, anche un **intervallo minimale** in cui tale soluzione sarà sicuramente definita (eventualmente poi l'intervallo reale di definizione potrebbe essere più grande) (teorema di esistenza ed unicità globale).

Tutti questi risultati si basano su proprietà di regolarità della funzione $f(t, y)$.

Teorema 55 (Esistenza di Peano)
Sia dato il problema di Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0. \end{cases}$$

Se f è continua in un intorno $U(t_0, y_0) \subset \mathbb{R}^2$ del punto (t_0, y_0), allora esiste un intorno $I_\delta = [t_0 - \delta, t_0 + \delta]$, $\delta > 0$ del punto t_0 nel quale è definita almeno una soluzione ϕ del problema di Cauchy.

Teorema 56 (Esistenza ed unicità locale)
Sia dato il problema di Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0. \end{cases}$$

Se $f \in C^1(U(t_0, y_0))$, allora esiste un intorno $I_\delta = [t_0 - \delta, t_0 + \delta]$, $\delta > 0$ del punto t_0 nel quale è definita un'unica soluzione ϕ del problema di Cauchy.

Prima di enunciare il terzo risultato, facciamo alcuni esempi.

Esempi:

1. Consideriamo il problema di Cauchy:

$$\begin{cases} y' = H(t) \\ y(0) = 1 \end{cases}$$

ove $H(t)$ è la funzione di Heaviside:

$$H(t) = \begin{cases} 1 & \text{se } t \geq 0 \\ 0 & \text{se } t < 0. \end{cases}$$

È un problema di ricerca di primitive. In questo caso si ha $f(t, y) = H(t)$ dunque non dipende esplicitamente da y. Ci chiediamo se f è continua in un intorno del punto $(t_0, y_0) = (0, 1)$. La funzione di Heaviside non è continua in $t = 0$, dunque la funzione di due variabili $f(t, y) = H(t)$ non è continua in $(0, y_0)$ per nessun $y_0 \in \mathbb{R}$. Dunque non possiamo applicare il teorema di Peano e non possiamo garantire a priori che esista almeno una soluzione del problema di Cauchy. Di fatto, una soluzione ϕ non esiste perché se così fosse, si avrebbe $\phi'(t) = H(t)$ e quindi in $t = 0$ la soluzione ϕ avrebbe per derivata una funzione con una discontinuità di tipo salto. Ma per il teorema di Lagrange questo non può avvenire. Il problema proposto non ha quindi soluzioni.

2. L'esempio che segue è particolarmente importante.

$$\begin{cases} y' = \sqrt[3]{y^2} \\ y(0) = 0. \end{cases}$$

Si tratta di una equazione differenziale autonoma, a variabili separabili. In questo caso $f(t, y) = \sqrt[3]{y^2} : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Si ha $f \in C(\mathbb{R}^2)$ ma non esiste la derivata parziale $\frac{\partial f}{\partial y}$ in tutti i punti $(t_0, 0)$, dunque in particolare **non** è di classe C^1 in un intorno del punto di passaggio $(0, 0)$. (Osserviamo infatti che $\frac{\partial f}{\partial y}(t, y) = \frac{2}{3} y^{-\frac{1}{3}}$ e questa non è definita per $y = 0$). Possiamo quindi applicare il teorema di Peano ed affermare a priori che esiste almeno una soluzione del problema di Cauchy, ma non possiamo applicare il teorema di esistenza ed unicità locale dunque non sappiamo a priori se tale soluzione sarà unica.

Sappiamo però risolvere esplicitamente questa equazione differenziale a variabili separabili. Una soluzione banale è la funzione costante $\phi_1(t) = 0$ per ogni $t \in \mathbb{R}$; inoltre, si ha la famiglia di soluzioni dipendenti da un parametro $y(t) = \left(\frac{1}{3}t + \frac{1}{3}C\right)^3$. Se poniamo ora $y(0) = 0$ troviamo un'altra soluzione $\phi_2(t) = \left(\frac{t}{3}\right)^3 : \mathbb{R} \longrightarrow \mathbb{R}$. Dunque il problema proposto ha almeno 2 soluzioni diverse.

D'altra parte, anche le funzioni

$$\phi_3(t) = \begin{cases} \left(\frac{t}{3}\right)^3 & t \geq 0 \\ 0 & t < 0, \end{cases}$$

e

$$\phi_4(t) = \begin{cases} \left(\frac{t}{3}\right)^3 & t \leq 0 \\ 0 & t > 0, \end{cases}$$

sono soluzioni ancora diverse.

3. Se consideriamo la stessa equazione differenziale con un altro problema di Cauchy

$$\begin{cases} y' = \sqrt[3]{y^2} \\ y(0) = 1 \end{cases}$$

la situazione è completamente diversa. Infatti, la funzione $f(t, y) = \sqrt[3]{y^2}$ è di classe C^∞ in un intorno del punto di passaggio $(t_0, y_0) = (0, 1)$, dunque possiamo applicare il teorema di esistenza ed unicità locale e dedurre che esisterà una sola soluzione per il problema di Cauchy proposto che sarà definita in un intorno del punto $t_0 = 0$ del tipo $[-\delta, \delta]$ con $\delta > 0$ che però a priori non sappiamo quanto vale.

Sappiamo risolvere esplicitamente anche questo problema e otteniamo come prima

$$y(t) = \left(\frac{1}{3}t + \frac{1}{3}C\right)^3$$

e imponendo la condizione $y(0) = 1$ otteniamo

$$1 = \left(\frac{1}{3}C\right)^3 \iff C = 3$$

da cui

$$\phi(t) = \left(\frac{1}{3}t + 1\right)^3$$

definita per $t \in \mathbb{R}$.

4.

$$\begin{cases} y' = 1 + y^2 \\ y(0) = 1. \end{cases}$$

È un'equazione autonoma, a variabili separabili. Si ha

$$f(t, y) = 1 + y^2 \in C^\infty(\mathbb{R} \times \mathbb{R})$$

dunque per il teorema di esistenza ed unicità locale, per ogni $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}$ esiste un'unica soluzione del problema di Cauchy

$$\begin{cases} y' = 1 + y^2 \\ y(t_0) = y_0. \end{cases}$$

Dunque, anche il nostro problema di Cauchy ha una e una sola soluzione che sarà definita in un intervallo $[-\delta, \delta]$ per un certo $\delta > 0$ che non possiamo prevedere *a priori*. Sappiamo però risolvere esplicitamente anche questo problema di Cauchy, poiché la famiglia delle soluzioni dell'EDO è in forma implicita

$$\arctan y = t + c, \quad c \in \mathbb{R}, \quad -\frac{\pi}{2} < t + c < \frac{\pi}{2}$$

e quindi la soluzione del problema di Cauchy è

$$\arctan \phi = t + \frac{\pi}{4}$$

cioè

$$\phi = \tan(t + \frac{\pi}{4}), \quad t \in (-\frac{3\pi}{4}, \frac{\pi}{4}).$$

Osserviamo che, nonostante la funzione $f(t, y)$ sia molto più regolare di quanto richiesto nel teorema (cioè sia addirittura $C^\infty(\mathbb{R}^2)$), ciò non basta a garantire che la soluzione sia definita su tutto $\mathbb{R}$.

5.
$$\begin{cases} y' = y^2 \\ y(0) = 1. \end{cases}$$

Anche in questo caso si ha

$$f(t, y) = y^2 \in C^\infty(\mathbb{R}^2)$$

e dunque il teorema di esistenza ed unicità locale si applica in ogni punto $(t_0, y_0) \in \mathbb{R}^2$. Tuttavia *a priori* non possiamo prevedere in quale intervallo sarà definita la soluzione. Anche in questo caso abbiamo una equazione a variabili separabili, autonoma. La famiglia delle soluzioni è

$$y = \frac{1}{-c - t}, \quad c \in \mathbb{R}$$

da cui l'unica soluzione del problema di Cauchy proposto è

$$\phi = \frac{1}{1 - t}, \quad t \in (-\infty, 1).$$

6.
$$\begin{cases} y' = y \\ y(0) = 0. \end{cases}$$

Anche qui, possiamo prevedere che esisterà una sola soluzione ed è banale verificare che è la soluzione costante $\phi(t) = 0$ per ogni $t \in \mathbb{R}$.

Veniamo al terzo risultato.

Teorema 57 (Esistenza ed unicità globale)
Sia dato il problema di Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0. \end{cases}$$

Se

1) f è definita in una striscia illimitata del piano $D = [\tau_1, \tau_2] \times \mathbb{R}$ con $[\tau_1, \tau_2]$ intervallo **limitato** *in $\mathbb{R}$;*

2) $f \in C^1(D)$;

3) $(t_0, y_0) \in (\tau_1, \tau_2) \times \mathbb{R}$;

4) è verificata almeno almeno una tra le due condizioni seguenti:

i) f è limitata su $[\tau_1, \tau_2] \times \mathbb{R}$, cioè esiste $M > 0$ tale che $|f(t,y)| \leq M$ per ogni $(t,y) \in [\tau_1, \tau_2] \times \mathbb{R}$;

ii) la derivata parziale $\frac{\partial f}{\partial y}$ è limitata su $[\tau_1, \tau_2] \times \mathbb{R}$, cioè esiste $M > 0$ tale che $\left|\frac{\partial f}{\partial y}(t,y)\right| \leq M$ per ogni $(t,y) \in [\tau_1, \tau_2] \times \mathbb{R}$;

allora esiste un'unica soluzione $\phi(t)$ del problema di Cauchy definita su $[\tau_1, \tau_2]$.

Esempi:

1. Consideriamo il problema di Cauchy

$$\begin{cases} y' = \dfrac{t}{1-y} \\ y(0) = 2. \end{cases}$$

Si ha $f(t,y) = \frac{t}{1-y} : \mathbb{R} \times (\mathbb{R} \setminus \{1\}) \longrightarrow \mathbb{R}$. Dunque, la condizione 1) del teorema non è verificata perché f non è definita su una striscia del piano del tipo $[\tau_1, \tau_2] \times \mathbb{R}$ e quindi NON possiamo applicare il teorema di esistema ed unicità globale. Tuttavia possiamo applicare il teorema di esistenza ed unicità locale perché si ha $f \in C^1(U(0,2))$ dunque sappiamo che esiste un'unica soluzione.

Ancora una volta, sappiamo risolvere esplicitamente il problema. L'integrale generale in forma implicita è

$$(1-y)^2 = -t^2 + C, \quad C \in \mathbb{R}$$

da cui la soluzione del problema di Cauchy si trova per $C = 1$ dunque

$$(1-y)^2 = -t^2 + 1$$

e quindi la soluzione è $\phi = 1 + \sqrt{-t^2+1}$ definita per $-t^2 + 1 > 0$ cioè $t \in (-1,1)$ (in $t = -1$ e $t = 1$ la funzione ϕ non è derivabile, dunque non è soluzione dell'equazione differenziale!)

2. Nell'esempio precedente

$$\begin{cases} y' = 1 + y^2 \\ y(0) = 1. \end{cases}$$

la funzione $f(t,y) = 1 + y^2$ non verifica né la condizione *i)* né la condizione *ii)* del teorema di esistenza ed unicità globale. Infatti $f(t,y) = 1 + y^2$ non è limitata in una striscia $[\tau_1, \tau_2] \times \mathbb{R}$, ma neanche $\frac{\partial f}{\partial y} = 2y$ è limitata in una striscia $[\tau_1, \tau_2] \times \mathbb{R}$. Infatti

$$\lim_{y \to +\infty} (1 + y^2) = +\infty, \quad \lim_{y \to +\infty} 2y = +\infty.$$

Dunque *a priori* non possiamo prevedere un intervallo minimo intorno al punto $t_0 = 0$ in cui la soluzione (unica, per il teorema di esistenza ed unicità locale) sarà definita. Poiché sappiamo risolvere esplicitamente questa equazione, possiamo *a posteriori* osservare che in effetti la soluzione $y = \tan\left(t + \frac{\pi}{4}\right)$ è definita solo in $(-\frac{3\pi}{4}, \frac{\pi}{4})$.

3.

$$\begin{cases} y' = t(\arctan y)^2 \\ y(0) = 1. \end{cases}$$

È una equazione a variabili separabili, non autonoma. Si ha

$$f(t,y) = t(\arctan y)^2 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$$

Inoltre $f \in C^\infty(\mathbb{R}^2)$; verifichiamo le altre ipotesi del teorema di esistenza ed unicità globale. Consideriamo una striscia in $\mathbb{R}^2$ del tipo $(\tau_1, \tau_2) \times \mathbb{R}$ con $0 \in (\tau_1, \tau_2)$ (τ_1, τ_2 per ora sono scelti in modo totalmente arbitrario; vedremo nel seguito se su τ_1 e τ_2 si devono imporre altre condizioni). Verifichiamo la condizione i) del teorema:

$$\left|t(\arctan y)^2\right| \leq \max(|\tau_1|, |\tau_2|)\left(\frac{\pi}{2}\right)^2 = M, \quad \forall (t, y) \in [\tau_1, \tau_2] \times \mathbb{R}.$$

Dunque, per il teorema di esistenza ed unicità globale, possiamo dedurre che il problema di Cauchy proposto ha un'unica soluzione $y = \phi(t) : [\tau_1, \tau_2] \longrightarrow \mathbb{R}$. Ora, possiamo ripetere la stessa argomentazione per ogni intervallo $[\tau_1, \tau_2]$ sempre più ampio, poiché la maggiorazione fatta precedente continua a valere (cambia però di volta in volta la costante M). Consideriamo ora due intervalli $[\tau_1, \tau_2] \subset [\tilde{\tau}_1, \tilde{\tau}_2]$ e le due rispettive soluzioni **dello stesso problema di Cauchy** proposto $\phi : [\tau_1, \tau_2] \longrightarrow \mathbb{R}$ e $\tilde{\phi} : [\tilde{\tau}_1, \tilde{\tau}_2] \longrightarrow \mathbb{R}$. Per l'unicità della soluzione, si ha che $\phi = \tilde{\phi}$ sull'intervallo più piccolo $[\tau_1, \tau_2]$ e quindi $\tilde{\phi}$ non è altro che un prolungamento della soluzione ϕ all'intervallo più grande $[\tilde{\tau}_1, \tilde{\tau}_2]$. In questo modo, si ricava che la soluzione (unica) risulta definita sull'unione di tutti gli intervalli $[\tau_1, \tau_2]$, cioè su tutto $\mathbb{R}$.

Osserviamo infine che non è semplice risolvere esplicitamente l'equazione proposta, dunque non si possono ricavare le stesse informazioni dedotte dal teorema dalla forma esplicita delle soluzioni come avevamo fatto per gli esempi precedenti.

4.
$$\begin{cases} y' = e^{(t^2-1)y^2} \\ y(0) = 2. \end{cases}$$

Anche in questo caso
$$f(t, y) = e^{(t^2-1)y^2} \in C^\infty(\mathbb{R} \times \mathbb{R}).$$

Vogliamo verificare se valgono le ipotesi del teorema di esistenza ed unicità globale. Se $t^2 - 1 \leq 0$ cioè $t \in [-1, 1]$ si ha per ogni $y \in \mathbb{R}$

$$\left|e^{(t^2-1)y^2}\right| \leq 1$$

dunque nella striscia $[-1, 1] \times \mathbb{R}$ è verificata la condizione i) del teorema e possiamo dedurre che l'unica soluzione $y = \phi(t)$ del problema di Cauchy sarà definita almeno su $[-1, 1]$. Possiamo tramite il teorema prevedere che l'intervallo di definizione sarà in effetti più ampio? No, poiché se $t > 1$ oppure $t < -1$ né la funzione $e^{(t^2-1)y^2}$, né la sua derivata parziale $2y(t^2-1)e^{(t^2-1)y^2}$ sono limitate se $y \in \mathbb{R}$. Di fatto, possiamo considerare il problema di Cauchy posizionato in $(1, \phi(1))$; qui vale il teorema di esistenza ed unicità locale e quindi possiamo prolungare la soluzione ϕ ad un intervallino aperto a destra di $t_0 = 1$, ma non abbiamo nessuna informazione su quanto grande sarà l'ampiamento dell'intervallo di definizione. Analogamente, si procede in $t_0 = -1$. In definitiva, la soluzione $\phi(t)$ sarà definita in un intervallo aperto contenente $[-1, 1]$. Anche in questo caso, non è possibile integrare esplicitamente l'equazione differenziale per verificare *a posteriori* l'esattezza di quanto previsto dal teorema.

L'ultimo risultato generale su un sistema di equazioni differenziali del primo ordine riguarda la sensibilità della soluzione di un problema di Cauchy a piccole variazioni sul dato iniziale.

Teorema 58 (Dipendenza continua dai dati)
Consideriamo i due problemi di Cauchy

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_1, \end{cases} \qquad \begin{cases} y' = f(t, y) \\ y(t_0) = y_2. \end{cases}$$

Se

1) f è di classe C^1 su un aperto $D \subset \mathbb{R}^2$ contenente entrambi i punti (t_0, y_1), (t_0, y_2);

2) la derivata parziale $\frac{\partial f}{\partial y}$ *è limitata su* D*, cioè esiste* $L > 0$ *tale che per ogni* $(t, y) \in D$

$$\left|\frac{\partial f}{\partial y}(t, y)\right| \leq L,$$

3) i grafici di $\phi_1(t) : I \longrightarrow \mathbb{R}$ *e* $\phi_2(t) : I \longrightarrow \mathbb{R}$*, soluzioni (uniche) dei due problemi di Cauchy, sono tutti contenuti in* D*,*

allora

$$|\phi_1(t) - \phi_2(t)| \leq |y_1 - y_2| e^{L|t-t_0|}, \quad \forall t \in I.$$

Esercizio: Si consideri l'equazione differenziale $y' = \cos y$.

1. Mostrare che per ogni dato iniziale il problema di Cauchy ha una ed una sola soluzione.
2. Mostrare che tale soluzione è $\mathcal{C}^\infty$ ed è definita su tutto $\mathbb{R}$.
3. Se φ_0 e φ_1 sono soluzioni corrispondenti ai dati iniziali $y(1) = \pi/2$ e $y(1) = \pi/3$, mostrare che $|\varphi_0(t) - \varphi_1(t)| \leq \frac{\pi}{6} e^{|t-1|}$ per ogni $t \in \mathbb{R}$.

I quattro teoremi visti permettono di stabilire in quale contesto un problema di Cauchy è un problema ben posto. Precisiamo questa nozione.

Definizione 59
Diremo che un problema (in un contesto matematico qualsiasi) è **ben posto nel senso di Hadamard** *se*

i) esiste un'unica soluzione;

ii) la soluzione dipende in modo continuo dai dati, cioè a piccole variazioni dei dati corrispondono piccole variazioni della soluzione.

I teoremi di esistenza ed unicità forniscono ipotesi sufficienti a garantire l'esistenza di una sola soluzione di un problema di Cauchy; il teorema di dipendenza continua dai dati fornisce condizioni aggiuntive che permettono di stimare per ogni t la distanza tra due soluzioni $|\phi_1(t) - \phi_2(t)|$ tramite la distanza dei punti di passaggio $|y_1 - y_2|$; chiaramente se $|t - t_0|$ è molto grande, anche la stima di $|\phi_1(t) - \phi_2(t)|$ diventa molto imprecisa (poiché $|\phi_1(t) - \phi_2(t)| \leq |y_1 - y_2| e^{L|t-t_0|}$), dunque il teorema fornisce una stima che diventa tanto peggiore, quanto più ci allontaniamo dall'istante iniziale t_0.

2.3 Sistemi lineari

2.3.1 Teoremi di esistenza, unicità e regolarità. Struttura delle soluzioni

Veniamo ora ai sistemi lineari del primo ordine in forma normale, cioè della forma

$$\underline{y}' = A(t)\underline{y} + \underline{b}(t)$$

ove $A : I \subset \mathbb{R} \longrightarrow M_{n\times n}(\mathbb{R})$ e $\underline{b} : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$ con I un intervallo di $\mathbb{R}$. Più precisamente, si tratta dei sistemi del tipo seguente:

$$\begin{cases} y_1' = a_{11}(t)y_1 + a_{12}(t)y_2 + \cdots + a_{1n}(t)y_n + b_1(t) \\ y_2' = a_{21}(t)y_1 + a_{22}(t)y_2 + \cdots + a_{2n}(t)y_n + b_2(t) \\ \vdots \\ y_n' = a_{n1}(t)y_1 + a_{n2}(t)y_2 + \cdots + a_{nn}(t)y_n + b_n(t) \end{cases}$$

cioè

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \dots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix}$$

Esempio:

$$\begin{cases} y_1' = \cos t\, y_1 + e^t\, y_2 + t \\ y_2' = 3y_1 - t^2\, y_2 \end{cases}$$

cioè

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} \cos t & e^t \\ 3 & -t^2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} t \\ 0 \end{pmatrix}.$$

Nella sezione precedente abbiamo enunciato i teoremi di esistenza, esistenza ed unicità locale e globale solo per equazioni del primo ordine. Potremmo ora enunciare i risultati nella forma più generale per un sistema del primo ordine. Tuttavia, poiché ci interesseremo soltanto di sistemi lineari, enunciamo il risultato valido solo per tali sistemi, in cui la forma della funzione $\underline{f}(t, \underline{y})$ è particolare.

Teorema 60 (Esistenza ed unicità globale; regolarità)
Siano $A(t) = [a_{i,j}(t)]_{i,j=1,\dots,n}$ e $\underline{b}(t) = [b_i(t)]_{i=1,\dots,n}$. Se a_{ij} e b_i sono funzioni di classe $C([a,b])$ (con $a < b$) per $i, j = 1, \dots, n$, allora per ogni $(t_0, \underline{y}_0) \in (a,b) \times \mathbb{R}^n$ esiste un'unica soluzione $\underline{y} = \underline{\phi}(t)$ del problema di Cauchy

$$\begin{cases} \underline{y}' = A(t)\underline{y} + \underline{b}(t) \\ \underline{y}(t_0) = \underline{y}_0 \end{cases}$$

definita su tutto $[a,b]$. Inoltre, se a_{ij} e b_i sono funzioni di classe $C^k([a,b])$ per $k \geq 1$, la soluzione $\underline{\phi}$ è di classe $C^{k+1}([a,b])$.

Denotiamo ora con $C^k(I, \mathbb{R}^n)$ lo spazio vettoriale delle funzioni di classe C^k su $I \subset \mathbb{R}$ a valori in $\mathbb{R}^n$. Fissata la matrice $A(t)$ a componenti tutte continue su un intervallo $I \subset \mathbb{R}$, l'operatore

$$\begin{array}{rcl} T : C^1(I, \mathbb{R}^n) & \longrightarrow & C(I, \mathbb{R}^n) \\ \underline{y} & \longmapsto & \underline{y}' - A(t)\underline{y}. \end{array}$$

è un operatore lineare. Vale la seguente proposizione, che abbiamo già dimostrato nell'analoga proposizione 24.

Proposizione 61 (Teorema di struttura delle soluzioni dell'equazione $\underline{\mathbf{y}}' = \mathbf{A(t)}\underline{\mathbf{y}} + \underline{\mathbf{b}}\mathbf{(t)}$)
L'insieme delle soluzioni del sistema omogeneo

$$\underline{y}' = A(t)\underline{y} \qquad (\iff \underline{y}' - A(t)\underline{y} = \underline{0})$$

è un sottospazio vettoriale di $C^1(I, \mathbb{R}^n)$ e coincide con il nucleo dell'operatore lineare T. L'insieme delle soluzioni del sistema completo

$$\underline{y}' = A(t)\underline{y} + \underline{b}(t) \qquad (\iff \underline{y}' - A(t)\underline{y} = \underline{b}(t))$$

è dato da $\underline{\phi}_0 + \operatorname{Ker} T = \{\underline{\phi} = \underline{\phi}_0 + \underline{\psi}, \quad \underline{\psi} \in \operatorname{Ker} T\}$, ove $\underline{\phi}_0$ è una qualunque soluzione (fissata) del sistema completo.

Più precisamente, vale il seguente

Teorema 62
Fissata la matrice $A(t) : I \subset \mathbb{R} \longrightarrow M_{n\times n}(\mathbb{R})$ a componenti tutte continue, lo spazio delle soluzioni del sistema omogeneo $\underline{y}' = A(t)\underline{y}$ ha dimensione n. In particolare, per ogni base $B = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ scelta in $\mathbb{R}^n$ e per ogni $t_0 \in I$ scelto, le n soluzioni $\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ degli n problemi di Cauchy:

$$\begin{cases} \underline{y}' = A(t)\underline{y} \\ \underline{y}(t_0) = \underline{v}_j, \quad j = 1, \dots, n \end{cases}$$

costituiscono una base dello spazio delle soluzioni.

Dimostrazione: Dobbiamo mostrare che la famiglia $\{\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n\}$ è generatrice ed indipendente.

Cominciamo a mostrare che è generatrice, cioè che presa una qualsiasi soluzione $\underline{\phi} : I \longrightarrow \mathbb{R}^n$ dell'equazione omogenea $\underline{y}' = A(t)\underline{y}$ esistono n coefficienti reali $\alpha_1, \alpha_2, \dots, \alpha_n$ tali che $\underline{\phi} = \alpha_1\underline{\phi}_1 + \alpha_2\underline{\phi}_2 + \dots + \alpha_n\underline{\phi}_n$.

Sia dunque $\underline{\phi}$ una soluzione e sia $\underline{\phi}(t_0) = \underline{v} \in \mathbb{R}^n$; poiché B è una base di $\mathbb{R}^n$ esistono $\alpha_1, \alpha_2, \dots, \alpha_n$ (definiti in modo unico) tali che $\underline{v} = \alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \dots + \alpha_n\underline{v}_n$. Consideriamo ora la funzione

$$\tilde{\underline{\phi}} = \alpha_1\underline{\phi}_1 + \alpha_2\underline{\phi}_2 + \dots + \alpha_n\underline{\phi}_n.$$

Per la linearità, $\tilde{\underline{\phi}}$ è soluzione dell'equazione omogenea $\underline{y}' = A(t)\underline{y}$ ed inoltre $\tilde{\underline{\phi}}(t_0) = \underline{v}$. Per l'unicità della soluzione del problema di Cauchy

$$\begin{cases} \underline{y}' = A(t)\underline{y} \\ \underline{y}(t_0) = \underline{v} \end{cases}$$

si ha $\tilde{\underline{\phi}} = \underline{\phi}$.

Mostriamo ora che $\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ sono indipendenti.

Sappiamo per ipotesi che $\underline{\phi}_1(t_0), \underline{\phi}_2(t_0), \dots, \underline{\phi}_n(t_0)$ sono vettori indipendenti in $\mathbb{R}^n$ e quindi, per la proposizione 5 le funzioni $\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ sono indipendenti come funzioni.

•

Osservazione importante: Nella dimostrazione precedente abbiamo visto che la famiglia $\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ genera lo spazio delle soluzioni partendo dall'ipotesi che $\underline{\phi}_1(t_0), \underline{\phi}_2(t_0), \dots, \underline{\phi}_n(t_0)$ genera $\mathbb{R}^n$. Più precisamente, abbiamo mostrato che, poiché

$$\underline{\phi}(t_0) = \alpha_1\underline{\phi}_1(t_0) + \alpha_2\underline{\phi}_2(t_0) + \dots + \alpha_n\underline{\phi}_n(t_0)$$

(che abbiamo riscritto

$$\underline{v} = \alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \dots + \alpha_n\underline{v}_n),$$

allora anche

$$\underline{\phi} = \alpha_1\underline{\phi}_1 + \alpha_2\underline{\phi}_2 + \dots + \alpha_n\underline{\phi}_n.$$

In altri termini, abbiamo dedotto dalla proprietà

le funzioni $\underline{\phi}, \underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ sono dipendenti in un punto t_0

la proprietà

le funzioni $\underline{\phi}, \underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_n$ sono dipendenti come funzioni.

È importante osservare che questa deduzione, vera per le n soluzioni di un sistema di equazioni differenziali lineare omogeneo, NON è affatto vera in generale per n funzioni definite da un intervallo I a valori in $\mathbb{R}^n$!! Abbiamo già avuto modo di sottolineare questo fatto dopo la proposizione 5 riguardante la dipendenza lineare in spazi di funzioni.

Definizione 63
Una famiglia $\{\underline{\phi}_1, \underline{\phi}_2, \ldots, \underline{\phi}_n\}$ *di* n *soluzioni indipendenti del sistema omogeneo*

$$\underline{y}' = A(t)\underline{y}$$

si dice **sistema fondamentale di soluzioni** *e la funzione matriciale*

$$\Phi(t) = \Big(\underline{\phi}_1(t)\,|\,\underline{\phi}_2(t)\,|\,\ldots\,|\,\underline{\phi}_n(t)\Big) \in C^1(I, M_{n\times n}(\mathbb{R}))$$

si dice **matrice fondamentale del sistema***.*

Si osservi che un sistema fondamentale e una matrice fondamentale non sono univocamente determinati (così come non è unica una base di uno spazio vettoriale). Poiché l'integrale generale del sistema omogeneo è l'insieme delle funzioni $\underline{\phi}$ della forma

$$\underline{\phi}(t) = c_1\underline{\phi}_1(t) + c_2\underline{\phi}_2(t) + \cdots + c_n\underline{\phi}_n(t), \quad c_j \in \mathbb{R}, \quad j = 1, \ldots, n$$

possiamo anche scriverlo in forma matriciale:

$$\underline{\phi}(t) = \Phi(t)\underline{c}, \quad \underline{c} \in \mathbb{R}^n.$$

Per il teorema di struttura, l'integrale generale del sistema non omogeneo

$$\underline{y}' = A(t)\underline{y} + \underline{b}(t)$$

è l'insieme delle funzioni $\underline{\phi}$ che si scrivono come

$$\underline{\phi}(t) = \Phi(t)\underline{c} + \underline{\phi}_0(t), \quad \underline{c} \in \mathbb{R}^n,$$

ove $\underline{\phi}_0$ è una soluzione particolare del sistema non omogeneo.

Osservazione importante: occorre prestare attenzione alla posizione del vettore $\underline{c}$ delle costanti arbitrarie nell'espressione matriciale dell'integrale generale $\underline{\phi}(t) = \Phi(t)\underline{c}, \quad \underline{c} \in \mathbb{R}^n$. Infatti, poiché il prodotto righe per colonne NON è commutativo, non è possibile scambiare l'ordine dei termini.

2.3.2 Sistemi lineari omogenei a coefficienti costanti

Ci proponiamo ora di calcolare esplicitamente l'integrale generale di un **sistema lineare omogeneo a coefficienti costanti**:

$$\underline{y}' = A\underline{y} \tag{2.5}$$

ove $A \in M_{n\times n}$ è una matrice costante.

Esempio:

$$\begin{cases} y_1' = 3y_1 + 4y_2 \\ y_2' = -y_1 + 5y_2 \end{cases}$$

che possiamo riscrivere in forma matriciale

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Osserviamo che un sistema lineare omogeneo a coefficienti costanti è sempre un sistema autonomo; inoltre, grazie al teorema 60, tutte le soluzioni di tale sistema sono definite su tutto $\mathbb{R}$, sono di classe C^∞ e vale esistenza ed unicità in ogni punto.

Sappiamo che l'integrale generale del sistema avrà la forma

$$\underline{y} = \Phi(t)\underline{c}, \quad \underline{c} \in \mathbb{R}^n$$

ove Φ è una matrice fondamentale data da

$$\Phi(t) = \left(\underline{\phi}_1(t)\,|\,\underline{\phi}_2(t)\,|\,\dots\,|\,\underline{\phi}_n(t)\right)$$

ove le funzioni $\underline{\phi}_j$ sono n soluzioni indipendenti.

Se la matrice A fosse diagonale

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

il problema sarebbe semplice; infatti il sistema sarebbe "disaccoppiato", cioè del tipo

$$\begin{cases} y_1' = \lambda_1 y_1 \\ y_2' = \lambda_2 y_2 \\ \vdots \\ y_n' = \lambda_n y_n \end{cases}$$

e ciascuna equazione avrebbe come famiglia delle soluzioni

$$y_k = c_k e^{\lambda_k t}, \quad c_k \in \mathbb{R}$$

da cui l'integrale generale sarebbe

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

e una matrice fondamentale sarebbe

$$\Phi(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}.$$

Data una matrice diagonale Λ, denotiamo

$$e^{\Lambda t} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}$$

da cui

$$\Phi(t) = e^{\Lambda t}.$$

Se però A non è diagonale, la situazione è più complicata.

2.3.3 Caso di una matrice diagonalizzabile su $\mathbb{R}$

Se A non è diagonale, ma **è diagonalizzabile su** $\mathbb{R}$, allora sappiamo che esistono una matrice diagonale Λ e una matrice invertibile S tali che

$$A = S \Lambda S^{-1}$$

ove Λ è la matrice che ha sulla diagonale principale gli autovalori λ_j di A (eventualmente ripetuti, a seconda della loro molteplicità) e S è la matrice che ha come colonne una base formata da autovettori $\underline{v}_j$ relativi ai rispettivi autovalori λ_j:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \qquad S = (\underline{v}_1 \,|\, \underline{v}_2 \,|\, \dots \,|\, \underline{v}_n)\,.$$

Il risultato che permette di calcolare facilmente una matrice fondamentale $\Phi(t)$ per il sistema lineare omogeneo

$$\underline{y}' = A\underline{y}$$

anche nel caso in cui A non sia diagonale ma sia diagonalizzabile su $\mathbb{R}$ è il seguente (che non dimostriamo):

Teorema 64
Se $A = S\Lambda S^{-1}$, allora una matrice fondamentale del sistema

$$\underline{y}' = A\underline{y}$$

è

$$\Phi(t) = Se^{\Lambda t}S^{-1} = S \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix} S^{-1}.$$

Dal risultato precedente, si ha che l'integrale generale del sistema omogeneo

$$\underline{y}' = A\underline{y}$$

con A matrice diagonalizzabile su $\mathbb{R}$ è l'insieme delle funzioni della forma

$$\underline{\phi}(t) = Se^{\Lambda t}S^{-1}\underline{c}, \quad \underline{c} \in \mathbb{R}^n$$

ove Λ è la matrice diagonale degli autovalori e S è la matrice invertibile degli autovettori relativi. Poiché però S è invertibile (e quindi anche S^{-1}) possiamo scrivere l'integrale generale nella forma

$$\underline{\phi}(t) = Se^{\Lambda t}\underline{\tilde{c}}, \quad \underline{\tilde{c}} \in \mathbb{R}^n;$$

infatti, per ogni $\underline{c} \in \mathbb{R}^n$ esiste $\underline{\tilde{c}} \in \mathbb{R}^n$ tale che $S^{-1}\underline{c} = \underline{\tilde{c}}$ e viceversa, per ogni $\underline{\tilde{c}} \in \mathbb{R}^n$ esiste $\underline{c} \in \mathbb{R}^n$ tale che $\underline{\tilde{c}} = S^{-1}\underline{c}$: basta considerare $\underline{c} = S\underline{\tilde{c}}$.

In definitiva, l'integrale generale del sistema omogeneo

$$\underline{y}' = A\underline{y}$$

con A matrice diagonalizzabile su $\mathbb{R}$ è l'insieme delle funzioni della forma

$$\underline{\phi}(t) = Se^{\Lambda t}\underline{\tilde{c}}, \quad \underline{\tilde{c}} \in \mathbb{R}^n$$

e

$$Se^{\Lambda t}$$

è una matrice fondamentale del sistema. Più precisamente:

$$\underline{\phi}(t) = (\underline{v}_1 \,|\, \underline{v}_2 \,|\, \dots \,|\, \underline{v}_n) \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

cioè

$$\underline{\phi}(t) = c_1 e^{\lambda_1 t} \underline{v}_1 + c_2 e^{\lambda_2 t} \underline{v}_2 + \cdots + c_n e^{\lambda_n t} \underline{v}_n$$

Esempi:

1. $$\underline{y}' = \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix} \underline{y}.$$

   Ricerchiamo gli autovalori di $A = \left(\begin{smallmatrix} 1 & 0 \\ 3 & -2 \end{smallmatrix}\right)$; si ha

   $$\det(A - \lambda I) = (1 - \lambda)(-2 - \lambda)$$

   dunque esistono due autovalori reali distinti $\lambda_1 = 1$ e $\lambda_2 = -2$. La matrice A è quindi diagonalizzabile su $\mathbb{R}$. Ricerchiamo ora due autovettori indipendenti. Per $\lambda_1 = 1$ si ha

   $$\begin{pmatrix} 0 & 0 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   da cui un autovettore è $\underline{v}_1 = (1, 1)$. Per $\lambda_2 = -2$ si ha

   $$\begin{pmatrix} 3 & 0 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   da cui un autovettore è $\underline{v}_2 = (0, 1)$. L'integrale generale del sistema omogeneo è quindi

   $$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

   cioè

   $$\begin{cases} y_1(t) = c_1 e^t \\ y_2(t) = c_1 e^t + c_2 e^{-2t}. \end{cases}$$

   Osserviamo in particolare che per $c_1 = c_2 = 0$ si ha la soluzione costante $(y_1(t), y_2(t)) = (0, 0)$; inoltre, per $c_1 = 0$ si hanno le infinite soluzioni

   $$\begin{cases} y_1(t) = 0 \\ y_2(t) = c_2 e^{-2t}, \quad c_2 \in \mathbb{R} \end{cases}$$

   mentre se $c_2 = 0$ si hanno le infinite soluzioni seguenti:

   $$\begin{cases} y_1(t) = c_1 e^t \\ y_2(t) = c_1 e^t, \quad c_1 \in \mathbb{R}. \end{cases}$$

2. (ESERCIZIO)

   $$\begin{cases} y_1' = y_1 + 3y_2 - 2y_3 \\ y_2' = 2y_2 + 4y_3 \\ y_3' = y_2 + 2y_3. \end{cases}$$

   Scrivere dapprima il sistema in forma matriciale e determinarne poi l'integrale generale.

2.3.4 Caso di una matrice diagonalizzabile su $\mathbb{C}$ *(fatto a esercitazione)*

Se la matrice A è diagonalizzabile su $\mathbb{C}$ ma non su $\mathbb{R}$, possiamo comunque trovare un sistema fondamentale di soluzioni reali, passando però attraverso soluzioni complesse. Se $\lambda_1, \dots, \lambda_n$ sono gli autovalori complessi di A e $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ è una base di autovettori (complessi) relativi, allora come nel caso reale, un sistema fondamentale di soluzioni è facilmente determinabile ed ha la stessa forma che nel caso reale. Ricordando che

$$e^{\alpha+i\beta} = e^{\alpha}(\cos\beta + i\sin\beta)$$

l'integrale generale del sistema omogeneo

$$\underline{y}' = A\underline{y}$$

è dato dalle funzioni (complesse)

$$\underline{\phi}(t) = c_1 e^{\lambda_1 t}\underline{v}_1 + c_2 e^{\lambda_2 t}\underline{v}_2 + \cdots + c_n e^{\lambda_n t}\underline{v}_n$$

ove i coefficienti c_j sono in generale complessi e il vettore $\underline{v}_j$ è un autovettore (complesso) relativo all'autovalore λ_j.

Tuttavia noi siamo interessati solo a quelle soluzioni della famiglia precedente che siano reali. Vorremmo cioè sostituire alle soluzioni complesse che formano il sistema fondamentale precedente, soluzioni reali che continuino a formare un sistema fondamentale di soluzioni.

Supponiamo allora che $\lambda = \alpha + i\beta$ sia un autovalore complesso di A. Poiché il polinomio caratteristico ha coefficienti reali, anche il coniugato $\bar{\lambda} = \alpha - i\beta$ è autovalore. Inoltre, è possibile scegliere come autovettori relativi agli autovalori λ e $\bar{\lambda}$ due vettori complessi coniugati, nel senso che se $\underline{v} = \underline{w} + i\underline{u}$ è autovettore relativo a λ, allora $\overline{\underline{v}} = \underline{w} - i\underline{u}$ è autovettore relativo a $\bar{\lambda}$.

Poiché le funzioni (a valori complessi)

$$\begin{aligned} e^{(\alpha+i\beta)t}\underline{v} &= e^{\alpha t}(\cos\beta t + i\sin\beta t)(\underline{w} + i\underline{u}) \\ e^{(\alpha-i\beta)t}\overline{\underline{v}} &= e^{\alpha t}(\cos\beta t - i\sin\beta t)(\underline{w} - i\underline{u}) \end{aligned}$$

sono soluzioni indipendenti del sistema omogeneo, per linearità, anche le due funzioni

$$\begin{aligned} \frac{1}{2}\left(\underline{v}e^{(\alpha+i\beta)t} + \overline{\underline{v}}e^{(\alpha-i\beta)t}\right) &= e^{\alpha t}\left(\underline{w}\cos\beta t - \underline{u}\sin\beta t\right) \\ \frac{1}{2i}\left(\underline{v}e^{(\alpha+i\beta)t} - \overline{\underline{v}}e^{(\alpha-i\beta)t}\right) &= e^{\alpha t}\left(\underline{u}\cos\beta t + \underline{w}\sin\beta t\right) \end{aligned}$$

sono soluzioni del sistema, indipendenti e **reali**. Allora, ad ogni coppia di soluzioni provenienti da autovalori complessi coniugati, possiamo sempre sostituire una coppia di soluzioni reali e si può dimostrare che la famiglia di soluzioni così formata è ancora un sistema fondamentale di soluzioni. A questo punto, sceglieremo solo coefficienti c_j reali ed avremo tutte e sole le soluzioni reali.

Esempio:

$$\underline{y}' = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix}\underline{y}.$$

Determiniamo autovalori ed autovettori di $A = \left(\begin{smallmatrix} 0 & -4 \\ 1 & 0 \end{smallmatrix}\right)$:

$$\det(A - \lambda I) = (-\lambda)^4 + 4,$$

per cui esistono due autovalori complessi coniugati $\lambda_1 = 2i$ e $\lambda_2 = -2i$. Ricerchiamo due autovettori complessi; per $\lambda_1 = 2i$ si ha

$$\begin{pmatrix} -2i & -4 \\ 1 & -2i \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff -2iy_1 - 4y_2 = 0 \iff y_2 = -\frac{i}{2}y_1.$$

Un autovettore è quindi $\underline{v}_1 = (2, -i) = (2, 0) + i(0, -1)$. Per $\lambda_2 = -2i$ si ha

$$\begin{pmatrix} 2i & -4 \\ 1 & 2i \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff 2iy_1 - 4y_2 = 0 \iff y_2 = \frac{i}{2} y_1.$$

Un autovettore è quindi $\underline{v}_2 = (2, i) = (2, 0) + i(0, 1) = \overline{\underline{v}_1}$.

L'integrale generale complesso del sistema omogeneo è

$$\begin{aligned} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} &= c_1 e^{2it} \underline{v}_1 + c_2 e^{-2it} \underline{v}_2 \\ &= c_1 e^{2it} \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) + c_2 e^{-2it} \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right), \quad c_1, c_2 \in \mathbb{C} \end{aligned}$$

cioè

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = c_1 (\cos 2t + i \sin 2t) \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) + c_2 (\cos 2t - i \sin 2t) \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right), \quad c_1, c_2 \in \mathbb{C}.$$

Possiamo anche scrivere l'integrale generale come

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \alpha_1 \left[\frac{1}{2} e^{2it} \underline{v}_1 + \frac{1}{2} e^{-2it} \underline{v}_2 \right] + \alpha_2 \left[\frac{1}{2i} e^{2it} \underline{v}_1 - \frac{1}{2i} e^{-2it} \underline{v}_2 \right], \quad \alpha_1, \alpha_2 \in \mathbb{C}$$

poiché le due soluzioni scelte sono ancora indipendenti, da cui facendo i conti:

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \alpha_1 \left[\begin{pmatrix} 2 \\ 0 \end{pmatrix} \cos 2t + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sin 2t \right] + \alpha_2 \left[\begin{pmatrix} 2 \\ 0 \end{pmatrix} \sin 2t + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \cos 2t \right], \quad \alpha_1, \alpha_2 \in \mathbb{C}.$$

Poiché le due soluzioni trovate sono reali, scegliendo $\alpha_1, \alpha_2 \in \mathbb{R}$ otteniamo solo le soluzioni reali, che possiamo scrivere anche

$$\begin{cases} y_1(t) = 2\alpha_1 \cos 2t + 2\alpha_2 \sin 2t \\ y_2(t) = \alpha_1 \sin 2t - \alpha_2 \cos 2t. \end{cases}$$

Capitolo 3

Serie di Fourier

Testo consigliato: M. Bramanti, C. D. Pagani e S. Salsa, MATEMATICA, Calcolo infinitesimale e algebra lineare, Zanichelli, nuova edizione 2004

3.1 Un esempio dalla fisica: il problema della propagazione del calore in una sbarra unidimensionale

In questo capitolo vogliamo dare alcune nozioni di base sull'argomento delle serie di Fourier e per fare ciò partiamo da uno dei problemi che ne hanno motivato l'introduzione. Supponiamo di volere studiare l'evoluzione nel tempo della temperatura di una sbarra monodimensionale di lunghezza π (la lunghezza non è rilevante, qualsiasi altro numero andrebbe bene) costituita da un materiale omogeneo ed isotropo. Supponiamo inoltre di conoscere la temperatura della sbarra punto per punto all'istante $t = 0$ e supponiamo inoltre che gli estremi della sbarra siano mantenuti in ogni istante di tempo a temperatura nulla. L'equazione differenziale (alle derivate parziali) che governa l'evoluzione della temperatura in una sbarra di questo tipo è la cosiddetta *equazione del calore* che ha la seguente forma:

$$\frac{\partial u}{\partial t}(x,t) = k\frac{\partial^2 u}{\partial x^2}(x,t) \quad (x,t) \in [0,\pi] \times \mathbb{R} \tag{3.1}$$

La costante k che compare nell'equazione differenziale è una caratteristica del materiale di cui è fatta la sbarra e possiamo, per i nostri scopi non fisici, porre $k = 1$. Le condizioni iniziali e al bordo che abbiamo imposto si traducono nelle condizioni seguenti:

$$\begin{cases} u(x,0) = f(x) & x \in [0,\pi] \\ u(0,t) = u(\pi,t) = 0 & t \in \mathbb{R}. \end{cases}$$

Osservazione: Possiamo osservare che l'equazione del calore è lineare: infatti, se u_1 e u_2 sono soluzioni dell'equazione (per ora non ci interessiamo alle condizioni aggiuntive) e se $\alpha \in \mathbb{R}$ è uno scalare reale, allora anche $u_1 + u_2$ e αu_1 sono soluzioni. Dunque, se saremo in grado di determinare alcune soluzioni dell'equazione differenziale, anche ogni combinazione lineare di esse fornirà una soluzione.

Cerchiamo di risolvere l'equazione del calore; non vogliamo ora trovare tutte le soluzioni, ma solo alcune e precisamente ci chiediamo se esistano soluzioni della forma

$$u(x,t) = F(x)G(t)$$

cioè applichiamo il cosiddetto "metodo di separazione delle variabili". A priori, data una equazione differenziale alle derivate parziali, non è detto che esistano soluzioni di questo tipo e non è neanche detto che tutte le soluzioni siano di questo tipo.

Se $u(x,t) = F(x)G(t)$, allora $\frac{\partial u}{\partial t}(x,t) = F(x)G'(t)$ e $\frac{\partial^2 u}{\partial x^2}(x,t) = F''(x)G(t)$, dunque u è soluzione dell'equazione del calore se e solo se

$$F(x)G'(t) = F''(x)G(t).$$

Supponendo ora $F(x)G(t) \neq 0$, possiamo dividere per $F(x)G(t)$ entrambi i membri ed otteniamo la condizione

$$\frac{G'(t)}{G(t)} = \frac{F''(x)}{F(x)}.$$

Poiché i due membri dell'uguaglianza sono funzioni che dipendono da variabili diverse, l'unica possibilità è che

$$\frac{G'(t)}{G(t)} = \frac{F''(x)}{F(x)} = c, \quad c \in \mathbb{R}$$

dunque l'equazione differenziale alle derivate parziali si separa in due equazioni differenziali ordinarie

$$G'(t) = c\, G(t), \qquad F''(x) = c\, F(x).$$

La prima ha come integrale generale

$$G(t) = \alpha e^{ct}, \quad \alpha \in \mathbb{R}$$

mentre la seconda ha un integrale generale la cui forma varia a seconda che c sia maggiore, minore o uguale a zero. Più precisamente:

1) se $c > 0$, allora l'integrale generale dell'equazione su F è

$$F(x) = \beta_1 e^{\sqrt{c}x} + \beta_2 e^{-\sqrt{c}x}, \quad \beta_1, \beta_2 \in \mathbb{R};$$

2) se $c = 0$, allora l'integrale generale dell'equazione su F è

$$F(x) = \beta_1 + \beta_2 x, \quad \beta_1, \beta_2 \in \mathbb{R};$$

3) infine, se $c < 0$, allora l'integrale generale dell'equazione su F è

$$F(x) = \beta_1 \cos(\sqrt{|c|}x) + \beta_2 \sin(\sqrt{|c|}x), \quad \beta_1, \beta_2 \in \mathbb{R}.$$

In definitiva, abbiamo trovato tre famiglie di soluzioni dell'equazione del calore:

$$\begin{aligned}
&A) \quad u(x,t) = e^{ct}\left(c_1 e^{\sqrt{c}x} + c_2 e^{-\sqrt{c}x}\right), \quad c > 0, \quad c_1, c_2 \in \mathbb{R};\\
&B) \quad u(x,t) = c_1 + c_2 x, \quad c_1, c_2 \in \mathbb{R};\\
&C) \quad u(x,t) = e^{ct}\left(c_1 \cos(\sqrt{|c|}x) + c_2 \sin(\sqrt{|c|}x)\right), \quad c < 0, \quad c_1, c_2 \in \mathbb{R}.
\end{aligned}$$

Tornando al problema iniziale, vogliamo determinare quali (se ne esistono) tra queste soluzioni verificano anche le condizioni aggiuntive. Cominciamo con le condizioni al bordo $u(0,t) = u(\pi,t) = 0$ per ogni t.

Per le soluzioni di tipo A), le condizioni al bordo diventano

$$\begin{cases} e^{ct}(c_1 + c_2) = 0, \quad \forall t \\ e^{ct}\left(c_1 e^{\sqrt{c}\pi} + c_2 e^{-\sqrt{c}\pi}\right) = 0, \quad \forall t \end{cases}$$

e l'unica soluzione possibile è $c_1 = c_2 = 0$, da cui $u(x,t) = 0$ per ogni (x,t).

Per le soluzioni di tipo B), le condizioni al bordo diventano

$$\begin{cases} c_1 = 0 \\ c_2 \pi = 0 \end{cases}$$

e l'unica soluzione possibile è ancora $c_1 = c_2 = 0$, da cui $u(x,t) = 0$ per ogni (x,t).
Infine, per le soluzioni di tipo C), le condizioni al bordo diventano

$$\begin{cases} e^{ct} c_1 = 0, & \forall t \\ e^{ct} c_2 \sin(\sqrt{|c|}\pi) = 0, & \forall t \end{cases}$$

Ora, ci sono due possibilità: se $\sqrt{|c|} \notin \mathbb{N}$, allora $c_1 = c_2 = 0$ e quindi ancora una volta abbiamo $u(x,t) = 0$ per ogni (x,t); se però $\sqrt{|c|} = k \in \mathbb{N}$, allora $c_1 = 0$ ma c_2 può assumere qualsiasi valore. Abbiamo quindi trovato infinite soluzioni non nulle: se scegliamo $c = -k^2$, con $k \in \mathbb{N}$ e c_2 qualsiasi in $\mathbb{R}$, abbiamo le infinite soluzioni

$$u_k(x,t) = c_2 e^{-k^2 t} \sin(kx), \quad c_2 \in \mathbb{R}.$$

(Per essere precisi, dovremmo verificare che le funzioni precedenti verificano l'equazione del calore anche nei punti in cui si annullano, ma ciò si può verificare direttamente).
Ora, per ogni $n \in \mathbb{N}$ possiamo considerare $0 \leq k \leq n$ e fissare per ogni k un valore c_k della costante che avevamo denotato c_2. Grazie alla linearità, anche le funzioni

$$u_n(x,t) = \sum_{k=1}^{n} c_k e^{-k^2 t} \sin(kx)$$

sono soluzioni dell'equazione del calore, soddisfacenti le condizioni al bordo (che sono esse stesse delle condizioni lineari).
Veniamo ora alla condizione iniziale $u(x,0) = f(x)$. Si tratta ora di capire per quali dati iniziali $f(x)$ è possibile determinare opportunamente i coefficienti c_k per ottenere una soluzione con il metodo precedente. Chiaramente, poiché le funzioni precedenti verificano l'uguaglianza

$$u_n(x,0) = \sum_{k=1}^{n} c_k \sin(kx),$$

se il dato iniziale $f(x)$ ha la forma

$$f(x) = \sum_{k=1}^{n} c_k \sin(kx),$$

avremo trovato una soluzione del problema proposto poiché basta scegliere i coefficienti c_k che compaiono nell'espressione di f.

Esempio: Per $f(x) = \frac{1}{2} \sin x - \pi \sin(2x) + 2 \sin(3x)$, una soluzione (non sappiamo se è l'unica) dell'equazione del calore compatibile con le condizioni iniziali e al bordo imposte è

$$u(x,t) = \frac{1}{2} e^{-t} \sin x - \pi e^{-4t} \sin(2x) + 2e^{-9t} \sin(3x).$$

Di fatto, si può dimostrare che, scelta una qualunque successione $\{c_k\}_{k \in \mathbb{N}}$ limitata (cioè una funzione $k \mapsto c_k$ da $\mathbb{N}$ in $\mathbb{R}$ tale che $|c_k| \leq M$ per ogni $k \in \mathbb{N}$), la serie

$$u(x,t) = \sum_{k=1}^{\infty} c_k e^{-k^2 t} \sin(kx)$$

converge per ogni $t > 0$ e per ogni $x \in \mathbb{R}$ e soddisfa l'equazione del calore in ogni punto. Poiché formalmente

$$u(x,0) = \sum_{k=1}^{\infty} c_k \sin(kx)$$

la domanda ora diventa: quali funzioni $f(x) : [0, \pi] \longrightarrow \mathbb{R}$ ammettono uno sviluppo in serie del tipo

$$f(x) = \sum_{k=1}^{\infty} c_k \sin(kx)$$

e in quale senso vale l'uguaglianza? Vedremo che la risposta ci verrà fornita dalla teoria delle serie di Fourier.

3.2 Polinomi e serie trigonometriche

Definizione 65
Si dice **polinomio trigonometrico di grado (o ordine)** n *ogni funzione del tipo*

$$P_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \sin(kx))$$

ove a_k *e* b_k *sono numeri reali.*

Osservazioni:

1. P_n è un polinomio trigonometrico di grado n se e solo se $P_n \in \text{vect}\{1, \cos(kx), \sin(kx)\}_{k=1,\dots,n}$.
2. Ogni polinomio trigonometrico è una funzione di classe $C^\infty(\mathbb{R})$, periodica di periodo 2π.

Definizione 66
Si dice **serie trigonometrica** *ogni espressione del tipo*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

ove a_k *e* b_k *sono numeri reali.*

Dati dei coefficienti a_k, con $k \in \mathbb{N}$ e b_k con $k \in \mathbb{N} \setminus \{0\}$, non è affatto detto che la serie scritta precedentemente converga. Potrebbe convergere per alcuni valori di x e non convergere per altri valori. Tuttavia, se tale serie converge per ogni $x \in \mathbb{R}$, allora definisce una funzione periodica di periodo 2π, poiché risulta essere il limite delle sue somme parziali, che sono funzioni periodiche di periodo 2π. Per quanto riguarda la regolarità, nonostante le somme parziali siano sempre funzioni di classe C^∞, ciò non implica affatto che la serie (quando converge) definisca una funzione regolare.

3.2.1 Coefficienti di Fourier

Il problema che affrontiamo ora è il seguente: data una funzione $f \in C(\mathbb{R})$ periodica di periodo 2π, quando si può scrivere come somma di una serie trigonometrica?

Cominciamo con il determinare qual'è (se ne esiste una) la serie trigonometrica "candidata" a rappresentare una funzione f periodica assegnata. Osserviamo che ogni funzione continua, periodica di periodo 2π si può identificare con una funzione appartenente a $C([0, 2\pi])$, in quanto tutta l'informazione sulla funzione f è contenuta in un periodo. Le funzioni periodiche sono quindi quelle particolari funzioni in $C([0, 2\pi])$ che verificano $f(0) = f(2\pi)$. Lo spazio $C([0, 2\pi])$ è uno spazio vettoriale reale; in esso possiamo considerare il prodotto scalare seguente: per ogni coppia di funzioni f, g si pone

$$(f, g) := \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x$$

(avevamo già visto questo tipo di prodotto scalare in un esempio). Di conseguenza, possiamo definire una norma in $C([0,2\pi])$:

$$\|f\| := (f,f)^{\frac{1}{2}} = \left(\int_0^{2\pi} f(x)^2 \,\mathrm{d}x\right)^{\frac{1}{2}}$$

e quindi una distanza:

$$\mathrm{dist}(f,g) := \|f-g\| = \left(\int_0^{2\pi} (f(x)-g(x))^2 \,\mathrm{d}x\right)^{\frac{1}{2}}. \tag{3.2}$$

Osserviamo che i polinomi trigonometrici di ogni grado n appartengono allo spazio $C([0,2\pi])$; in particolare, per ogni $n \geq 0$ fissato (per $n=0$ stiamo considerando unicamente le funzioni costanti), lo spazio dei polinomi trigonometrici di grado minore o uguale a n è un sottospazio vettoriale, che denotiamo T_n e più precisamente:

$$T_n = \mathrm{vect}\{1, \cos(kx), \sin(kx)\}_{k=1,\dots,n}.$$

Abbiamo già enunciato e dimostrato la proposizione seguente nel capitolo riguardante i prodotti scalari e le basi ortonormali.

Proposizione 67
I vettori $\{1, \cos(kx), \sin(kx)\}_{k=1,\dots,n}$ sono indipendenti ed ortogonali a due a due. In particolare, la famiglia

$$B_n = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos(kx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}}\right\}_{k=1,\dots,n}$$

è una base ortonormale di T_n, lo spazio dei polinomi trigonometrici di grado minore o uguale a n.

Ora, data una funzione $f \in C([0,2\pi])$, dato $n \in \mathbb{N}$ e considerato lo spazio dei polinomi trigonometrici di grado minore o uguale ad n, che per la proposizione precedente possiamo caratterizzare tramite la base ortonormale

$$T_n = \mathrm{vect}\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos(kx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}}\right\}_{k=1,\dots,n},$$

qual'è l'elemento $S_n(f) \in T_n$ che meglio approssima f rispetto alla distanza definita in (3.2)? È quello che minimizza la distanza $\mathrm{dist}(f, P_n)$, con $P_n \in T_n$ ed è la proiezione ortogonale di f sul sottospazio T_n. Ha quindi la seguente espressione:

$$S_n(f) = \left(f, \frac{1}{\sqrt{2\pi}}\right)\frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{n}\left[\left(f, \frac{\cos(kx)}{\sqrt{\pi}}\right)\frac{\cos(kx)}{\sqrt{\pi}} + \left(f, \frac{\sin(kx)}{\sqrt{\pi}}\right)\frac{\sin(kx)}{\sqrt{\pi}}\right]. \tag{3.3}$$

Definizione 68
Data f continua su $\mathbb{R}$ e periodica di periodo 2π, i **coefficienti di Fourier di** *f sono definiti come segue:*

$$\begin{aligned}
\frac{a_0}{2} &= \left(f, \frac{1}{\sqrt{2\pi}}\right)\frac{1}{\sqrt{2\pi}} = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x,\\
a_k &= \left(f, \frac{\cos(kx)}{\sqrt{\pi}}\right)\frac{1}{\sqrt{\pi}} = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\cos(kx)\,\mathrm{d}x, \quad k \in \mathbb{N}\setminus\{0\}\\
b_k &= \left(f, \frac{\sin(kx)}{\sqrt{\pi}}\right)\frac{1}{\sqrt{\pi}} = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\sin(kx)\,\mathrm{d}x, \quad k \in \mathbb{N}\setminus\{0\}.
\end{aligned}$$

Chiameremo **serie di Fourier** *associata a f la serie trigonometrica*

$$S(f) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos(kx) + b_k\sin(kx))$$

dove a_0, a_k e b_k per $k \in \mathbb{N} \setminus \{0\}$ sono i coefficienti di Fourier e

$$S_n(f) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \sin(kx))$$

è una somma parziale n-esima della serie di Fourier.

Osservazione: Per l'argomento sul polinomio trigonometrico approssimante e per come sono definiti i coefficienti di Fourier, la somma parziale $S_n(f)$ è proprio il polinomio trigonometrico di grado minore o uguale a n che si trova a distanza minima da f (che possiamo pensare come elemento di $C([0, 2\pi])$); la distanza

$$\mathrm{dist}(f, g) = \left(\int_0^{2\pi} (f(x) - g(x))^2 \, \mathrm{d}x \right)^{\frac{1}{2}}$$

viene detta anche **scarto quadratico medio**.

Osservazioni:

1. I coefficienti di Fourier possono anche essere scritti come

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) \, \mathrm{d}x, \quad k \in \mathbb{N} \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) \, \mathrm{d}x, \quad k \in \mathbb{N} \setminus \{0\}. \end{aligned}$$

   poiché le funzioni $f(x)$, $\cos(kx)$ e $\sin(kx)$ sono periodiche di periodo 2π e quindi l'integrale su un qualunque intervallo di lunghezza 2π è costante.

2. Se f è pari su $\mathbb{R}$, allora per ogni $k \in \mathbb{N} \setminus \{0\}$ la funzione $f(x)\sin(kx)$ è dispari e quindi si ha $b_k = 0$; se f è dispari su $\mathbb{R}$, allora la funzione $f(x)\cos(kx)$ è dispari e quindi per ogni $k \in \mathbb{N}$ si ha $a_k = 0$. Dunque, se f è pari, la sua serie di Fourier diventa

$$S(f) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx)$$

   mentre se f è dispari, allora

$$S(f) = \sum_{k=1}^{\infty} b_k \sin(kx).$$

3. **Osservazione importante:** abbiamo introdotto i coefficienti di Fourier solo per funzioni continue su $\mathbb{R}$ e periodiche di periodo 2π. Tuttavia, essi possono essere definiti nello stesso modo anche se f periodica è meno regolare, poiché è sufficiente che le funzioni $f(x)$, $f(x)\cos(kx)$ e $f(x)\sin(kx)$ siano integrabili su un periodo (ad esempio, si può dimostrare che basta che la funzione f sia continua su un periodo tranne che in un numero finito di punti e sia limitata).

Esempi:

1. Consideriamo l'**onda quadra**: per $A > 0$ poniamo

$$f(x) = \begin{cases} -A & -\pi < x < 0 \\ A & 0 \le x \le \pi \end{cases}$$

e consideriamo poi il prolungamento periodico di periodo 2π a tutto $\mathbb{R}$. Si tratta di una funzione dispari, discontinua in $x = m\pi$ con $m \in \mathbb{Z}$. Per calcolare la serie di Fourier, sappiamo già che $a_k = 0$ per ogni $k \in \mathbb{N}$. Calcoliamo i coefficienti b_k, per $k \in \mathbb{N} \setminus \{0\}$:

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\sin(kx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} f(x)\,\sin(kx)\,\mathrm{d}x \\ &= \frac{2A}{\pi}\int_0^{\pi} \sin(kx)\,\mathrm{d}x = \frac{2A}{\pi}\left(-\frac{1}{k}\right)[\cos(kx)]_0^{\pi} \\ &= -\frac{2A}{\pi k}(\cos(k\pi) - \cos 0) = \begin{cases} 0 & k = 2h, \quad h \in \mathbb{N}\setminus\{0\} \\ \frac{4A}{\pi k} & k = 2h+1, \quad h \in \mathbb{N}. \end{cases} \end{aligned}$$

In definitiva, la serie di Fourier associata all'onda quadra è

$$S(f) = \frac{4A}{\pi}\sum_{h=0}^{\infty}\frac{\sin((2h+1)x)}{2h+1}.$$

2. Consideriamo ora

$$f(x) = |x|, \quad -\pi \le x \le \pi$$

e consideriamo poi il prolungamento periodico di periodo 2π a tutto $\mathbb{R}$. Si tratta di una funzione continua su $\mathbb{R}$, pari. Sappiamo quindi che i coefficienti di Fourier di tipo b_k sono tutti nulli. Calcoliamo quindi a_k per $k \in \mathbb{N}$:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} f(x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\,\mathrm{d}x = \pi$$

e per $k \in \mathbb{N} \setminus \{0\}$:

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,\cos(kx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\cos(kx)\,\mathrm{d}x \\ &= \frac{2}{\pi}\left(\left[\frac{1}{k}\sin(kx)x\right]_0^{\pi} - \frac{1}{k}\int_0^{\pi}\sin(kx)\,\mathrm{d}x\right) \\ &= -\frac{2}{k\pi}\left[-\frac{1}{k}\cos(kx)\right]_0^{\pi} = \frac{2}{k^2\pi}(\cos(k\pi) - \cos 0) = \begin{cases} 0 & k = 2h, \quad h \in \mathbb{N}\setminus\{0\} \\ -\frac{4}{k^2\pi} & k = 2h+1, \quad h \in \mathbb{N}. \end{cases} \end{aligned}$$

La serie di Fourier associata a f è quindi:

$$S(f) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{h=0}^{\infty}\frac{\cos((2h+1)x)}{(2h+1)^2}.$$

3.2.2 Convergenza puntuale della serie di Fourier

Fino ad ora non abbiamo ancora considerato la questione della convergenza della serie di Fourier associata ad una funzione data. Non è vero che ogni serie di Fourier converge puntualmente alla funzione di partenza, ad esempio si può dimostrare che per ogni $x_0 \in \mathbb{R}$ esiste una funzione periodica, continua su $\mathbb{R}$ per la quale la serie di Fourier associata non converge in x_0. Diamo allora la definizione di una classe di funzioni (periodiche) per le quali potremo poi enunciare un teorema di convergenza puntuale ovunque della serie di Fourier associata.

Definizione 69
Una funzione periodica (di periodo 2π), limitata si dice **monotona a tratti** *se l'intervallo $[-\pi, \pi]$ può essere suddiviso in un numero finito di intervalli nei quali f è monotona (crescente o decrescente).*

Esempio: La funzione $f(x) = |x|$ su $[-\pi, \pi]$ ed estesa ad $\mathbb{R}$ per periodicità è monotona a tratti. Anche la funzione $f(x) = x$ su $[-\pi, \pi]$ ed estesa ad $\mathbb{R}$ per periodicità è monotona a tratti.

Osservazione: Si può dimostrare che se f è periodica, limitata, monotona a tratti, allora per ogni $x_0 \in \mathbb{R}$ esistono il limite destro e sinistro di f in x_0 (osserviamo che f potrebbe non essere continua in x_0); denotiamo

$$f(x_0^+) = \lim_{x \to x_0^+} f(x), \quad f(x_0^-) = \lim_{x \to x_0^-} f(x).$$

Inoltre, si può dimostrare che $f(x)$, $f(x)\cos(kx)$ e $f(x)\sin(kx)$ sono integrabili su $[-\pi, \pi]$ e quindi si può definire la serie di Fourier associata a f.

Vale il teorema di convergenza seguente.

Teorema 70 (Teorema di convergenza puntuale della serie di Fourier)
Sia f una funzione periodica di periodo 2π, limitata e monotona a tratti. Allora, la serie di Fourier

$$S(f) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

converge per ogni $x \in \mathbb{R}$. Più precisamente, per ogni $x_0 \in \mathbb{R}$ si ha

$$S(f)(x_0) = \frac{f(x_0^+) + f(x_0^-)}{2}.$$

In particolare, se f è continua in x_0, allora

$$S(f)(x_0) = f(x_0)$$

cioè la serie di Fourier in x_0 converge al valore della funzione.

Esempi:

1. Se f è l'onda quadra considerata nell'esempio precedente, f è una funzione periodica di periodo 2π, monotona a tratti (anzi, costante a tratti) e continua tranne che nei punti $x_0 = m\pi$ per $m \in \mathbb{Z}$. Consideriamo ora la serie di Fourier associata

$$S(f) = \frac{4A}{\pi} \sum_{h=0}^{\infty} \frac{\sin((2h+1)x)}{2h+1};$$

è banale constatare la validità del teorema precedente nei punti di discontinuità, poiché si ha che se $x_0 = m\pi$, allora

$$\frac{f(x_0^+) + f(x_0^-)}{2} = 0$$

ed anche

$$S(f)(x_0) = \frac{4A}{\pi} \sum_{h=0}^{\infty} \frac{\sin((2h+1)x_0)}{2h+1} = 0.$$

Nei punti di continuità invece il teorema fornisce la somma della serie di Fourier: per ogni $x_0 \neq m\pi$ si ha

$$\frac{4A}{\pi} \sum_{h=0}^{\infty} \frac{\sin((2h+1)x_0)}{2h+1} = f(x_0);$$

in particolare, se $x_0 \in (0, \pi)$ si ha $f(x_0) = A$ e quindi

$$A = \frac{4A}{\pi} \sum_{h=0}^{\infty} \frac{\sin((2h+1)x_0)}{2h+1}$$

da cui

$$\frac{\pi}{4} = \sum_{h=0}^{\infty} \frac{\sin((2h+1)x_0)}{2h+1}.$$

Possiamo sostituire a x_0 un qualsiasi valore nell'intervallo $(0, \pi)$, ad esempio $x_0 = \frac{\pi}{2}$, da cui:

$$\frac{\pi}{4} = \sum_{h=0}^{\infty} \frac{\sin\left((2h+1)\frac{\pi}{2}\right)}{2h+1} = \sum_{h=0}^{\infty} \frac{(-1)^h}{2h+1}.$$

Osservazione: Grazie al teorema sulla convergenza puntuale di una serie di Fourier, abbiamo anche potuto calcolare la somma di una serie a termini di segno alterno.

2. La funzione di un esempio precedente $f(x) = |x|$ su $[-\pi, \pi]$ e poi estesa per periodicità è continua su $\mathbb{R}$ e monotona a tratti; la sua serie di Fourier è

$$S(f) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{h=0}^{\infty} \frac{\cos((2h+1)x)}{(2h+1)^2}.$$

Poiché f è continua su $\mathbb{R}$, per il teorema sulla convergenza puntuale della serie di Fourier si ha per ogni $x_0 \in \mathbb{R}$:

$$f(x_0) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{h=0}^{\infty} \frac{\cos((2h+1)x_0)}{(2h+1)^2},$$

dunque in particolare, per $x_0 \in [0, \pi]$ si ha $f(x_0) = x_0$ e quindi:

$$x_0 = \frac{\pi}{2} - \frac{4}{\pi} \sum_{h=0}^{\infty} \frac{\cos((2h+1)x_0)}{(2h+1)^2}$$

e per $x_0 = 0$ si ottiene

$$\frac{\pi}{2} = \frac{4}{\pi} \sum_{h=0}^{\infty} \frac{1}{(2h+1)^2}$$

cioè

$$\frac{\pi^2}{8} = \sum_{h=0}^{\infty} \frac{1}{(2h+1)^2}$$

che fornisce ancora la somma di una serie.

3. (ESERCIZIO) Sia $f(x) = x^2$ su $[-\pi, \pi]$ e poi estesa per periodicità a tutto $\mathbb{R}$;

   (a) calcolare la serie di Fourier associata;

   (b) sfruttando opportunamente il teorema di convergenza puntuale della serie di Fourier, dedurre il valore della serie

$$\sum_{k=1}^{\infty} \frac{1}{k^2}.$$

3.2.3 Sviluppo in serie di soli seni o di soli coseni *(non fatto a lezione - non è in programma)*

Consideriamo ora una funzione $f : [0, \pi] \longrightarrow \mathbb{R}$. Vogliamo calcolare la serie di Fourier associata e quindi dobbiamo estenderla ad una funzione periodica di periodo 2π; possiamo farlo in almeno due modi: in modo pari oppure dispari.

Attenzione: Potremmo anche estenderla ad una funzione periodica di periodo π, semplicemente ripetendola su ogni intervallo del tipo $[k\pi, (k+1)\pi]$ con $k \in \mathbb{Z}$ ma abbiamo trattato esplicitamente solo il caso di funzioni periodiche di periodo 2π e non π. Di fatto, tutto quello che abbiamo detto fino ad ora sulla serie di Fourier associata ad una funzione periodica di periodo 2π si può ripetere per una funzione periodica di periodo T qualsiasi, ma bisogna modificare opportunamente la definizione dei coefficienti di Fourier e della serie di Fourier. Non sviluppiamo i dettagli in questo senso, rimandando al testo di riferimento (Bramanti–Pagani–Salsa).

Consideriamo allora due estensioni di f:

$$\tilde{f}_p(x) = \begin{cases} f(x) & x \in [0, \pi] \\ f(-x) & x \in (-\pi, 0) \end{cases}$$

e poi estendiamo $\tilde{f}_p$ per periodicità 2π a tutto $\mathbb{R}$; $\tilde{f}_p$ è una funzione pari, mentre

$$\tilde{f}_d(x) = \begin{cases} f(x) & x \in [0, \pi] \\ -f(-x) & x \in (-\pi, 0) \end{cases}$$

estesa per periodicità 2π a tutto $\mathbb{R}$ è una funzione dispari. Possiamo allora calcolare la serie di Fourier associata a $\tilde{f}_p$ e quella associata a $\tilde{f}_d$. Poiché $\tilde{f}_p$ è pari, si sviluppa in serie di soli coseni:

$$S\tilde{f}_p = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx)$$

ove (grazie alla parità di $\tilde{f}_p(x)\cos(kx)$):

$$a_k = \frac{2}{\pi}\int_0^{\pi} \tilde{f}_p(x)\cos(kx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} f(x)\cos(kx)\,\mathrm{d}x, \quad k \in \mathbb{N}$$

mentre $\tilde{f}_d$ è dispari e si sviluppa in serie di soli seni:

$$S\tilde{f}_d = \sum_{k=1}^{\infty} b_k \sin(kx)$$

ove (grazie alla parità di $\tilde{f}_d(x)\sin(kx)$):

$$b_k = \frac{2}{\pi}\int_0^{\pi} \tilde{f}_d(x)\sin(kx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} f(x)\sin(kx)\,\mathrm{d}x, \quad k \in \mathbb{N}\setminus\{0\}.$$

Osservazione: Se $f : [0, \pi] \longrightarrow \mathbb{R}$ è continua (compresi gli estremi $x = 0$, $x = \pi$), allora la funzione $\tilde{f}_p$ è continua su tutto $\mathbb{R}$, mentre la funzione $\tilde{f}_d$ è sicuramente continua su $\mathbb{R}$ tranne che nei punti $x_0 = -\pi, 0, \pi$ e corrispondenti per periodicità; sarà continua anche i questi punti se e solo se $f(-\pi) = f(\pi) = f(0) = 0$. In ogni caso, si ha sempre:

$$\frac{\tilde{f}_d(x_0^+) + \tilde{f}_d(x_0^-)}{2} = 0, \quad \text{per } x_0 = -\pi, 0, \pi$$

Inoltre, se f è monotona a tratti, anche $\tilde{f}_d$ e $\tilde{f}_p$ sono monotone a tratti; dunque si può applicare il teorema di convergenza puntuale della serie di Fourier, ottenendo:

$$\forall x \in [0, \pi], \quad f(x) = \tilde{f}_p(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx)$$

e

$$\forall x \in (0, \pi), \quad f(x) = \tilde{f}_d(x) = \sum_{k=1}^{\infty} b_k \sin(kx).$$

In definitiva: partendo da una funzione **continua**, monotona a tratti, definita in $[0, \pi]$ e tale che $f(0) = f(\pi) = 0$ è possibile costruire uno sviluppo in serie di soli seni oppure di soli coseni, convergente per ogni $x \in [0, \pi]$.

3.3 Tornando al problema del calore *(non fatto a lezione - non è in programma)*

Se consideriamo allora l'equazione del calore in una sbarra unidimensionale trattato all'inizio di questo capitolo:

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) = \dfrac{\partial^2 u}{\partial x^2}(x,t), \quad (x,t) \in [0,\pi] \times \mathbb{R} \\ u(x,0) = f(x), \quad x \in [0,\pi] \\ u(0,t) = u(\pi,t) = 0, \quad t \in \mathbb{R}. \end{cases}$$

abbiamo già determinato che se $\{c_k\}_{k \in \mathbb{N}}$ è una successione limitata, allora la serie di funzioni

$$u(x,t) = \sum_{k=1}^{\infty} c_k e^{-k^2 t} \sin(kx)$$

è una soluzione dell'equazione del calore e soddisfa le condizioni al bordo. Ora, per quali funzioni $f : [0, \pi] \longrightarrow \mathbb{R}$ è possibile determinare una successione $\{c_k\}_{k \in \mathbb{N}}$ in modo che la soluzione $u(x,t)$ soddisfi anche la condizione iniziale $u(x, 0) = \sum_{k=0}^{\infty} c_k \sin(kx) = f(x)$? Si tratta di capire quali funzioni definite sull'intervallo $[0, \pi]$ si possono sviluppare in serie di soli seni. Per quanto visto nel paragrafo precedente, se f è continua, monotona a tratti e tale che $f(0) = f(\pi) = 0$, essa è sviluppabile in serie di soli seni e la successione $\{c_k\}_{k \in \mathbb{N}}$ che definisce la serie $u(x,t)$ è proprio quella dei coefficienti di Fourier $\{b_k\}$!

Esempio: Sia

$$f(x) = x(\pi - x), \quad x \in [0, \pi];$$

si tratta di una funzione continua su $[0, \pi]$, monotona a tratti e $f(0) = f(\pi) = 0$. Costruiamo la sua estensione periodica dispari:

$$\tilde{f}_d(x) = \begin{cases} x(\pi - x), & x \in [0, \pi] \\ -x(\pi - x), & x \in (-\pi, 0) \end{cases}$$

e poi estesa a tutto $\mathbb{R}$ in modo periodico 2π. Si ha che $\tilde{f}_d(x)$ è continua su $\mathbb{R}$, limitata e monotona a tratti. Costruiamo allora la serie di Fourier di $\tilde{f}_d$: per $k \in \mathbb{N} \setminus \{0\}$

$$\begin{aligned}
b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} \tilde{f}_d(x) \sin(kx)\,\mathrm{d}x = \frac{2}{\pi} \int_0^{\pi} \tilde{f}_d(x) \sin(kx)\,\mathrm{d}x = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx)\,\mathrm{d}x \\
&= \frac{2}{\pi} \int_0^{\pi} (\pi x - x^2) \sin(kx)\,\mathrm{d}x \\
&= \frac{2}{\pi} \left(\left[-\frac{\pi}{k} x \cos(kx) + \frac{1}{k} x^2 \cos(kx) \right]_0^{\pi} + \int_0^{\pi} \frac{\pi}{k} \cos(kx)\,\mathrm{d}x - \frac{1}{k} \int_0^{\pi} 2x \cos(kx)\,\mathrm{d}x \right) = \\
&= \frac{2}{\pi} \left(\left[\frac{\pi}{k^2} \sin(kx) \right]_0^{\pi} - \frac{2}{k} \left[\frac{1}{k} x \sin(kx) \right]_0^{\pi} + \frac{2}{k^2} \int_0^{\pi} \sin(kx)\,\mathrm{d}x \right) \\
&= \frac{2}{\pi} \left(\frac{2}{k^2} \left[-\frac{1}{k} \cos(kx) \right]_0^{\pi} \right) = \frac{2}{\pi} \frac{2}{k^2} \left(-\frac{1}{k} (-1)^k + \frac{1}{k} \right) = \begin{cases} 0 & k = 2h,\ h \in \mathbb{N} \setminus \{0\} \\ \frac{8}{\pi k^3} & k = 2h+1,\ h \in \mathbb{N}. \end{cases}
\end{aligned}$$

Dunque, la serie di Fourier associata a $\tilde{f}_d$ è

$$S\tilde{f}_d = \sum_{h=0}^{\infty} \frac{8}{\pi (2h+1)^3} \sin\left((2h+1)x\right)$$

e per il teorema di convergenza puntuale si ha per ogni $x \in [0, \pi]$:

$$x(\pi - x) = \sum_{h=0}^{\infty} \frac{8}{\pi (2h+1)^3} \sin\left((2h+1)x\right).$$

Una soluzione dell'equazione differenziale del calore con i dati assegnati è quindi:

$$u(x,t) = \sum_{h=0}^{\infty} \frac{8}{\pi (2h+1)^3} e^{-(2h+1)^2 t} \sin\left((2h+1)x\right).$$